

ISSN: 1346-535X

Proceedings of the 14th Conference of Pan-Pacific Association of Applied Linguistics

第十四回環太平洋応用言語学会

31st July – 2nd August, 2009

COOP INN Kyoto, Kyoto, Japan

環太平洋応用言語学会

Pan-Pacific Association of Applied Linguistics

Proceedings of the 14th Conference of Pan-Pacific Association of Applied Linguistics  
Edited by Pan-Pacific Association of Applied Linguistics

Bldg 16, Rm. 621. 1-6-1 Nishi-Waseda, Shinjuku-ku, Tokyo, Japan 169-8050  
<http://paaljapan.org/>

ISSN: 1346-535X

Copyrights © 2009 PAAL All rights reserved.

PRINTED IN JAPAN

**Co-Chairs:**

Michiko Nakano (Waseda University)

Bok-Myung Chang (Namseoul University)

**Conference Committee:**

Kazuharu Owada (Tokyo College of Music), Norifumi Ueda (Mejiro University), Junko Negishi (Waseda University), Masanori Oya (Waseda University), Eiichiro Tsutsui (Hiroshima International University), Yusuke Kondo (Ritsumeikan University), Kota Wachi (Shiba Junior/Senior High School), and Remi Murao (Waseda University)

Proceedings of the 14th Conference of Pan-Pacific Association of Applied Linguistics

第十四回環太平洋応用言語学会

31st July – 2nd August, 2009

COOP INN Kyoto, Kyoto, Japan

環太平洋応用言語学会

Pan-Pacific Association of Applied Linguistics



## Preface

We feel honored that we can hold the 14th Conference of Pan-Pacific Association of Applied Linguistics in Kyoto. The theme of the conference is 'Fundamental Issues in Applied Linguistics' which can cover any controversial topics in our profession. We have received more than 100 papers. We had a screening meeting on the 9th of May. The acceptance notifications were immediately sent by Professor Yusuke Kondo, Ritsumeikan University. This year, the length of Proceeding papers was double-columned 4-6 pages, which stands for 8-12 pages in the standard format. The organizing committee accepted, however, papers in double-columned 2-10 pages. The deadline for the Proceeding papers was on the 30th of June.

It is a great honor to receive Professor Rod Ellis as the Keynote speaker for this conference. Research interests of Professor Rod Ellis covers Second Language Acquisition Researches, in particular, analysis of Learner Language, Task-based Instruction Researches, etc. His books are widely studied by professionals in Applied Linguistics. Although his achievements are outstanding and too busy to spend time with us, he was kind enough to give us Workshop as well.

The 14th Conference of PAAL is funded in part by Grant-in-Aid for Scientific Research (B) 20320085, printing of this proceedings, travel fees, accommodation and honorarium for Professor Rod Ellis. Since due to the space, the acknowledgements were omitted in our papers, we would like to acknowledge that researches among graduate students and former graduate students at Japan-side of PAAL (specifically at Waseda) were made possible by Grant-in-Aid for Scientific Research (B) 20320085.

Finally, as a convention chair, I would like to thank our organizing committee members:

Yusuke Kondo, Norifumi Ueda, Junko Negishi, Eiichiro Tsutsui, Kazuharu Owada, Kota Wachi, Remi Murao, and Masanori Oya.

Michiko Nakano, Waseda University  
Bok-Myung Chang, Numseoule University

# Contents

A-1	Relative Clause Acquisition and EFL Learners Supakorn Phoocharoensil (Language Institute Thammasat University)	1
A-2	Intralexemic Sense Development of 'of': Evidence from the First Language Acquisition Jong Sup Jun (Hankuk University of Foreign Studies)	5
A-3	L2 Acquisition of Unaccusative Verbs by Japanese and Korean Learners of English Kenichi Yamakawa (Yasuda Women's University), Naoki Sugino (Ritsumeikan University), Yuko Shimizu (Ritsumeikan University), Michiko Nakano (Waseda University), Hiromasa Ohba (Joetsu Univeristy of Education)	9
A-4	Aspectual Use of the Negative Forms: A contrastive analysis of Japanese and Turkish Aydin Ozbek ((Canakkale Onsekiz Mart University)	15
B-1	National Policy and the Ideology of English: The Case of Taiwan Han-Yi Lin (The Overseas Chinese Institute of Technology)	19
B-2	Student Presentation as a Means of Learning English Lee Eun-pyo (College of Medicine, Eulji University)	25
C-1	Seeking a Culturally Appropriate Classroom Model for Japan William Baber (Ritsumeikan University)	29
C-2	Construction and implementation of automatic L2 speech evaluation system Yusuke Kondo (Ritsumeikan University) and Michiko Nakano (Waseda University)	33
D-1	"Lion Heart's Legacy": Personification of the Japanese Prime Minister in Cabinet E-mail Magazines Riikka Lansisalmi (Leiden University)	39
D-2	Movie Reviews in Japanese and English Newspapers Published in Japan: Rhetorical Preferences in Each Discourse Community Tama Kumamoto (Nagoya University of Foreign Studies)	45
E-1	Teaching Business Japanese of Distant Course through Internet Technology: Importance of Teacher Training Remi Kakiyama (Sophia University) and Yoko Okita (Tokyo Medical and Dental University)	51
E-2	Intercultural Communication through Distance Learning Among Korean and Japanese University Students Bok-Myung Chang (Namseoul University)	55
E-3	University Students' Recognition and Use of Strings in L2 Writing Victoria Muehleisen and Sean Wray (Waseda University)	59
F-1	Links Between Research on L2 Phonological Variation and Pedagogy of L2 Phonology: A Study of Phonological Variation in the English Interlanguage of Cantonese Speakers Yuet Hung Cecilia Chan (City University of Hong Kong)	65

F-2	Raising Awareness of English Prosody among Thai University Students Soisithorn Isarankura (Dhurakij Pundit University)	71
G-1	Has the English Ability of the Japanese Improved?-The National Curriculum Standards and Embarrassment for English Teachers and Students in Japan Katsura Yuasa (Kyoto Prefectural Kizu High School)	77
G-2	Epistemic Modality by Korean University Students: Perception and Production Hera Chu (Yong In University)	81
I-1	Analysis of Listening Test Items in Korean CSAT Haeng Jung (Honam University)	87
I-2	The Development and Validation of the In-house Can-do Statements for Required Writing Courses Kahoko Matsumoto (Tokai University)	91
J-1	Evaluating a Newly Developed Index for Readability Measurement of Japanese EFL Textbooks1 Toshiaki Ozasa and George R S Weir (Fukuyama Heisei University, University of Strathclyde)	95
J-2	Verb form usage in Japanese EFL Texts George R S Weir and Toshiaki Ozasa (University of Strathclyde, Fukuyama Heisei University)	101
J-3	Conjunctive Adverbials in English Academic Writing by Chinese Speakers: A Corpus Approach Tung-Yu Kao and Li-Mei Chen (National Cheng Kung University)	105
K-1	An Acoustic Study of English Language Learning Transfer from Japanese: a case study in CNN English Conversation Yi-jun Liu (National Cheng Kung University)	109
K-2	The effects of conference participation on communication apprehension Rieko Matsuoka (National College of Nursing)	115
K-3	Taiwanese College Students' Perceptions of Plagiarism: Is it a Cultural Issue? Shih-Chieh Chien(National Sun Yat-sen University)	117
L-1	Importance of Teaching Gender-neutral Language in EFL Classrooms: Students' Survey Results Chiyo Myojin (Kochi University of Technology)	121
L-2	Metaphorical use in lay and expert cancer discourse and conception of cancer in Thai culture Worawanna Petchkij (Chulalongkorn University and the Royal Golden Jubilee, Thailand Research Fund)	125
L-3	Emotions, Visual Rhetoric, and Pragmatic Inferencing in Campaigning Discourse Vincent Tao-Hsun Chang (National Chengchi University)	129

M-1	Developing ESP Course Material for International Economics Students	Tony Cripps (Ritsumeikan University)	135
M-2	Effects of Review Activities on EFL Learning	Hsiao-fang Cheng and Nina Chiulan Lin (Takming University of Science and Technology)	139
N-1	A study on the stress experienced by native English teachers in Korea	Jongbok Lee (Mokwon University)	145
N-2	Partitive-self constructions: Lexico-grammatical resources for constructing sociocultural and individual aspects of the self	Kam-yiu S. Pang (University of Macau)	149
O-1	Extremes of L2 Learning: Neurolinguistic Evidence	Loraine K. Obler (The Graduate Center of The City University of New York) and Elizabeth Ijalba (Queens College, CUNY)	157
O-2	An Investigation into the Prevalence of Voice Strain in Chinese University Teachers	Zhou Gang and Niu Xiaochun (Dalian University of Technology)	163
O-3	Memory Strategy Instruction, Contextual Learning and Vocabulary Recall	Somayeh Kazemi Tari (Science and Research University)	169
P-1	Do Visual Supports Enhance EFL Listeners' Self-confidence? An Empirical Study	Hsiao-fang Cheng (National United University)	175
P-2	SEL-Hi: Current and Future English in Japan	Hideyuki Kumaki (Waseda University)	179
P-3	Bilingual Education: What can Malaysia learn from the US Experience?	Juliana Binti Othman (University of Malaya)	185
Q-1	Suggestion for Teaching English Based on the Concept of Internalization	Naomi Kakihara (Hosei University)	191
Q-2	The Development and Implementation of Task-based Writing Performance Assessment for Japanese Learners of English: (3) Main Experiment 1	Yoshihito Sugita (Yamanashi Prefectural University)	195
R-1	Incorporating Information Gap Activities in EFL Elementary Classrooms	Chiou-hui Chou (National Hsinchu University of Education)	201
R-2	Greetings and EFL Teachers	James Martin Ronald (Hiroshima Shudo University) and Timothy Floyd Hawthorne (Hiroshima International University)	207
S-1	Interaction of Native and Target Phonology in Learner Speech: A Study of Korean English	Yunju Heo, Jin Hyun Kim and Jong-mi Kim (Kangwon National University)	213
S-2	A corpus study of passive unaccusative verbs in L2 English produced by advanced Japanese learners	Kazuharu Owada (Tokyo College of Music), Victoria Muehleisen and Hajime Tsubaki (Waseda University)	219

S-3	A corpus analysis of intransitive verbs used in junior high school English textbooks in Japan	Kazuharu Owada (Tokyo College of Music), Hajime Tsubaki, and Daisuke Yoshizaki (Waseda University)	225
S-4	Effectiveness of Student Perception on Need-based Grouping	Handan Girginer (Anadolu University)	229
S-6	Elaboration of Tasks as Means to Mediate Classroom Learning and the Outside World: Analyses of Language Learning Activities in Three EFL Classrooms for Children	Hirokatsu Kawashima (Nagasaki University of Foreign Studies), Naoki Sugino (Ritsumeikan University), and Yuya Koga (Shukugawa Gakuin College)	233
S-7	A Case Study of the Development of Communication Skills through Cross-Cultural Distance Learning	Akiko Watanabe (Waseda University) and Annette Karseras Sumi (Waseda University International Cooperation)	239
S-8	Vocabulary Frequency in Japanese Entrance Examinations from 2005 to 2008	Takane Yamaguchi (Waseda University)	245
S-9	Syntactic Categorization of Ditransitive and Dative Verbs by Japanese EFL Learners	Naoki Sugino (Ritsumeikan University), Simon Fraser (Hiroshima University), Chika Ikeda (Aichi Prefectural University), Hirokatsu Kawashima (Nagasaki University of Foreign Studies), Yuya Koga (Shukugawa Gakuin College) and Naoki Sugimori (Ritsumeikan University)	251
S-10	Quantitative Analysis of Speaking in Group Oral Interaction – Range of Speaking on the Basis of CEF Assessment –	Junko Negishi (Waseda University)	255
S-11	Towards the Assessment of the Use of Prosodic Cues in Speech Recognition	Remi Murao and Michiko Nakano (Waseda University)	260
S-12	The effect of two types of feedback on Iranian EFL students' tense-marker production	Somayeh Kazemi Tari (Science and Research University)	263
S-14	Information Sharing in a Pre-hospital Care Setting – A time related corpus based approach	Keiko Tsuchiya, Miche Kawashima, Tetsuo Yukioka1, Shoichi Ota, Shiro Mishima, Jun Oda, Chikako Kawahara (Tokyo Medical University), Masaki Onishi, and Ikushi Yoda (National Institute of Advanced Industrial Science and Technology)	269
S-15	A Study of English Language Textbooks in Japan: With a focus on consistency	Koji Uenishi (Setsunan University)	275
S-16	Acquisition of New Phonemes in Learner Speech: A Study of Korean English	Eun sil Bang, Oh Seon Kweon, and Jong-mi Kim (Kangwon National Univrsity)	281
S-18	College Teacher Training for Tertiary Students Towards Becoming Effective English Teachers	Byung-Bin Im (Kongju National University)	287

S-19	Effect of Using English Storybooks for Improving Korean English Learners' Reading and Writing Skills and Their Learning Attitudinal Factors	Mun-koo Kang (Kongju National University)	291
S-20	The Development of English Passive Constructions by Korean Learners	Hyesook Park (Kunsan National University)	297
S-21	The Association between EFL Learners' Reading Motivation and Their L2 Reading Behavior	Eun-Mi Yang (Kkottongnae Hyundo University of Social Welfare)	303
S-22	A Discussion of Understandable English	Kyunghee Choi (Hanyang Women's College)	305
S-23	Development of Negation in Child Language and Learning Principles	Ho Han (Ajou University)	311
S-25	Insertion Preferred to Deletion in Learner Speech: A Study of Korean English	Gyu-jeong Han and Jong-mi Kim (Kangwon National University)	317
S-27	The Effects of Dictation Practice in English Listening Classes	Myeong-Hee Seong (Eul-Ji University) and Eun-Hee Nam (Chung-Ang University)	323
S-28	An Analysis of the Interrelationship among Learner Variables of Dictation Practice in English Listening Classes	Myeong-Hee Seong (Eul-Ji University) and Eun-Hee Nam (Chung-Ang University)	327
S-29	The Development of Namseoul-Waseda Cross-Cultural Distance Learning Project from 2005 to 2009	Bok-Myung Chang (Namseoul University)	333
S-30	A Method of Automatic Acquisition of Typed-Dependency Representation of Japanese Syntactic Structure	Masanori Oya (Waseda University)	337
S-31	How intelligible and acceptable are 'Japanese English' pronunciations?	Nozomi Kato (Tokiwa University)	341
S-32	A Case Study on Developing a Vocabulary Testing	Norifumi Ueda (Mejiro University), Eiichiro Tsutsui (Hiroshima International University), Yusuke Kondo (Ritsumeikan University) and Michiko Nakano (Waseda University)	347
S-33	Range Analysis of Cross-Cultural Distance Learning (CCDL) Reflection Work Sheet: A Pilot Study	Satoshi Yoshida, Hideyuki Kumaki and Akiko Watanabe (Waseda University)	351
S-34	Cross-Cultural Distance Learning (CCDL) and the Learner's Motivation toward the CCDL CMC Activities: A Survey on Three Types of CCDL Classes	Satoshi Yoshida and Michiko Nakano (Waseda University)	357
S-35	Assessing Japanese EFL Learner's Social Skills in Cross-Cultural Distance Learning (CCDL) Context: a Longitudinal Study among Waseda University Students	Satoshi Yoshida and Michiko Nakano (Waseda University)	363

S-36	Lesson Review Tests and CEFR Can-do Statements Michiko Nakano (Waseda University) Hikaru Sugiyama (Waseda University International Co.), Manabu Itoh (Waseda University International Co.), Yusuke Kondo (Ritsumeikan University) and Hajime Tsubaki (Waseda University)	369
S-37	English Tutorials, CEFR and ACPA Michiko Nakano (Waseda University) Kazuharu Owada (Tokyo College of Music) Eiichiro Tsutsui (Hiroshima International University) and Yusuke Kondo (Ritsumeikan University)	375
S-38	Preliminary assessment of Cross-Cultural Distance Learning (CCDL) competence using Discourse Completion Tasks (DCT) Michiko Nakano (Waseda University), Junko Negishi (Waseda University), Kazuharu Owada (Tokyo College of Music), Remi Murao (Waseda University), Masanori Oya (Waseda University) Tae Yamazaki, Miyasaka, N., and Norifumi Ueda (Waseda University)	379
S-39	Cyber course on World Englishes and ELF: Some tentative evidence Michiko Nakano and Yuko Haraguchi (Waseda University)	385
S-40	Assessment of Teaching Practice of an EFL teacher – A Working Model of a Practical Scheme Taehee Choi (Guryong Middle School)	393
S-41	A Research on Children's Motivation Eunpyo Lee (Eulji University)	397
S-42	Award-giving as a Means of Motivation and English Grades Eunpyo Lee (Eulji University)	401
T-1	Durational Analysis of the Acoustical Characteristics of English Speech by Japanese Learners based on the Contrast between the Stressed and the Unstressed Shizuka Nakamura (Waseda University)	405
T-2	The relation between familiarity rating and productive knowledge of academic words Toru Nakamura (Tokyo Metropolitan Aoyama High School)	409
T-3	A Study of the Effectiveness of the CALL Program, 'Adjective Sommelier', as a Learning Tool to Improve Learners' Analytical Approach to the Polysemous Senses of TL Adjectives Noriko Aotani (Tokai Gakuen University), Taichi Kameyama (Gifu National College of Technology), Naoki Sugino (Ritsumeikan University) and Yuko Amaya (Tokai Gakuen University)	415
T-4	Learning Strategies: A Theoretical Assumption Wakako Kobayashi (Chuo University)	419

T-5	The Development and Implementation of Task-based Writing Performance Assessment for Japanese Learners of English: (1) A Pilot Experiment Yoshihito Sugita (Yamanashi Prefectural University) and Michiko Nakano (Waseda University)	423
T-6	The Development and Implementation of Task-based Writing Performance Assessment for Japanese Learners of English: (2) How to improve rating scales Yoshihito Sugita (Yamanashi Prefectural University) and Michiko Nakano (Waseda University)	429
T-7	The Development and Implementation of Task-based Writing Performance Assessment for Japanese Learners of English: (4) Main Experiment 2 Yoshihito Sugita (Yamanashi Prefectural University) and Michiko Nakano (Waseda University)	435
T-8	The Development and Implementation of Task-based Writing Performance Assessment for Japanese Learners of English: (5) Comparison of Two Main Experiments Yoshihito Sugita (Yamanashi Prefectural University)	441
T-9	A Qualitative Look at Japanese and Native Speaker Teachers' Views on the University EFL Classroom Richard Silver, Satoko Ito, and William Baber (Ritsumeikan University)	447
T-10	Consonant clusters by Japanese learners of English Aya Kitagawa (Waseda University)	453
T-11	Using, not knowing traditional Chinese characters affect Kanji recognition by Japanese: A MEG case study Yoko Okita (Tokyo Medical and Dental University)	457
T-12	Investigating the Effects of English E-Tutoring in a Junior High Afterschool Program Hung-Hsuan Kao (Jinwen University of Science & Technology)	461
T-13	The Acquisition of Restrictive Relative Clauses by Japanese and Korean Learners of English Hiromasa Ohba (Joetsu University of Education), Kenichi Yamakawa (Yasuda Women's University), Naoki Sugino (Ritsumeikan University), Yuko Shimizu (Ritsumeikan University), Michiko Nakano (Waseda University)	467
T-14	A Study on Learner Factors in Acquiring English Pronunciation Yuko Tominaga (University of Tokyo)	471
T-15	Examination of Foreign Language Anxiety Constructs between Japanese and Korean College Students Learning English Nami Iwaki (Nagoya University) and Hyun Jin Kim (Cheongju National University of Education)	475
T-16	Supporting and Assessing L2 Learners' Self-Regulated Learning Eiichiro Tsutsui (Hiroshima International University), Kazuharu Owada (Tokyo College of Music), Norifumi Ueda (Mejiro University) and Michiko Nakano (Waseda University)	479

T-17	An analysis of basic verbs in Japanese junior and senior high school textbooks	
	Norifumi Ueda (Mejiro University) and Kota Wachi (Shiba Junior/Senior High School)	483
T-18	Listening Comprehension Strategies of Turkish EFL Students and the Effects of Proficiency Level on Strategy Use	
	Funda Gercek (Anadolu University)	487
S-26	Students' perception of intercultural communication through new distance learning model	
	Bok-Myung Chang and Eunjung Bang (Namseoul University)	491

## Time Table

### Day 1: 31st of July

09:00-09:30	Registration			
09:30-10:00	Opening Address			
10:00-12:00	Session A	Session B		
	A-1	B-1		
	A-2	B-2		
	A-3			
A-4				
12:00-13:30	Break			
13:30-15:00	Session C	Session D		
	C-1	D-1		
	C-2	D-2		
		D-3		
15:00-15:30	Break		Poster Session I (14:00-16:00)	
15:30-17:00	Session E	Session F		
	E-1	F-1		
	E-2	F-2		
	E-3			
18:00-19:30	Board Meeting			

### Day 2: 1st of August

	Room 1	Room 2	Room 3
9:00-10:30	Session G	Session H	Undergraduate session
	G-1	H-1	
	G-2	H-2 H-3	
10:30-11:00	Break		
11:00-12:00	Keynote speech		
12:00-13:30	Break		
13:30-15:00	Session I	Session J	
	I-1	J-1	
	I-2	J-2 J-3	
15:00-15:30	Break		
15:30-17:00	Workshop by Prof. Rod Ellis		
19:30-21:30	Dinner at RIHGA ROYAL HOTEL KYOTO		

**Day 3: 2nd of August**

	Room 1	Room 2	Room 3
9:00-10:30	Session K	Session L	
	K-1	L-1	
	K-2	L-2	
	K-3	L-3	
10:30-11:00	Break		
11:00-12:00	Session M	Session N	
	M-1	N-1	
	M-2	N-2	
12:00-13:30	Break		
13:30-15:00	Session O	Session P	Poster Session II (13:30-15:30)
	O-1	P-1	
	O-2	P-2	
	O-3	P-3	
15:00-15:30	Break		
13:30-15:00	Session Q	Session R	
	Q-1	R-1	
	Q-2	R-2	
16:30-17:00	Closing Address		



word for persons as well as things (Swan, 2005), as in (7) and (8). An omission of relativizer (zero or '·') is allowed when a relative marker serves as a RC object, as in (3) and (6).

- (3) A teacher *who/whom/that/·* every student respects is smiling now.
- (4) A teacher *who/that* like syntax is keen on teaching grammar.
- (5) The dogs *which/that* is sleeping near me is Alex's.
- (6) The car *which/that/·* she purchased is very expensive.
- (7) I know the boy *whose* bicycle was stolen.
- (8) The house *whose* kitchen had been repaired was sold.

## 2 The Noun Phrase Accessibility Hierarchy (NPAH)

*The Noun Phrase Accessibility Hierarchy* (NPAH) resulted from Keenan and Comrie's (1977) considerable effort to find a language universal that explains how RCs are acquired in first language acquisition. Having exhaustively studied fifty languages across the world, they proposed a world-famous hierarchy, NPAH, as outlined below: SU > DO > IO > OBL > GEN > OCOMP<sup>1</sup>

- (9) That's the man [who ran away]. (SU)
- (10) That's the man [whom I saw yesterday]. (DO)
- (11) That's the man [to whom I gave the letter]. (IO)
- (12) That's the man [whom I was talking about]. (OBL)
- (13) That's the man [whose sister I know]. (GEN)
- (14) That's the man [whom I am taller than]. (OCOMP)

(Keenan & Comrie, 1977)

According to the NPAH, two claims are worth being discussed. First, it is assumed that any language that has relativization can relativize on subjects. In addition, predictions can be made such that if a language has a RC type x, then it will also have any RC type to the left of x. For instance, Thai is found to allow relativization on IO, This means Thai is claimed to permit relativization on DO and SU as well. It is unlikely for relativization possibility to skip any type of RC along the way from x to SU.

As regards second language acquisition of RCs, the NPAH can also be used to predict the order of acquisition in such a way that the left RC types, which are less marked, tend to be acquired before those on the right. It is also proposed that resumptive pronouns<sup>2</sup> are more likely to be used in marked RC types on the NPAH (Keenan & Comrie,

<sup>1</sup> N.B. '>' = 'is more accessible to relativization than'

SU = subject, DO = direct object, IO = indirect object, OBL = oblique, GEN = genitive, OCOMP = object of comparative

<sup>2</sup> A within-RC pronoun that has the same reference as the head noun

1977).

Despite being not allowed in English, resumptive pronouns exist in many other languages. Therefore, learners of English whose mother tongues permit pronoun retention are often predicted to commit an error when supplying this kind of pronoun in ERCs (Celce-Murcia, 1999).

Several studies found support to the NPAH (e.g. Schumann, 1978; Gass, 1979, 1982; Hawkins, 1994, 1999).

## 3 Avoidance

Avoidance is L2 learners' behavior caused by the differences between the native language and the target language. The greater the difference, the more likely they are to avoid. When learners are not sure about a grammatical structure in L2, they are inclined not to produce it on purpose (Richards, 2002). Schachter (1974) discovered avoidance of using ERCs, which are head-initial, by Chinese and Japanese speakers, whose RCs in their L1s precede heads. Although these learners seemed to make fewer errors on ERCs than Arabic and Persian speakers, whose native languages, like English, have head-initial RCs, this was because Chinese and Japanese subjects also produced much fewer ERCs than the others since the different head directions of RCs in English was not familiar to them. Kleinmann (1977) attributed such a phenomenon to the considerable difference in head directions between L1 and L2, which caused the learners difficulty and made them employ an avoidance strategy.

Gass (1980) also investigated avoidance in EFL learners' writing and found that they intentionally avoided marked RC types on the NPAH. In other words, the more marked a type of RC, the more difficulty they have. For Gass, avoidance is closely related to high degree of markedness. Marked RC types are claimed to lead learners to avoidance of those types. Zhao (1989) used a translation to compare the frequency of RCs in English and Chinese, asking Chinese EFL learners to translate ERCs into Chinese. The study revealed that the subjects avoided employing RCs in their Chinese translation because some special functions of ERCs cannot be translated into Chinese RCs. Additionally, Chinese does not have non-restrictive RCs as English does. Consequently, it was necessary for the learners to use some other constructions to represent the meaning in the source text.

According to Li (1996), which also supports Zhao (1989), Chinese ESL learners obviously avoided producing ERCs not because of their major difference in head direction between both languages. Their avoidance is claimed to result from the fact that RCs in English have certain pragmatic

functions to which Chinese speakers are not accustomed. The learners, hence, relied on some other structures which are closer to the corresponding structures in L1 so as to serve these functions.

## 4. Research Methodology

### 4.1 Data collection

The participants of the study are 45 Thai intermediate EFL learners who are first-year undergraduate students at Thammasat University, Thailand. They speak Thai as L1 and have learned English in classroom setting for at least twelve years before being recruited for the present study. The participants were asked by the researcher, who also acted as the instructor for their English foundation course, to write four descriptive 200-word essays and submit each in class every two weeks. They were also informed that their essays would not be graded according to grammatical accuracy so that they would feel free and relaxed to produce their writing which truly represents their authentic linguistic competence.

### 4.2 Data analysis

The framework of the analysis is based on the NPAH (Keenan & Comrie, 1977). As the present research aims at investigating the learners' use of ERCs according to the types proposed in the NPAH, the relative adverbs were excluded. Of the four essay topics, only one was selected here for the preliminary data analysis as a pilot study.

Table1: Distribution of relative markers in different RC types

ERC types	<i>SU</i>	<i>DO</i>	<i>OPREP</i>	<i>Total</i>	<i>%</i>
Relativizers					
<i>who</i>	42	2	-	44	68.75
<i>which</i>	2	-	-	2	3.13
<i>that</i>	4	10	1	15	23.44
<i>∅</i>	-	2	1	3	4.69
<b>total</b>	<b>48</b>	<b>14</b>	<b>2</b>	<b>64</b>	
<b>%</b>	<b>75</b>	<b>21.88</b>	<b>3.13</b>		

The findings from the pilot study indicate, as illustrated in table 1, that the most frequently produced RC type is subject (SU) with 75 %. The second most frequent type is direct object (DO) with 21.88 %, while object of preposition (OPREP) is ranked third with 3.13 %. No indirect object (IO) type was used. Further, the two most marked types, genitive (GEN) and object of comparative (OCOMP), are not found either.

When relative markers are taken in account, *who* was made with the highest frequency (68.75 %). The next most frequent marker is *that*

(23.44 %). Zero relative (*∅*) is the third (4.69 %), and the least frequent one is *which* (3.13 %).

The overall number of RCs found is 64, with 59 restrictive RCs (92.19 %) and 5 non-restrictive RCs (7.81 %).

## 5. Discussion of the findings

It is discovered that the findings of the study conform to the predictions of the NPAH. That is, SU was used with the highest frequency since it is the least marked type. DO, which is more marked than SU, occurs with the second most frequency. OPREP is the most marked type produced and it appears with the lowest frequency. The results at first seem to go against the NPAH in that IO, which is predicted to be easier to acquire than OPREP, is not found at all, while OPREP was used by the participants. However, upon careful consideration, it is not unusual to see more frequency of OPREP (3.13%) than IO (0 %). Gass (1979, 1982) explained this occurrence, showing that IO and OPREP can be merged together since they are close in nature. An IO relative requires a proposition inside the RC, as does an OPREP one. Because of this, these two RC types should be treated as one and the same, and it is not surprising to find only OPREP in place of IO especially from the small amount of the pilot study data.

Moreover, as predicted, the most marked RC types, GEN and OCOMP, are not seen perhaps because of two reasons. First, these two types are too difficult for them to use; they are probably unaware of how to use them, particularly OCOMP, which is the most marked and sometimes viewed as ungrammatical for some native speakers (Ellis, 1994). Second, even though the learners have seen some instances of GEN or OCOMP, they were probably not certain about its proper use, thus avoiding such RC types.

With respect to the relative markers, the participants, used *who* with the highest frequency (68.75 %) probably because the topic of the writing was *my best friend*. This means the learners needed to use relative words referring to persons. In this case, *who* is the most common and easiest for them to employ since *who* can occur in the subject as well as the object positions of a RC. Second to it in frequency is *that* (23.44 %), used more as objects than as subjects. In addition, only 3 tokens or 4.69 % of *that* was used with human heads, while the other (18.75%) was used to refer to non-human heads. This result is in line with Biber et al. (1999), which found more use of *that* with non-human antecedents. In addition, since *that* can occur in free variation in terms of grammatical functions in RCs and animacy of heads, the learners probably took advantage of this convenience, avoiding taking a

risk in using other relativizers, the usage of which they are not certain about.

The third most frequent is zero relativizer (4.69 %). The learners omitted relative markers in DO (3.13 %) and OPREP (1.56%). There appears no deletion of relative markers used as subjects, which is ungrammatical in standard English. The least frequently-occurring relative word is *which* (3.13 %), entirely used as subjects of RCs. This is not unexpected because the topics are related to persons and may have necessitated more use of *who* as a human relativizer.

An interesting point lies in pronoun retention (7.81 %) used by the participants despite the fact that their native language, Thai, does not allow it. The use of pronominal reflexes is claimed to be universal since they are produced by most L2 RC acquirers irrespective of their L1s. Whether learners' L1 has this kind of pronoun, they often supply such a resumptive pronoun to clarify the meaning of a RC (Gass & Selinker, 2001).

## 8 Conclusion

The present study does demonstrate that Thai EFL learners are similar to others in that their acquisition of RCs in English can be explained by the NPAH. The degree of markedness, in other words, determines their ERC acquisition in such a way that the learners tend to use RC types with low level of markedness, which are easy to acquire. To be specific, subject relatives (SU), the least marked type, occur with the highest frequency, followed by direct objects (DO) and objects of preposition (OPREP) respectively. More marked RC types were avoided. It is found that avoidance is connected with high degree of markedness. It seems that such markedness has in turn brought about typological universals, e.g. NPAH, of RC acquisition, which mostly prove true. In addition, pronoun retention is shown to be a feature universally shared by most L2 learners regardless of their native languages.

## References

- Azar, B. (2003). *Understanding and using English grammar*. New Jersey: Prentice hall.
- Biber, D, S. Johassan, Leech G, Conrad S. & Finegan E. (1999). *Longman grammar of spoken and written English*. Pearson: Singapore.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Gass, S. (1979). Language transfer and universal grammatical relations. *Language Learning*, 29(2), 327-344.
- Gass, S. (1980). An investigation of syntactic transfer in adult second language learners. In R. Scarcella & S. Krashen (Eds.), *Research in second language acquisition*. Rowley, MA: Newbury House.
- Gass, S. (1982). From theory to practice. In M. Hynes & W. Rutherford (Eds.), *On TESOL '81: Selected papers from the fifteenth Annual Conference of teachers of English to Speakers of Other Languages* (pp.129-139). Washington, DC: TESOL.
- Gass, S. & Selinker L. (2001). *Second language acquisition: An introductory course*. LEA: New Jersey.
- Keenan, E. & Comrie, B. (1977). Noun phrase accessibility hierarchy and universal grammar. *Linguistic Inquiry*, 8, 63-99.
- Kleinmann, H, (1977). Avoidance behavior in adult second language acquisition. *Language Learning*, 27, 93-107.
- Li, J. (1996). Underproduction does not necessarily mean avoidance: Investigation of underproduction using Chinese ESL learners. In Lawrence F. Bouton (Ed.). *Pragmatics and Language Learning: Monograph series*, 7, 171-187. University of Illinois at Urbana-Champaign.
- Richards, J. C. (2002). *Dictionary of language teaching and applied linguistics*. Longman: Pearson Education.
- Schachter, J. (1974). An error in error analysis. *Language Learning*. 24, 205-214.
- Sornhiran, P. (1978). *A transformational study of relative clauses in Thai*. PhD. Dissertation, University of Texas at Austin.
- Swan, M. (2005). *Practical English usage*. Oxford: Oxford University Press.
- Zhao (1989). A discourse analysis of relative clauses in Chinese and English: an error in 'An error in error analysis'. *IDEAL*, 6, 105-17.

# Intralexemic Sense Development of ‘of’: Evidence from the First Language Acquisition\*

Jong Sup Jun

Dept. of Linguistics and Cognitive Science, Hankuk University of Foreign Studies

jongsupjun@korea.com

## Abstract

Contrary to the common belief that *of* is a dummy case-assigner (Chomsky, 1981), *of* imposes strict semantic restrictions on the preceding and following nominals. For instance, in ‘A of B’, A can be a part, and B can be a whole (e.g. *the legs of the table*), but not vice versa (*\*the table of the legs*). This paper aims to spell out semantic properties of *of* from the perspective of language acquisition. To achieve this goal, I analyzed longitudinal data of eight children at age points of 3, 4, 5, and over 9 from the CHILDES (= Child Language Data Exchange System) database (MacWhinney & Snow, 1985, 1990). Using the CLAN software, I extracted all utterances of the subject children that include the uses of *of*, and analyzed the data with reference to pre-defined sense categories of *of*. Surprisingly, the well-known examples that are cited as evidence for a dummy case assigner are not found at all in the child language data. Rather, English-speaking children use a handful number of fixed expressions that contain *of*, most of which are part/quantity phrases or simple collocations. Based on this finding, I argue that *of* is a relational functor that defines a restricted number of semantic relations over a concept network.

## Keywords

Functional category, case, dummy case-assigner, language acquisition, CHILDES

## Introduction

The standard treatment of *of* is syntactic by nature; i.e. *of* lacks semantic content, and makes a functional contribution to grammar. According to Chomsky (1981), *of* makes a purely functional contribution to grammar as a dummy case assigner that is inserted into a phrase like *destruction of Rome* to satisfy the Case Filter. Here, the lexical NP *Rome* needs case as a complement of *destruction*; but *destruction* fails to assign case to *Rome* because of its [+N] feature. Instead, *Rome* gets case by *of*.

A less discussed use of *of* is *the legs of the table*. In this example, *table* is a lexical NP, and gets case

by virtue of the functional preposition *of*. But why can’t we reverse the order of the preceding and following nominals of *of*, and say *\*the table of the legs*? In A of B, A can be a part of B, but not vice versa. On the other hand, in the Product-Material A of B construction, either ‘Product of Material (e.g. *a table of wood*)’ or ‘Material of Product (*wood of the table*)’ is allowed. This shows that the preposition of imposes strict selection restrictions on the preceding and following nominals.

Then, how can we characterize the semantic properties of *of*? This paper aims to answer this important question from the perspective of language acquisition. In this paper, I analyzed longitudinal data of eight children from the CHILDES (= Child Language Data Exchange System) database (MacWhinney & Snow, 1985, 1990). The eight children in this study are carefully selected, since their speech transcripts are available for the same discourse situations at ages 3, 4, and 5. The database also provides transcripts obtained at age 9 for six of the eight children, which allows us to trace important features of language development at later years.

When we look into the actual data spoken by English-speaking children, we are surprised by the fact that children never utter nominalized verbal complement structures like *destruction of Rome*. In fact, the representative examples that are cited as evidence for a dummy case assigner do not constitute a representative portion of child language data at all. Rather, when English-speaking children utters *of*, it is used most of the time in part/quantity constructions (e.g. *all of / a cup of / two of*) or in collocation structures (e.g. *kind of / out of / a lot of*). The intralexemic sense development of *of* strongly suggests that *of* is best described as a relational functor that defines a restricted number of semantic relations over a concept network rather than as a dummy case-assigner.

---

\* This paper is an abridged version of my latest work. For a more detailed discussion of the data, see J. S. Jun (2009).

# 1 Children's Acquisition of 'of'

Child language data provide essential tools to investigate the grammatical status of *of*, in that the acquisition of a functional category has been a long-standing issue of the acquisition literature (Radford, 1990; O'Grady, 1997; Whitman, 1997).

## 1.1 Research Design

### 1.1.1 Data Source and Subjects

This paper reports the finding from the speech transcripts of eight English-speaking children in the the CHILDES (= Child Language Data Exchange System) database (MacWhinney & Snow, 1985, 1990). In particular, I focus on the HSLLD (The Home-School Study of Language and Literacy Development; Dickinson & Tabors, 2001) corpus.

In the HSLLD corpus, speech transcripts were obtained from various modalities, i.e. discourse situations. The modalities include BR (Book Reading), ER (Elicited Report), MT (Mealtime), TP (Toy Play), and so on. In this paper, eight children are selected from the HSLLD corpus, and they are coded by eight acronyms, i.e. Brt, Geo, Kar, May, Mel, Sar, Tri, and Vic. These children's speech transcripts are available for the same modalities of BR, ER, MT, and TP at ages 3, 4, and 5. The database also provides transcripts obtained at age 9 for six of the eight children, which allows us to trace important features of language development at later years.

### 1.1.2 Procedure

The speech transcripts are reorganized in terms of age groups of 3, 4, 5, and 9. Then, using the KWAL command on the CLAN program, I extracted all usages of *of* for the eight children. The uses of *of* are classified into (sub-)sense categories of *of*.

The sense categories of *of* are taken and edited from the Collins COBUILD Electronic Dictionary (Ver. 3.1). Hence, the senses in (1) make up a relational ontology that defines semantic relations among ontological objects.

#### (1) The Relational Ontology of 'of'

- a. property: *the population of this town*
- b. affiliation: *the prince of Wales*
- c. part/quantity: *the other side of the square / a cup of tea*
- d. material: *local decorations of wood and straw*
- e.  $\theta$ -roles: *the death of their father; the reduction of trade union power inside the party*
- f. explanation: *a little cry of pain*
- g. idiom/collocation: *bunch of, more of a problem, type of, kind of, out of*
- h. predicate-complement: *dream of, be proud of*

- i. after property adjectives: *It is kind of you to give me this present.*
- j. before an abstract noun: *of kindness*

In (1), the uses of '(e)  $\theta$ -roles,' '(i) after property adjectives,' and '(j) before an abstract noun' never appear in the speech transcripts of the eight children.

## 1.2 Results

Overall, children do not make use of *of* very often: at age 3, 0.8 % of utterances contain the use of *of*; at age 4, 1.6 %; at age 5, 2.0 %; and at age 9, 3.7 %. At ages 3 and 4, children's uses of *of* are restricted to part/quantity and idiom/collocation phrases, as in (2) and (3).

#### (2) Part/Quantity (Brt 3;8)

- \*MOT: you wanna [: want to] help me?
- \*CHI: want a glass of juice?
- \*MOT: yeah that'd be nice.

#### (3) Idiom/Collocation (Geo 3;8)

- \*MOT: # <who did Peter> [/] who did Peter live with?
- \*CHI: he [/] he live in a cage.
- \*MOT: Peter ?
- \*CHI: because when the wolf come out of the forest # &a and you shoot [!] the wolf.

By the time they pass their fifth birthday, children's uses of *of* become refined and expanded. At age 5, part/quantity expressions include *the end of this book*, *the top of the bridge*, etc.; and idiom/collocations include *a whole bunch of*, *out of*, etc. Besides these two representative uses of *of*, the material use comes into play.

#### (4) Material (Geo 5;3)

- \*CHI: ## make a paper airplane out of this [!] I suppose.
- \*CHI: it willn't go that # far .

Also, the affiliation and predicate-complement uses are observed in examples like *the King of the jungle*, and *She was a scared of it* respectively.

Finally, at age 9, language development reaches its mature stage, which is shown by the fact that the uses of *of* are expanded to diverse semantic types like material, predicate-complement, explanation, and property. (5) shows representative utterances for various types in the corpus.

#### (5) Semantic Types at Age 9

- a. part/quantity: *the rest of the day* (Geo 9;8)/ *one of the dangerous stunts in the whole wide world* (Sar 9;2)/ *we get this book at the day*

- before the last day of school (Tri 9;10)/ in the middle of the year (Tri 9;10 )
- b. idiom/collocation: in front of (Geo 9;8)/ instead of (Geo 9;8)/ I don't run out of space (May 9;4)
- c. material: we have to make them out of a paper bag (Tri 9;10)/ a graph out of it (Tri 9;10)
- d. predicate-complement: be afraid of (Geo 9;8)/ think of (May 9;4)/ be careful of (Tri 9;10)
- e. explanation: we can play a nice quiet game of checkers (May 9;5)/ they are doing a project of a show (May 9;5)
- f. property: in the light of the moon (Mel 9;11)

Table 1 summarizes our findings so far with exact figures.

Table 1: Intraleximic sense development of 'of'

	Age 3	Age 4	Age 5	Age 9
part/quantity	14	18	29	32
idiom/collocation	7	19	17	20
material			3	4
affiliation			1	
Predicate-complement			1	4
explanation	1			2
property				3
overgeneration	1	4		6
Total # of utterances (# of children)	2718 (8)	2568 (8)	2596 (8)	1920 (6)
Percentage of <i>of</i>	0.8 %	1.6 %	2.0 %	3.7 %

It is clear from Table 1 that children's early uses of *of* are restricted to part/quantity and idiom/collocation.

Although other uses like material, affiliation, and predicate-complement come into play, the part/quantity and idiom/collocation types still dominate the distribution of *of* at age 5. When we look into specific expressions used by each child in the part/quantity and idiom/collocation types, we are surprised to see that children use a handful number of fixed expressions like *a glass of*, *one of*, *some of*, *all of*, *one slice of*, *one piece of*, *two cans of*, *the end of*, *the top of*, etc. for part/quantity, and

*sort of*, *kind of*, *out of*, *because of*, *a lot of*, etc. for idiom/collocation.

### 1.3 Discussion

The results in 1.2 are important in at least three respects. First, children's uses of *of* are restricted to the limited domains of part/quantity and idiom/collocation as late as at age 5, and hence it is hard to believe that *of* is inserted for syntactic reasons.

Second, children make use of a handful number of fixed expressions repeatedly. Geo, for instance, is a talkative child with the MLU of 3.35 at age 3. But he utters only *one of--* and *out of--* over and over at ages 3 and 4; and *one slice of--*, *some of--*, *one of--*, *pack of--*, *out of--*, *because of--*, and *a lot of--* at age 5. The (seemingly functional) preposition *of* does not look like a functional preposition; rather, it is part of a small number of fixed expressions, and the fixed expressions denote a restricted set of semantic relations.

Finally, at age 9, children's uses of *of* are refined and expanded, but the distribution is best explained by such semantic relations as part/quantity, material, explanation, property, and so on. More importantly, the well-known examples that are cited as evidence for a dummy case assigner do not constitute a representative portion of child language data at all. We cannot find phrases like *destruction of Rome* in the child language database, nor can we find more plausible phrases analogous with *destruction of Rome*, e.g. *writer of (the book)*. Notice that *of* is frequently used in predicate-nominal structures at age 9, but all the observed instances of *of* in predicate-nominal structures are part of idiomatic verbal/adjectival constructions like *be afraid of* (Geo 9;8), *think of* (May 9;4), and *be careful of* (Tri 9;10).

From these facts follow my claims: (a) the preposition *of* has its own semantics; (b) the semantics of *of* is best described by a limited number of semantic relations, i.e. a relational ontology in (1); (c) in acquisition, *of* is used as part of a handful number of fixed expressions that denote a limited set of semantic relations at early ages; (d) at later ages, the uses of *of* are refined and expanded, but they are still semantic by nature; and (e) children are not willing to use *of* as a pure dummy case-assigner. In short, *of* is a relational functor that defines a restricted set of semantic relations over ontological objects.

## 2 Concluding remarks

Contrary to the common belief that *of* is a dummy case-assigner, children do not seem to insert *of* before a caseless nominal. Instead, children use *of*

as part of a small number of fixed phrases to express restricted semantic relations.

Then, what about the case-assigning function of *of*? We have to remember that all prepositions including *of* have the case-assigning function. In other words, the fact that an oblique argument follows *of* does not provide evidence for the wide-spread belief that *of* is a dummy case-assigner. The finding of this paper is important, in that it suggests a future direction for a case theory and any theory of grammar in general. Given that Case Filter is a crucial part of the GB theory, the claim that *of* is not purely grammatical can be a strong motivation to revise the theory of nominalization in generative linguistics.

The present study is, however, limited in the scope of data collection, in that the CHILDES transcripts mainly consist of production data, and not comprehension data. That is, we can say that English-speaking children say such and such things at ages 3, 4, and 5; but we cannot say whether the same children can comprehend certain uses of *of* even though these uses do not appear in their spontaneous speech. We have observed that children do not produce nominalized verb complement structures like *destruction of Rome* or *writer of the book*. But can they comprehend nominalized verb complement structures when they hear these phrases? If the answer to this question is *Yes*, why do they not produce such phrases? Or, in case the children seem to understand such phrases, does it necessarily mean that *of* is used as a dummy case-assigner? These questions call for a more comprehensive study that covers both the production and comprehension data. Clearly, the current study will serve as an important starting point of the future research.

## References

- Chomsky, N. (1981). *Lectures on government and binding: the Pisa lectures*. Dordrecht: Foris.
- Dickinson, D. K. & Tabors, P. O., eds. (2001). *Beginning literacy with language: young children learning at home and school*. Baltimore: Paul Brookes Publishing.
- Jun, J. S. (2009). A Relational Function in Acquisition: The Ontology of 'of'. *Korean Journal of Linguistics* 34-2, 399-421.
- MacWhinney, B. & Snow, C. E. (1985). The Child Language Data Exchange System. *Journal of Child Language* 12, 271-96.
- MacWhinney, B. & Snow, C. E. (1990). The Child Language Data Exchange System: an update. *Journal of Child Language* 17, 457-72.
- O'Grady, W. (1997). Kinung pemcwuwa sengswuk:

Hankwuke suptuk. In Y-J Kim, et al., eds., *Enehak ilonkwa hankwuke uymi thongsa kwuco suptuk I*, 155-170. Seoul: Minumsa.

Radford, A. (1990). *Syntactic theory and the acquisition of English syntax*. Oxford: Blackwell.

Whitman, J. (1997). Phyeywaypulyu pemcwuuy mokwuke suptuk. In Y-J Kim, et al., eds., *Enehak ilonkwa hankwuke uymi thongsa kwuco suptuk I*, 171-198. Seoul: Minumsa.

# L2 Acquisition of Unaccusative Verbs by Japanese and Korean Learners of English

Kenichi Yamakawa<sup>1</sup>, Naoki Sugino<sup>2</sup>, Yuko Shimizu<sup>3</sup>, Michiko Nakano<sup>4</sup> and Hiromasa Ohba<sup>5</sup>

<sup>1</sup>Faculty of Letters, Yasuda Women's University, <sup>2</sup>College of Information Science and Engineering,

Ritsumeikan University, <sup>3</sup>College of Economics, Ritsumeikan University,

<sup>4</sup>Faculty of Education and Integrated Arts and Sciences, Waseda University,

<sup>5</sup>Graduate School of Education, Joetsu University of Education

kyamakaw@yasuda-u.ac.jp, gwisno@is.ritsumeik.ac.jp,

yukos@ec.ritsumeik.ac.jp, nakanom@waseda.jp, hohba@juen.ac.jp

## Abstract

The present study focuses on the acquisition of unaccusative verbs by Japanese and Korean learners of English. Previous studies have accumulated with regard to the characteristics observed in L2 acquisition of English unaccusative verbs. It has been reported that learners with various L1 backgrounds make a distinction among intransitive verbs and that they feel greater difficulty with unaccusatives than with unergatives. However, few studies have examined and compared acquisition processes of unaccusative verbs by learners with different L1s on a common scale.

34 Japanese and 34 Korean EFL learners participated in the study. Based on the results of a test aiming to measure the general proficiency in the English grammar, they were assumed to be of equal proficiency. The 68 learners, then, were given a grammatical judgment task in order to explore their knowledge of unaccusative/unergative verbs. Their task performances were analyzed in terms of 1) the distinction between unaccusative and unergative verbs, 2) possible causes of overpassivization errors in unaccusative verbs, 3) the varying degrees of difficulty of unaccusative verbs, and 4) comparison of the two L1 groups. The results showed the overall similarity of the two L1 groups, confirming most of the findings reported in the previous research.

## Keywords

Unaccusative verb, unergative verb, Japanese EFL learner, Korean EFL learner, the Unaccusative Hypothesis, the Unaccusative Hierarchy, SLA

## 1 Unaccusative and unergative verbs

The Unaccusative Hypothesis (Perlmutter, 1978)

claims that there are two distinct classes of intransitive verbs known as unergatives and unaccusatives, which exhibit different argument structures:

- (1) a. Unergatives: [NP<sub>1</sub> [VP V] ]  
(e.g., [ Mary [VP cried] ] )
- b. Unaccusatives: [empty [VP V NP<sub>2</sub>] ]  
(e.g., [empty [VP happen the accident] ] )
- c. The accident happened 15 years ago.

The unergatives (e.g., *cry*, *dance*) originally have a logical subject (NP<sub>1</sub>) as an external argument, which bears the participant role AGENT (the instigator of an event) (1a). The unaccusatives (e.g., *fall*, *happen*), on the other hand, originally lack a logical subject, and only have a logical object (NP<sub>2</sub>) as an internal argument, which assumes the participant role THEME (a participant affected by an event) (1b). The internal argument (NP<sub>2</sub>), then, is moved to the surface subject position in order to satisfy the English syntactic requirement which stipulates that the subject position must be filled with a lexical item (1c). As a result, the grammatical subject of the unaccusative verb originates as the logical object. At first sight it becomes difficult to make the unergative/unaccusative distinction on the surface level because both have the same surface structure (i.e., NP + V).

## 2 L2 Acquisition of Unaccusative verbs

### 2.1 Errors in unaccusative verbs made by EFL learners

Many researchers have noted that L2 learners of English often extend passive formation rules to unaccusatives and produce the following types of ungrammatical sentence:

- (2) a. \*Most of people are fallen in love and

marry with somebody. (Japanese L1)  
(Zobl, 1989: 204)

- b. \*...terrorism is happened very often...  
(Korean L1) (Oshita, 2000: 312)

Unlike unaccusatives, unergatives rarely undergo this inappropriate passivization process. In addition, these nontarget sentences are observed in L2 English with various L1 backgrounds, and are particularly noticeable among intermediate/advanced learners.

## 2.2 Two major accounts of passivized unaccusative verbs

Two major accounts of the nontarget phenomena have been advanced so far: the NP movement account and the lexical causativization account. The former account points out that the argument structures of an unaccusative (1b) and a passive construction (3a) are almost identical in that both lack an external argument (logical subject) and that an internal argument (logical object) is moved to the surface subject position. One difference is that only the passive construction can take the *be* + *p.p.* marker to signal the NP movement (1c and 3b):

- (3) a. [empty [<sub>VP</sub> V NP<sub>2</sub>]]  
(e.g., [empty [<sub>VP</sub> be spoken English]])  
b. English is spoken in many countries.

However, some learners notice the similarity of the two and also apply the passive formation rules to unaccusatives in order to signal the NP movement, which results in inappropriate passives as in (2).

According to the lexical causativization account, on the other hand, L2 learners treat an unaccusative verb as transitive and temporarily create a causer of the event (4a). Then the verb is passivized with the suppression of the nonce causer (4b):

- (4) a. \*The driver happened the accident 15 years ago.  
b. \*The accident was happened 15 years ago.

## 2.3 Variation in the acquisition of unaccusatives

As we have seen, there is amounting evidence in the literature that the Unaccusative Hypothesis can be validated in L2 acquisition because L2 learners experience greater difficulty with unaccusatives than with unergatives (e.g., Hirakawa, 2003; Oshita, 1997; Shomura, 1996; Yamakawa et al., 2003). Thus, in the relevant literature, unaccusatives are treated differently from unergatives, and those two types are regarded as subclasses of intransitive verbs, each forming a single category. From this view, members belonging to the same category should display the same grammatical characteristics. However, it is also observed that L2 learners experience varying degrees of difficulty with

unaccusatives, which could question the notion of unaccusatives as a single linguistic (and psychological) category.

Unaccusative verbs can be divided into two subcategories: alternating and non-alternating verbs. The former have transitive counterparts (e.g., *break*, *open*), whereas the latter lack such corresponding transitives (e.g., *appear*, *die*). Research findings have presented contradictory results with regard to relative difficulty with the two subcategories of unaccusatives. Some studies report that L2 learners found alternating unaccusative verbs more difficult (e.g., Hirakawa, 2003; Hwang, 1999, 2001; Kwak, 2003), whereas Shomura (1996) reports the opposite. Furthermore, Ju (2000) argues his Chinese subjects showed no difference in difficulty level. Kondo (2005) claims that different degrees of difficulty depend on the learner's native language.

Varying degrees of difficulty of individual unaccusative verbs have also been reported. L2 learners avoided intransitive counterparts of such alternating verbs as *break* and *dry* (Hirakawa, 2003), *break*, *hang* and *open* (Shomura, 1996), *break*, *close* and *freeze* (Ju, 2000), and *close*, *dry* and *freeze* (Kondo, 2005). In addition, L2 learners also regarded as transitives such non-alternating verbs as *emerge*, *fall* and *occur* (Shomura, 1996) and *appear*, *disappear*, *die*, *emerge* and *vanish* (Kondo, 2005).

No & Chung (2006) also noted different rates of overpassivization errors in unaccusatives. They gave a grammaticality judgment task to 112 Korean learners of English, and found that such errors were attributable to multiple factors: English inherent factors, L1 influence, and semantic factors. More specifically, more overpassivization errors were observed when the verb was an alternating unaccusative verb (English inherent factors), when the corresponding Korean unaccusative verb had a lexical/derived passive morpheme (L1 influence), and when the subject of the unaccusative verb was inanimate (semantic factors).

Sorace (1993, 1997) notes that, based on observations of various Western languages, different unaccusative (and unergative) verbs display different syntactic behavior cross-linguistically and intra-linguistically, and proposes the Unaccusative Hierarchy, which is later called the Split Intransitive Hierarchy (Sorace & Shomura, 2001: Table 1). Based on aspectual parameters, the hierarchy identifies the notion of telic dynamic change at the core of unaccusativity and the notion of agentive nonmotional activity at the core of unergativity. The upper and lower extremes of the hierarchy consist of core unaccusative/unergative monadic verbs, which display least variation of syntactic characteristics, whereas peripheral verb types between the extremes are susceptible to

variable syntactic behavior.

Table 1: The Split Intransitive Hierarchy (Adapted from Sorace & Shomura, 2001: 250)

Unaccusative (least variation)	
↑	Change of location (e.g., <i>arrive, fall, go</i> )
	Change of condition (e.g., <i>appear, die, happen</i> )
	Continuation of a pre-existing condition (e.g., <i>last</i> )
	Existence (e.g., <i>exist, suffice</i> )
	Uncontrolled process
	[emission] (e.g., <i>rattle</i> )
↓	[involuntary reaction] (e.g., <i>tremble</i> )
	Controlled motional process (e.g., <i>swim</i> )
	Controlled nonmotional process (e.g., <i>work</i> )
Unergative (least variation)	

Sorace (1997) hypothesizes that learners' knowledge is sensitive to the relative hierarchical position of a verb, and that acquisition takes place in an order consistent with the hierarchy with core verbs first, and peripheral verbs at later stages.

Yamakawa et al. (2007) examined the plausibility of the Unaccusative Hierarchy, using grammatical judgment tasks administered to 655 Japanese EFL learners. The data was analyzed with the use of Item Response Theory, which yielded *b*-parameter (difficulty parameter) values of the test items. The comparison of the *b*-parameters of the test items did not lend support to the difficulty (or acquisition) order that the hierarchy predicted.

Kim (2006) investigated how semantic nature of intransitive verbs could be related to the L2 overpassivization errors. A total of 126 Korean high school students were given a grammatical judgment task consisting of seven semantically different intransitive verbs (four semantic types for unaccusatives and three semantic types for unergatives). The results showed that the subjects experienced different degrees of difficulty with the seven semantic types, which did not accord with what the Split Intransitive Hierarchy predicted. It was also argued that the simple dichotomy of unaccusative/unergative distinction did not suffice for the explanation of L2 overpassivization.

### 3 The study

As we have seen in Section 2, research findings have accumulated with regard to L2 acquisition of English unaccusative verbs; however, few studies have thus far compared acquisition processes of unaccusative verbs by learners with different L1s on a common scale. In this section, we will report a study conducted to Japanese and Korean learners of English.

### 3.1 Purpose

The present study aims to pursue the following three research questions (RQs):

RQ1) Do learners of English make a distinction between unaccusative and unergative verbs (the Unaccusative Hypothesis)?

RQ2) Which account (i.e., the NP movement account and the lexical causativization account) is plausible in order to explain the possible cause of overpassivization errors in unaccusatives?

RQ3) Does any variation in the acquisition of unaccusatives and unergatives exist? If so, does it support the Split Intransitive Hierarchy?

RQ4) Do learners of English with different L1s display any differences in the acquisition of unaccusative and unergative verbs, especially in terms of the three research questions above?

### 3.2 Subjects and materials

The authors had prepared several grammatical (judgment) tests and gathered data from more than 1,200 university-level Japanese learners of English to order to investigate the overall development of the English proficiency (For details, see Yamakawa et al., 2008). One of the tests, Unit Y, which required judgments of 48 sentences with unaccusatives/unergatives, and another test, MEG05 (Measure of English Grammar 2005: see Shimizu et al., 2006), which consisted of 35 discrete-point items to measure learners' general proficiency of the English grammar, had already been used. Out of the 1,200 learners, 367 learners had taken both Unit Y and MEG05.

In the present study, the same tests were given to 57 university-level Korean learners of English. Based on the test scores of MEG05, 34 Japanese learners out of 367 and 34 Korean learners out of 57 were selected for analysis. The difference between their average scores of MEG05 ( $M=20.92$ ,  $SD=5.58$  for the Japanese learners;  $M=21.74$ ,  $SD=5.08$  for the Korean learners) was not statistically significant ( $t(66) = -.636$ ,  $p=.527$ ). The two groups were assumed to be of the equal proficiency in the English grammar.

Unit Y contained six nonalternating unaccusative verbs (*appear, arrive, die, exist, fall, happen*), six unergative verbs (*cry, dance, laugh, play, sing, work*) and six alternating unaccusative verbs<sup>1</sup> (*break, burn, close, dry, grow, melt*). Each verb was placed in the following three sentence constructions: NP+V, NP+be+p.p. and NP+V+NP (Table 2).

<sup>1</sup> In the present study, the alternating verbs were not included in the analysis because Unit Y did not contain test items where alternating verbs were used in the NP+V construction.

Table 2: Sample sentences used in Unit Y

(Unaccusative verbs: e.g., *arrive*)

**NP+V (Category A):** Your letter arrived yesterday.

**\*NP+be+p.p. (Category C):** \*Because of the rain, the train was arrived late.

**\*NP+V+NP (Category E):** \*Finally the waitress arrived the salad to us.

(Unergative verbs: e.g., *cry*)

**NP+V (Category B):** Her father cried at her wedding ceremony.

**\*NP+be+p.p. (Category D):** \*He was cried when he heard of his mother's death.

**\*NP+V+NP (Category F):** \*The boy hit his little sister and cried her.

### 3.3 Procedure

The Japanese learners took Unit Y and MEG05 in the computer-scored answer sheet style on different days, whereas the Korean learners took the tests in the on-line fill-in-the-blank style on the same day. In either case, it took about 20-30 minutes to complete the two tests. The results of Unit Y, which was based on a five-point scale (1-5), were converted to 0-4 because we intended to measure the “distances” of learners’ judgments from the correct answers. All the statistical analyses were conducted with the use of ANOVA 4 (Kiriki, 2002: <http://www.hju.ac.jp/~kiriki/anova4/>).

### 3.4 Results

#### 3.4.1 Comparison of verb categories

(RQs1 and 2)

Tables 3 and 4 below show the results of Unit Y for the Japanese and Korean learners, respectively. Figure 1 shows the mean scores of the six verb categories.

Table 3: Mean acceptability judgments (Japanese)

	Unaccusatives			Unergatives		
Mean	2.31			2.52		
SD	1.26			1.28		
Category	A	C	E	B	D	F
Mean	2.63	1.81	2.49	3.07	2.14	2.36
SD	1.28	1.33	1.17	1.14	1.42	1.28

Table 4: Mean acceptability judgments (Korean)

	Unaccusatives			Unergatives		
Mean	2.53			2.98		
SD	1.52			1.33		
Category	A	C	E	B	D	F
Mean	3.07	1.70	2.83	3.55	2.49	2.89
SD	1.43	1.67	1.45	0.90	1.68	1.41

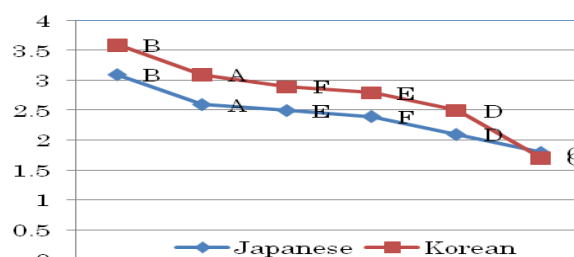


Figure 1: Mean scores of the six verb categories

A two-way repeated measures ANOVA with two independent variables (L1 and verb category) was conducted. There was a main effect both of L1 ( $F(1, 66)=7.39, p<.01$ ) and of verb category ( $F(5, 330)=42.65, p<.0000$ ). It can be claimed that the Korean learners performed better overall than the Japanese learners. A post hoc comparison test (Ryan's method) indicated the following difficulty order:  $B > A, E, F > D > C$  (“>” means there is a statistically significant difference). Both groups felt the least difficulty with unaccusatives/unergatives used in the NP+V construction (Categories A and B), and the most difficulty in the NP+be+p.p. construction (Categories C and D). They always felt more difficulty with unaccusatives than with unergatives used in the same construction except for Categories E and F. This bears out the Unaccusative Hypothesis (RQ1).

With regard to the possible cause of overpassivization of unaccusatives, it can be argued that the results above would support the NP movement account (RQ2) because Category C was significantly worse than the other categories for both L1 groups: The lexical causativization account presupposes that an overpassivization error is derived from the corresponding transitive sentence. This implies that if the lexical causativization had been at work, both Categories C and E should have been worse than the other categories and that Categories C and E should not have been significantly different; however, this was not the case here<sup>2</sup>.

#### 3.4.2 Variation in individual verbs (RQ3)

Table 5 and Figures 2-5 show that the mean scores of correct judgments of individual unaccusatives and unergatives used in three constructions.

A two-way repeated measures ANOVA with two independent variables (six individual verbs and three construction types (A-C-E and B-D-F)) was conducted for the Japanese and Korean learners

<sup>2</sup> The Japanese learners produced relatively more overpassivization errors in such unergatives as *cry*, *dance*, *laugh* than the Korean learners. This can be counter-evidence to the NP movement account because the NP movement account does not predict such overpassivization errors in unergatives.

separately. It was shown that there were significant interactions between verbs and constructions for all the four combinations: Unaccusatives for the Japanese learners, ( $F(10, 330)=6.85, p<.0000$ : Figure 2) and for the Korean learners ( $F(10, 330)=2.83, p<.005$ : Figure 3); Unergatives for the Japanese learners ( $F(10, 330)=3.74, p<.0001$ : Figure 4) and for the Korean learners ( $F(10, 330)=3.93, p<.0000$ : Figure 5). Multiple post hoc comparison tests were conducted to determine which contrasts were significant.

Table 5: Mean scores of individual verbs

M (SD)	NP+V(A)	NP+be+p.p.(C)	NP+V+NP(E)
appear	2.50 (1.40)	1.84 (1.40)	2.50 (1.11)
arrive	1.94 (1.74)	2.24 (1.56)	3.21 (1.07)
die	3.47 (0.66)	1.21 (1.20)	2.68 (1.30)
exist	2.53 (1.44)	2.18 (1.22)	2.18 (1.19)
fall	2.56 (1.35)	1.97 (1.24)	1.65 (1.35)
happen	2.79 (1.10)	1.47 (1.38)	2.71 (1.00)
M (SD)	NP+V(B)	NP+be+p.p.(D)	NP+V+NP(F)
cry	3.44 (0.89)	2.21 (1.49)	2.09 (1.38)
dance	3.06 (1.07)	2.09 (1.36)	2.00 (1.28)
laugh	3.06 (1.15)	1.53 (1.44)	1.94 (1.21)
play	2.79 (1.51)	2.32 (1.47)	2.32 (1.47)
sing	3.44 (0.79)	2.77 (1.44)	2.65 (1.43)
work	2.65 (1.41)	1.94 (1.35)	3.15 (0.89)

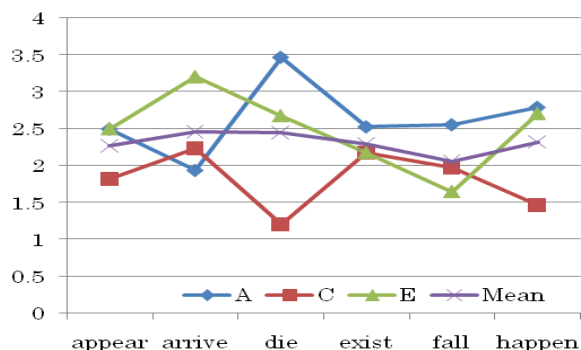


Figure 2: Mean scores of individual unaccusative verbs for the Japanese learners

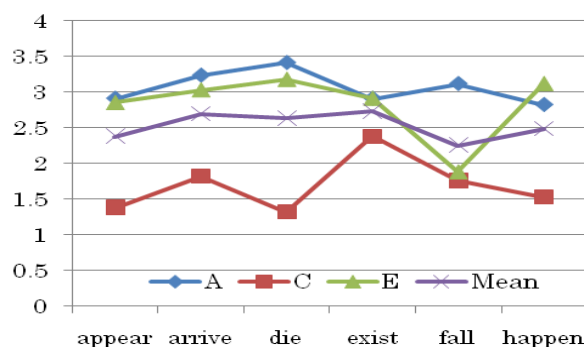


Figure 3: Mean scores of individual unaccusative verbs for the Korean learners

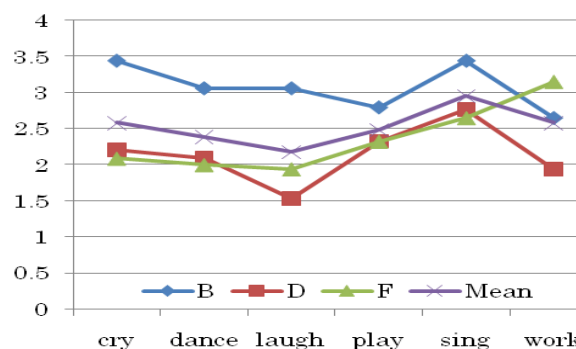


Figure 4: Mean scores of individual unergative verbs for the Japanese learners

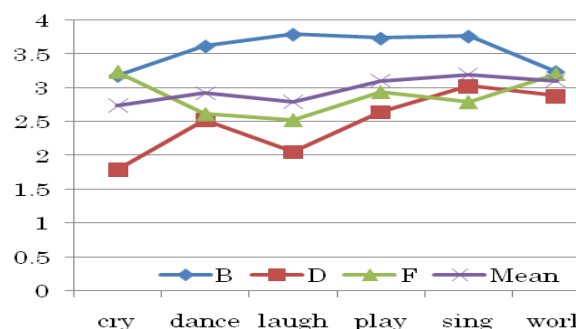


Figure 5: Mean scores of individual unergative verbs for the Korean learners

As shown in Table 5 and Figures 2-5 above, some unaccusative/unergative verbs behaved differently from the others: For example, for the Japanese learners, *die* was significantly better in the NP+V construction while the same verb was significantly worse in the NP+be+p.p. construction. In addition, *fall* was significantly worse in the NP+V+NP construction for the two L1 groups. However, there was not a clear indication of statistical differences that could support the Split Intransitive Hierarchy.

There was a significant difference among almost all of the individual verbs according to the three construction types: All the unaccusatives except *appear* and *exist* for the Japanese learners, and all the unaccusatives except *exist* for the Korean learners; all the unergatives except *play* for the Japanese learners, and all the unergatives except *work* for the Korean learners.

As illustrated in Section 3.4.1, the distinction between unaccusatives and unergatives was made in L2 acquisition; however, there was enough evidence that would support the claim that there was such a wide variation in between-unaccusatives and between-unergatives respectively, and in between-construction types (A-C-E and B-D-F, respectively) that it might be a little too simplistic to categorize those different verbs under the same labels such as “unaccusatives” and “unergatives” (Also see Kim, 2006). Therefore, it is necessary, at

least, to give due attention to the characteristics of individual verbs as well as those of linguistic categories at the same time.

### 3.4.3 Japanese and Korean learners (RQ4)

As shown in Figure 1, the Korean learners performed better overall than the Japanese learners. A two-way repeated measures ANOVA with two independent variables (L1 and construction type) was conducted on each of the six unaccusatives and the six unergatives. As for unaccusatives, there was almost no significant difference between the two L1 groups except one verb: the Korean learners performed significantly better on *arrive* used in the NP+V construction. As for unergatives, on the other hand, the Korean learners performed significantly better on *dance*, *laugh*, *play*, and *work*.

## 4 Discussion and conclusion

The present study showed the following findings: 1) The Unaccusative Hypothesis was validated in L2 acquisition; 2) the NP movement account was a more plausible candidate for the possible cause of overpassivization of unaccusatives; 3) there was a wide range of difference among individual verbs in terms of between-verbs and between-construction types; however, evidence of such variation did not support the Split Intransitive Hierarchy; 4) the findings 1-3 above were applicable to both the Japanese and Korean learners. In addition, the Korean learners performed better than the Japanese learners on unergatives; on the other hand, there was not much difference between their performances on unaccusatives.

### Select Bibliography

Hirakawa, M. (2003). *Unaccusativity in second language Japanese and English*. Tokyo: Hituzi Syobo.

Hwang, J-B. (1999). Learnability and the L2 development of English unaccusative verbs. *Journal of the Applied Linguistics Association of Korea*, 15, 65-87.

Hwang, J-B. (2001). Focus on form and the L2 learning of English unaccusative verbs. *English Teaching*, 56, 111-133.

Ju, M. K. (2000). Overpassivization errors by second language learners. *Studies in Second Language Acquisition*, 22, 85-111.

Kim, J-T. (2006). Unaccusativity and L2 passive construction. *English language & Literature Teaching*, 12, 69-89.

Kondo, T. (2005). Overpassivization in second language acquisition. *IRAL*, 43, 129-161.

Kwak, H-Y. (2003). A study of Korean high school students' overpassivization errors. *Journal*

*of Pan-Pacific Association of Applied Linguistics*, 7, 169-187.

No, G-H., & Chung, T. (2006). Multiple effects and the learnability of English unaccusatives. *English Teaching*, 61, 19-39.

Oshita, H. (1997). *The unaccusative trap: L2 acquisition of English intransitive verbs*. Ph.D. dissertation. University of Southern California, LA, California.

Perlmutter, D. M. (1978). Impersonal passives and the unaccusative hypothesis. *Proceedings of the Berkeley Linguistics Society*, 4, 157-189.

Shimizu, Y., Yamakawa, K., Sugino, N., Ohba, H., & Nakano, M. (2006). Developing a sample-free grammatical proficiency test for SLA research. *Proceedings of the 10th Conference of Pan-Pacific Association of Applied Linguistics*, 227-236.

Sorace, A. (1997). Acquiring linking rules and argument structures in a second language. In L. Eubank, L. Selinker, & M. Sharwood Smith (Eds.), *The current state of interlanguage* (pp.153-175). Amsterdam: John Benjamins.

Sorace, A., & Shomura, Y. (2001). Lexical constraints on the acquisition of split intransitivity. *Studies in Second Language Acquisition*, 23, 247-278.

Yamakawa, K., Sugino, N., Kimura, S., Nakano, M., Ohba, H., & Shimizu, Y. (2003). The development of grammatical competence of Japanese EFL learners: Focusing on unaccusative/unergative verbs. *Annual Review of English Language Education in Japan*, 14, 1-10.

Yamakawa, K., Sugino, N., Ohba, H., Nakano, M., & Shimizu, Y. (2007). Variation in the acquisition of unaccusative verbs by Japanese EFL learners. *Proceedings of the 12th Conference of Pan-Pacific Association of Applied Linguistics*, 134-137.

Yamakawa, K., Sugino, N., Ohba, H., Nakano, M., & Shimizu, Y. (2008). Acquisition of English grammatical features by adult Japanese EFL learners: The application of Item Response Theory in SLA research. *Electronic Journal of Foreign Language Teaching (e-FLT)*, 5, 13-40.

Zobl, H. (1989). Canonical typological structures and ergativity in English L2 acquisition. In S. M. Gass, & J. Schachter (Eds.), *Linguistic perspectives on second language acquisition* (pp.203-221). Cambridge: Cambridge University Press.

# Aspectual Use of the Negative Forms: a contrastive analysis of Japanese and Turkish

Aydin OZBEK

Faculty of Education, Canakkale Onsekiz Mart University-Turkey

aydinozbek@gmail.com

## Abstract

In this study we aim to illustrate how Turkish and Japanese have different aspectual use and why the past tense marker is interpreted as an aspect marker in the negative forms of Turkish.

In languages such as English and Russian, perfectual aspect expressions are clearly interpreted by the hearers. However, when we consider Turkish or Japanese, there is no clear distinction between the perfectual aspect and past tense in affirmative constructions. Most of the aspectual expressions have ambiguous interpretations and do not have an overt or marked form to be identified. However, negative forms exhibit a clear distinction between the temporal and the aspectual usage. In this sense, Japanese and Turkish show a similarity.

## Keywords

Negation, Aspect, Perfect, Interchange-ability, principle of contrast

## Introduction

In an article entitled 'A pragmatic Analysis of Japanese Negative Verbal Aspect Forms', Szatrowski (1983) presents a well-argued analysis of the conditions determining the choice between Japanese past tense marker –TA, and progressive aspect marker –Teiru in negative answers as responses to questions about the past which end with the –TA form. Following Szatrowski (1983), Teramura (1984)'s seminal work is the one of the leading studies pointing out the distinction between perfect and past marking, especially in negative answer constructions. Traditional textbooks illustrate that Answer sentences should have the same temporal or aspectual marking as that of Question sentences. However, Teramura noticed that, in sentences such as (1), this simple method fails.

### (1) Question

Moo hirumesi wo tabe-TA-ka?  
Already lunch Acc eat-Past-Q  
"Did you eat (your) lunch?"

Answer:

a) Iya, (mada) tabe-tei-nai  
No, yet eat-Asp-neg  
"No, I haven't eat yet"

b) \*Iya, tabe-nakat-TA  
No, eat-neg-Past

### (2) Question

Kinoo, hirumesi wo tabe-TA-ka  
Yesterday lunch –Acc eat-Past-Q  
"Did you eat (your) lunch yesterday?"

Answer:

a) Iya, tabe-nakat-TA  
No, eat-neg-Past  
"No, I didn't eat"

b) \*Iya, tabe-tei-nai  
No, eat-Asp-neg

Teramura (1984:321,322)

By using the negation test, Teramura (1984) proposes that the past tense suffix –TA has a perfective meaning as exemplified in (1). On the other hand, the same suffix plays its original role as a temporal past tense marker in the example (2). This is why (a) should be chosen as the answer for (1), which has an aspectual adverb *moo*. Pragmatically ill-formed (b) should not be chosen because it expresses only past tense interpretation.

As well as in Japanese, verbs in main clauses in Turkish are made up of a verb stem followed by a tense or aspectual suffix. Turkish also has a past tense marker –di and a progressive marker –yor. Although they have syntactic and semantic similarities with their Japanese simple affirmative counterparts, they show clear pragmatic differences in negative constructions. As I will illustrate in the following sections, I would like to investigate the structures and pragmatic properties of negative answer sentences in Turkish by using the contrastive analysis method.

# 1 Turkish Perfect

The Turkish aspectual system has been analyzed by a number of linguists. Aksu Koç (1988), Kornfilt (1997), Slobin and Aksu (1982), Göksel and Kerslake (2005) provide a detailed treatment of tense and aspect in Turkish which were unnoticed in previous researches. Especially, Kornfilt deals with the various properties of the Turkish perfect. In this section, I will refer to some of these works in order to clarify the aspect and tense system in Turkish.

## 1.1 Theoretical Implications

### 1.1.1 Basic Definitions

Kornfilt (1997) and Göksel and Kerslake (2005) argue that Turkish does not have special suffixes which function only for aspectual situations in all morphological contexts. This does not mean that Turkish lacks an aspectual system. For instance, the perfect aspect, which indicates the continuing present relevance of a past situation (Comrie 1978:52), is usually expressed by the past tense morpheme *-di*.

- (3) Hasan balig-i ye-di.  
Hasan fish –Acc eat-Past  
“Hasan ate the fish”  
“Hasan has eaten the fish”  
(Kornfilt 1997: 349)

As exemplified in Kornfilt’s example (3), the past tense morpheme has ambiguous interpretations. This ambiguity can be removed by using temporal or adverbial adverbs such as *saat ikide* (at two o’clock) or *coktan* (already).

- (4) Hasan balig-i saat ikide ye-di. (Past)  
Hasan fish-Acc hour two-Loc eat-Past  
“Hasan ate the fish at two”  
(5) Hasan baligi çoktan ye-di. (Perfect)  
Hasan fish-Acc already eat-Past  
“Hasan has already eaten the fish”

It is interesting to observe the fact that constructions such as (3), without a temporal or an aspectual adverb, can bear both past tense and aspectual interpretation.

As already mentioned, Turkish has a progressive marker *-yor*. However, this suffix is used not only for the continuative aspect of stative verbs, but also for the progressive aspect (continuative aspect of a nonstative verb) and ingressive aspect, as illustrated in examples (6), (7) and (8).

- (6) Continuous aspect of a stative verb  
Ali Ayse-yi tan-iyor.  
Ali Ayse –Acc know-Prog  
“Ali knows Ayse”  
(7) Continuous aspect of a non-stative verb  
Ali televizyon izli-yor.  
Ali television watch-Prog  
“Ali is watching television”  
(8) Ingressive aspect  
Ali yat-ıyor.  
Ali lie down –Prog  
“Ali is going to bed (now)”

As mentioned in Kornfilt (1997:357), it must be noted here that “in English, the progressive form ‘was knowing’ in the translation of this example with a stative verb would be ungrammatical. Likewise, verbs of inert perception cannot appear in the progressive in English while they can do so in Turkish.”

### 1.1.2 Inter-changeability between the Past and the Continuative aspect

As I will demonstrate, *-yor* is not the only suffix which expresses the continuative aspect. This is also proposed by Göksel and Kerslake (2005). Continuative aspects often appear in a past tense marked form corresponding to genuine progressive construction. These forms are formed by the past tense marker *-di*. Now, let us consider these affirmative structures.

- (9) Bura-da ol-ma-na sevin-di-m.  
Here-Loc be-N-Dat be glad-Past-1stSng  
“I am glad you are here”  
(10) Bura-da ol-ma-na sevin-iyor-um.  
Here-Loc be-N-Dat be glad-Prog-1stSng  
“I am glad you are here”

Here, the example (9) which is the past tense form of (10) shows almost exactly the same interpretation that the speaker’s state (to be glad) is still continuing. According to Göksel and Kerslake (2005:334), while examples such as (9) express an “entry into a state” meaning, examples such as (10) express the “state itself”. However, this is true only for some psychological verbs such as *sevin-* (be glad), *üşü-* (be cold).

In the case of affirmative constructions of non-stative verbs such as *koş-* (run), *ye-* (eat), *taşı-* (carry), *düş-* (fall down), this kind of inter-changeability as illustrated below will not give

identical interpretations<sup>1</sup>.

## 2 Aspect in Negative Constructions

There has so far been less investigation of the expression of the continuative aspect of Turkish negative constructions. As mentioned above, interchangeability between past tense and aspect is only acceptable under special conditions. As Kornfilt (1997) states, in Turkish *-den beri* constructions are used to convey a situation which began in the past and is still continuing, which is expressed by “have been –ing” in English. Her statement also proposes that to convey this meaning, a present progressive marker *-(I)yor* is needed.

- (11). Beş saat-ten beri sen-i bekl-iyor-um.  
Five hours-Abl after you-Acc wait-Prog-1sng  
‘I have been waiting for you for five hours’

This statement should be considered to be true to express the present perfect aspect in Turkish. However, not only the set of an ablative marker *-den*, a postpositional *beri* (after) and a progressive tense marker *-(I)yor* gives us this interpretation, but also a past tense marker *-di* can be used with the *-den beri* set only in the negative constructions as shown in (12b).

Here, an interesting fact about the semantics of negative constructions in Turkish is that the past form *-di* can also be inflected to a stative verb to refer to continuing situations. Now let us consider the following examples;

- (12)  
a. Sabah-tan beri hiçbir şey ye-mi-yor-um.  
morning-Abl after any thing eat-Neg-Prog-1sng  
‘I have not been eating anything since morning’  
(Speaker is still hungry)

<sup>1</sup> In Turkish youth-language, interchangeability between the past tense and the progressive marker is widely seen as below. These forms are, however, considered to be highly informal and criticized by many scholars with flat denial.

- |  |  |
|--|--|
| i. Tesekkür et-ti-m<br>thank do-Past-1stsng<br>“I thank you” | ii. Tesekkür ed-iyor-um<br>thank do-Prog-1stsng<br>“I thank you” |
|--|--|

Or,

- |  |   |
|--|---|
| iii. Ben git-ti-m.<br>I go-Past-1stsng<br>“I am going (now)” | iv. Ben gid-iyor-um<br>I go-Prog-1stsng<br>“I am going (now)” |
|--|---|

- b. Sabah-tan beri hiçbir şey ye-me-di-m.  
morning-Abl after any thing eat-Neg-Past-1sng  
‘I have not been eating anything since morning’  
(Speaker is still hungry)

In (12a) and its counterpart example (12b), despite the past tense marker in the construction, the speaker’s situation has not changed, which suggests that it has begun in the past and still is continuing. Yet, past tense interpretation of (12b) will exhibit an ill-formed situation. Negative polarity items (*hiçbir şey* /anything) gives a higher negation modality to the constructions. However, the affirmative counterpart of (12b) without a negative interpretation would be ill-formed, as exemplified in (13b).

- (13)  
a. Sabah-tan beri yi-yor-um.  
morning-Abl after eat-Prog-1sng  
‘I have been eating since morning’

- b. ??Sabah-tan beri ye-di-m.  
morning-Abl after eat-Past-1sng  
‘I have not been eating anything since morning’

Let us now focus on the acceptability of (12b). The first observation to be made is that there is a clear pragmatic difference between the meaning of (12a) and past tense marked (12b). In both constructions the situation of “not eating” is still continuing, and the speaker is still hungry at the time of speaking. My informants found both sentences fully acceptable. Yet, while (12a) gives us a modal meaning in which the “not eating” situation is achieved by the speaker’s own will, in (12b) the “not eating” situation is only expressed as a topic. This distinction could be briefly explained by Clark (1987)’s principle of contrast.

*The Principle of Contrast: Every two forms contrast in meaning.*

According to Clark (1987), this principle states that “any difference in form in a language marks a difference in meaning”

Note that both the progressive marker and past tense marker share the notion of “continuative aspect” in Turkish negative structures, even though both structures *prima facie* have the same aspectual formations. However, as expressed in the Principle of Contrast, they are naturally distinguished depending on whether the speaker has a strong will or not. In the next section, I will attempt to explain why this notion is possible in Turkish.

## 2.1 The Proposed Solution

Negation or so-called negative assertion is briefly described in Givon (2001:302) “*The proposition is strongly asserted to be false, most commonly in contradiction to the hearer’s explicit or assumed beliefs. A challenge from the hearer is anticipated, and the speaker has evidence or other strong grounds for backing up their strong belief.*”

This proposal also gives us another hint about negative constructions. It is not possible to consider a large number of further examples here, but what is clear is that the –yor in negative constructions does not only give the construction a continuative aspect meaning but also gives a modality interpretation which exhibits the speaker’s “own will” or “strong denial”. Let us consider following (non-stative verbal) examples;

### (14) Question

Kardes-in universite-ye git-ti-mi?  
 Brother-2nd Gen college-Dat go-Past-Q  
 “Has your brother gone to the college?”

### Answer

a. Hayir, git-mi-yor.  
 No, go-Neg-Prog  
 “No, he hasn’t” (He doesn’t want to go)

b. Hayir, git-me-Di  
 No, go-Neg-Past  
 “No, he hasn’t” (event)

(16) John oda-sı-nı topladı-Di-mi?  
 John room-Acc tidy-Past-Q  
 “Has John tidied up his room?”

### Answer

a. Hayir, topladı-mı-yor  
 No tidy-Neg-Prog  
 “No he hasn’t” (And he doesn’t want to do)

b. Hayir, topladı-ma-Di.  
 No tidy-Neg-Past  
 “No, he hasn’t” (event)

It seems, therefore, that the inter-changeability process seems a false approach that is contrary to the principle of contrasts. It is interesting to observe that the past tense marker always has a continuative aspectual interpretation, while the actual continuative marker –yor bears another role, to express the agent’s will. Past tense markers in negative constructions are filling the slot which is emptied by the progressive –yor.

## 3 Conclusion

The foregoing analysis of the behavior of the past tense and progressive markers has illustrated aspectual and modality notions such as the continuative and expressing the agent’s will. Showing syntactic similarities to Japanese, it was clarified that Turkish has a completely different strategy for negative constructions. In Turkish, the past tense marker asserts that the unchanged situation is continuing in the speech time; while on the other hand, the progressive marker implies the agent’s behaviors and bears a modal role. In this sense, Turkish and Japanese negative constructions need to be investigated with further studies.

## References

- Aksu-Koç, A. (1988). *The Acquisition of Aspect and Modality*. Cambridge University Press.
- Clark, E. V. (1987). The *Principle of Contrast*: A constraint on language acquisition. In MacWhinney, B., editor, *Mechanisms of Language Acquisition*, pp 1-33.
- Givon, T. (2001). *Syntax: An Introduction Volume I*. John Benjamins Publishing Co.
- Göksel, A & Kerslake, C. (2005) *Turkish: A Comprehensive Grammar*. London & New York: Routledge
- Kornfilt, J (1997). *Turkish* (Descriptive Grammars) Routledge, London
- Slobin, I. and Ayhan A. Aksu (1982). Tense, Aspect and Modality in the Use of the Turkish Evidential. In *Tense-Aspect* Hopper, Paul J. (ed.), 185 ff.
- Szatrowski, P. (1983). A pragmatic Analysis of Japanese Negative Verbal Aspect Forms. *Tsukuba working papers in linguistics* 2 pp. 48-64
- Teramura, H. (1984). *Nihongo no sintakusu to imi II* (Japanese syntax and semantics II) Kurosio Publishing.

# National Policy and the Ideology of English: The Case of Taiwan

Han-Yi Lin

Department of Applied English, the Overseas Chinese Institute of Technology

`hanyilin@ocit.edu.tw`

## Abstract

This study argues that the spread of English in the world involves not only the effect of policy making and education practices but also the subtle influence of discourses regarding English. By presenting a textual analysis of national policy, the ideological character of the English language in Taiwan is investigated. Through the employment of Critical Discourse Analysis and by analyzing how English is presented in the governmental document, it examines the presentation and construction of discourses on English. This research suggests that discourses concerning English in Taiwan are conditioned by the external global structure and Taiwan's socio-cultural context. Through discourses regarding the necessity of English for Taiwanese society, the ideology of English consequently functions as part of the mechanisms involved in the global spread of English.

## Keywords

Ideology of English, global English, English in Taiwan, Taiwan six-year national plan, critical discourse analysis

## Introduction

The research proposed here is a case study of the impact of the global spread of English in an EFL country, focusing on discursive practices and ideological effects. Although the global spread of English has been investigated by many scholars employing various perspectives, its impact and consequences are still controversial (see Davies, 1996; Phillipson 1992; Seidlhofer, 2003). This is not only because the phenomenon of global English is unprecedented, but also because it is complex and contextually varied. This research intends to offer a study of the underlying knowledge and power relations between global English and national policy in an East Asian country, Taiwan, in order to obtain a more complete and detailed picture of this global phenomenon.

In the case of Taiwan, the spread of English involves not only the effect of policy making and

education practices but also the subtle influence of discourses regarding English. In order to uncover the ideological character of the English language in Taiwan, this research focuses on the prevailing beliefs, values and propositions relating to English as a global language and the spread of English in the country. By analyzing how English is presented in national development plan, it examines the presentation and construction of discourses of English. First, I look at some significant research on the ideology of English and the role of the English language in *Taiwan Six-Year National Development Plan* (2002-2008). The second part of this study outlines the methodological background of this research. In the third part, the construction of discourses concerning English and ideological assumptions regarding the English language in Taiwan is investigated. Finally, discourses and ideological assumptions identified in the texts and their implications are discussed.

## 1 Global English, Ideology and Taiwan's National Plan

### 1.1 Criticism of Global English and the Ideology of English

Many scholars have done research on the critical investigation of the global spread of English (e.g. Canagarajah, 1999; Kubota, 1998; Pennycook, 1994; Phillipson, 1992; Tollefson, 2000; Tsu and Tollefson, 2007). Two significant works were published by Phillipson (1992) and Pennycook (1994). Both Phillipson and Pennycook advocate a critical examination of the global spread of English and expose the correlations between the English language and socio-economic inequality. However, under Phillipson's theoretical framework, the hegemony of English is examined and explained through the notion of imperialism and oppression and thus inevitably falls into the pattern of oppressor vs. the oppressed where the oppressed are powerless victims of oppressors (Anglophone speakers vs. non-Anglophone speakers in the case of English linguistic imperialism). This kind of politico-economic analysis to a certain extent might

effectively explain language policies and language planning in a collective sense but will inevitably encounter difficulties dealing with issues such as the confrontation between English and local languages, the resistance to English, the localization of English and the subjectivity of individuals. By imposing a powerless position on non-Anglophone speakers, the effect of power is unidirectional and hard to reverse, and the roles of the individual are fixed and controlled.

However, in fact, we are encountering rather complex and divergent phenomena of the global spread of English where roles of the individual are shifting, i.e. a non-Anglophone speaker in Taiwan could be a member of a minority ethnic group and meanwhile an influential official in the government, and power is uncentered as well as multiple-directional. In terms of the subjectivity of the individual, Phillipson's accounts to a certain degree lose its effectiveness for being unable to explain the practices and resistance from the oppressed side and the dispersal struggles and negotiations of power in discourse of the English language. Comparatively Pennycook offers a more effective investigation of the global spread of English. Without a clear agency of power and an evident causal relationship between intentions of native English speakers and outcomes of English spread, power struggles and the effect of ideology are exercised in discourse of English where there are no clear-cut oppressors to blame or victims to save. The focus has been shifted from the oppression of imperialism to the effect of power in processes of discourse. Therefore, rather than under a grand social structure, the ideology of English needs to be investigated in a more contextual and discursive way in order to effectively explain the divergent phenomena of global English and the complex relations of power.

## 1.2 English and Taiwan's National Plan

In the past few decades, the spread of English learning has been growing dramatically not only because of globalization, but also due to a fundamental change of perception and learning culture toward the English language. English is no longer regarded as a foreign language, but an indispensable part of national policy.

*Taiwan Six-Year National Development Plan* (2002-2008) includes wide-ranging projects focusing on economic growth and environmental protection, i.e. three major reforms (i.e. government, banking, and finance) and four major investments (cultivating talent; research, development, and innovation; international logistics; and a high-quality living environment). The plan was formulated by the Taiwanese government in order to face the technological challenges and a loss of

investment and skilled managers to Mainland China.

The English language plays a central role in the project for cultivation of talent for the E-generation. It aims to cultivate manpower in order to meet the challenges of globalization and internationalization (GIO, 2005). The project has three parts: (1) establishing an internationalized living environment and enhancing national English ability; (2) cultivating the young generation with vitality through sports and arts; (3) creating a society of lifelong learning (MOE, 2005). The first part of the project focuses on the promotion of English under the basic tactic of the six-year national development plan— 'global connection, local action'. In that, English has been seen as the key element of globalization and modernization in Taiwan.

## 2 Research Methodology

As Eagleton (1994) and Lull (2000) suggest, ideology, as a system of ideas expressed in communication, is shaped by as well as shaping our actions. Ideology has persuasive force only when these ideas can be represented and communicated. The mass media and all other large-scale institutions play a vital role in the circulation of ideologies. According to this assumption, the system of ideas about English is conditioned by as well as conditioning the social situation in the process of communication. The values, predispositions and orientations towards English in Taiwan are embedded and thus identifiable in the discourses produced by cultural institutions including the government and the education system.

As far as the ideology of English is concerned, the hypothesis proposed here is that there exists a system of ideas about the English language embedded in social practices of Taiwanese society through cultural institutions, e.g. the government. And discursive practices as the linguistic aspect of the social order play a significant role in the production and reproduction of the ideology of English. Under these assumptions, a discussion of CDA as research method, the rationale of text selection, and the procedure of text analysis are explained in this section.

### 2.1 Critical Discourse Analysis as Research Method

Critical discourse analysis (CDA), with a particular interest in the relation between language and power, regards 'language as social practice' (Fairclough and Wodak, 1997) and considers the context of language use to be crucial. Language as social practice is 'a key instrument in socialization, and the means whereby society forms and permeates the

individual's consciousness (Hodge and Kress, 1993: 1).' Hodge and Kress define the constructed consciousness of the society as ideology – 'a systematic body of ideas, organized from a particular point of view' (ibid, p.6). For CDA, ideology is regarded as an important means of establishing and maintaining power relations. The relationship of language and power where 'language mediates ideology in a variety of social institutions' (Weiss and Wodak, 2003: 14) therefore is given considerable attention by CDA analysts. For language is not powerful on its own, rather it indicates, expresses, and even challenges power. It is thus important for CDA to investigate the expressions and manipulations of power by analyzing linguistic forms in context.

According to the aim and hypothesis of this research, that is, there is a need to investigate the role of discursive practice in the maintenance of the predominant status of global English and the ideological effects of discourses on the development of ELT in Taiwan. Here, CDA is a useful instrument for analyzing the relations between the English language, power and ideology and is therefore adopted for analyzing governmental discourses in order to investigate the cultural and ideological meanings of global English in Taiwan.

## 2.2 The Rationale of Text Selection

In terms of text selection, although every instance of language use can be a reflection of social order, some instances such as governmental documents are more influential due to the process of their production and consumption. The assumption here is that the texts created by official institutions can be regarded as manufactured and highly-circulated discourses which produce, reproduce or challenge the existing power relations in Taiwanese society and therefore are more influential and prevailing at the socio-cultural level. The selected text is an extract from a significant governmental document, *Taiwan Six-Year National Development Plan*, which represents the blueprint of national development from 2002 to 2008 and emphasizes the importance of English at national level. This text is selected because of its significant influence on Taiwan's government policies and educational practices of English.

## 2.3 Application of CDA Methods in this Research

By employing Fairclough's (1992, 1995, 2001a, 2001b) analytical framework of CDA, Systemic Functional Grammar (Halliday and Matthiessen, 2004), and discourse/text analysis, this research examines the features at the three levels of analysis

in the selected texts: textual analysis (the stage of description), discursive practices (the stage of interpretation), and social practices (the stage of explanation).

### 2.3.1 Textual analysis

At the level of textual analysis, the focus is on content, form and textual organization. Therefore, by focusing on representations, social relations and identities, and cohesion and coherence (Fairclough, 1995; Richardson, 2007), the stage of description includes three levels of analysis: lexical analysis, sentence structures, and textual structures.

### 2.3.2 Analysis of discursive practices

Discursive practice mediates between text and social practice and focuses on 'how authors of texts draw on already existing discourses and genres to create a text and on how receivers of texts also apply available discourses and genres in the consumption and interpretation of the texts' (Jørgensen and Phillips, 2002: 69). In other words, the interpretation of a text is not only constructed by the formal features of a text but also depends on background assumptions of the reader, 'the members' resource' (MR) (Fairclough, 2001a:118). Therefore, the analysis of discursive practice involves not only an explanation of how a text is produced and interpreted but also 'what genres and discourses were drawn upon in producing the text, and what traces of them are there in the text' – intertextuality (Fairclough, 1995: 61).

### 2.3.3 Analysis of social practices

The analysis of social practices assumes a dialectical relationship between society and texts. At this stage, the MR is regarded as ideologies because they are associated with assumptions about culture, social relationships, and identities. It is thus important to investigate 'what elements of MR which are drawn upon have an ideological character' (Fairclough, 2001a:138). By focusing on processes of social struggles or relations of power, the analysis of social practices examines determinants and effects of discourse at three levels: societal, institutional and situational.

## 3 Analysis of Text: Cultivate Talent for the E-Generation

Text 1 (see Appendix A) is an extract of *Taiwan Six-Year National Development Plan*. With its title – *Cultivate Talent for the E-Generation*, the main idea of this project focuses on manpower. This project is the first of the ten projects of *Taiwan Six-Year National Development Plan* and is

designed to be implemented as an independent project. Text 1 is an extract from an English edition of *Taiwan Six-Year National Development Plan* which is provided by Government Information Office and available on both the Government Information Office Website (<http://www.gio.gov.tw/taiwan-website/4-0a/20020521/2002052101.html>) and the Washington State Taiwan Office Website.

In terms of textual analysis, Text 1 is preoccupied with manpower and development by wording including the nouns *cultivation, talent, e-generation, manpower, development, abilities* and the verbs *cultivate, enhance, establish, expand, promote, integrate*. Most sentences of the text are formulated in modality forms through non-modalized categorical assertions, e.g. '*manpower is the basis of all development*' (in sentence 1). or through the use of modal auxiliary verbs, e.g. *should* (in sentence 2 and 5), *will* (in sentence 3), and *must* (in sentence 8).

According to Fairclough (2001a), the use of modality is related to author's authority or power in terms of relations between participants or the truth of a proposition. Non-modalized categorical assertions can be regarded as a form of modality that expresses the author's strongest commitment to the truth of the proposition and that affirmatively supports a view of the world as transparent and self-evident. The modality forms used here represent a declarative mode, where the subject position of the author (i.e. the government) is the subject position of an information provider and the audience's position is a receiver (Fairclough, 2001a: 104-105). The forms manifest a commitment to the truth of the statements and to the implementation of policies. The authority and commitment of the government can also be identified through sentences in modality form with *the government* or *the ROC* as subject.

According to Fairclough (2001a:109-110), the content of the subordinate clauses (*since...*) are presuppositions. Two presuppositions can be identified in Text 1: '*manpower is the basis of all development*' (in sentence 1) and '*English is the language that links the world*' (in sentence 5). Through the use of nominalization, *the basis of all development* and *the language that links the world* are presupposed. By using relational processes, the relationship between the elements of the two sentences (e.g. *manpower* and *the basis of all development*) is set up with causality left unclear. These two presuppositions are treated as clear and undoubted facts or truths. However, if we closely examine these two presuppositions, they may be seen not as unquestionable truths but common sense assumptions. *Manpower* is of great importance for

the *development* of a nation, especially economic development. However, it is questionable if *the basis of all development* is *manpower*. It is quite possible that *all development* refers to national development (since it is a presupposition in a project under *Taiwan Six-Year National Development Plan*) and therefore focuses on economic development according to Taiwan's socio-historical background. *English as the language that links to the world* implies that English is the communicative instrument to interact with people, institutions and countries around the world. This presupposition is also related to sentence 2 in the first paragraph. *The ROC should first enhance the abilities of its people* in order to *meet the future challenges of globalization and internationalization*. The ability to master English or, in other words, the ability to link to the world is especially emphasized for facing *the challenges of globalization and internationalization*. However, it is not necessarily true that one cannot *link to the world* without the English language, and language is only one facet of the challenges of globalization and internationalization.

The linkage between national planning, manpower and English can also be identified through the analysis of discursive practice. The following discourses can be identified in Text 1: a discourse of national development and manpower; a discourse of the inevitability of globalization and internationalization; a discourse of English as the global language to link to the world. Manpower is considered the basis of national development and needs to acquire some skills required for facing the future challenge. Globalization mainly refers to economic globalization which is a process of integrating national economies into the international economy through trade, capital flows, foreign investment, migration and spread of technology. It can also be considered as the internationalization of everything related to foreign countries. The process of globalization and internationalization is not only regarded as an inevitable trend but also a challenge for Taiwanese society in the face of the competitive nature of the international economy. English as the global language is therefore treated as one of the required skills for connecting to the world and facing this future challenge. By combining these three discourses, a strong association between national development and the English language is constructed.

In the analysis of socio-cultural contexts, *Taiwan Six-Year National Plan* was proposed by DPP (Democratic Progressive Party) government in 2002 – a midterm policy of DPP's first four-year term of office (from 2000 to 2004) which intended to raise public support, by including almost all

important issues related to economic growth and national development. The promotion of English is presented as one of the elements for cultivating talent for the e-generation. Actually, English is the main focus of this project as articulated in the text. Accordingly, English will be promoted in two ways: by giving it an official status and by expanding its use in daily life. Although the statements are simple, the implications of these two methods of English promotion are considerable. They involve the development of Englishization nationwide; a change of English education; enhancing quality and quantity of English teachers; internationalization of universities and colleges; recruitment of foreign students and encouragement of studying overseas (MOE, 2005).

Furthermore, the significance of this project can especially be identified through other discourses in English education, for example, the *Teachers' Manual for English Language Teaching in Elementary and Junior High Schools*, which provides teachers with directions and references in their teaching. It can be regarded as discourse following the aims and rationale of English promotion in this project.

## 4 Discussion and Implications

### 4.1 Discourse of English

This research has analyzed how English is discursively presented in national policy. Common-sense assumptions in the text not only indicate the existing power relations and power struggles, but also sustain or change power relations. The underlying ideological assumptions of English are closely linked to the current trends in globalization or internationalization, the concepts of economic development, and competition. In addition, the importance of the modern state in the global capitalist system also has implicit effects on the discourse of English. It results in discourses such as national planning and manpower, and a political discourse that configures Taiwan in the world. They are drawn on for constructing the need for English at the national level. And through the mixture of the above mentioned discourses and some widely-accepted assumptions in the field of English language teaching such as 'the younger the better', the national and personal need for adequate English education is reinforced. These discourses can be regarded as dominant discourses of English, though subject to text selection in this research, reflecting prevailing assumptions of English in Taiwanese society. In other words, these discourses are used as 'background knowledge' and thus constrain the production and interpretation of the discourse of English.

### 4.2 Englishization as Globalization: A New Form of Modernization

The national policy of English promotion in Taiwan can be further examined within the wider context of East Asia and the global world. Taiwan and other East Asian countries have experienced a rapid growth in the ELT industry and changes in national policy for improving national English ability since the 1990s (Ho and Wong, 2003; Nunan, 2003). For these countries, the rationale for promoting English and problems regarding English education are similar in many respects. Similar to the process of modernization in the late 19th and the 20th century, globalization and internationalization are the predominant themes in the late 20th and the early 21st centuries. This process has also been seen as the so-called 'Englishization' of the world's political, economic and social communications. It is assumed that this trend is necessary in today's globalized world and perceived as instrumental in stimulating economic competition. In the competitive environment existing between countries such as China, Japan, South Korea and Taiwan, it is assumed that national levels of English proficiency can enhance human capital as a resource, and thus make the country more competitive. Similar to other East Asian countries, Taiwan faces the pressures of globalization and economic competition, and Englishization is adopted as a new, indispensable means for international competition in the globalized world.

## 5 Conclusion

In conclusion, it would be useful to consider the claim of Foucault (1988), that power is everywhere. The power of English can no longer be identified with a center-border dichotomy, i.e. the powerful native English speakers vs. the powerless non-English speakers. In the context of Taiwan, English is not imposed but spreads through global trends and the discourse of national plan. At the global level, this research argues that the spread of English can be regarded as a continuing manifestation of modernization which started in the late nineteenth century among East Asian countries. It aims to strengthen national politico-economic capability and to compete with Western powers and other countries in East Asia. In the case of Taiwan, the intention of improving the national level of English is simply to acquire the language, engage and become successful in global economic competition, and share the benefits. Through discourses regarding the necessity of English for Taiwanese society, the ideology of English consequently functions as part of the mechanisms involved in the global spread of English.

## References

- Canagarajah, A. S. (1999) *Resisting Linguistic Imperialism in English Teaching*. Oxford: Oxford University Press.
- Davies, A. (1996) Reviewing Article: Ironising the Myth of Linguicism. *Journal of Multilingual and Multicultural Development* 17, 6, 485-496.
- Eagleton, T. (1994) Introduction. In Eagleton, T. (ed.) *Ideology*. London: Longman.
- Fairclough, N. (1992) *Discourse and Social Change*. Cambridge: Polity Press.
- Fairclough, N. (1995) *Media Discourse*. London: Edward Arnold.
- Fairclough, N. (2001a) *Language and Power* (2nd Edition). Harlow: Longman.
- Fairclough, N. (2001b) Critical discourse analysis as a method in social scientific research. In Wodak, R. & Meyer, M. (eds.) *Methods of Critical Discourse Analysis*. London: SAGE.
- Fairclough, N. & Wodak, R. (1997) Critical discourse analysis. In van Dijk, T.A. (ed.) *Discourse as Social Interaction*. London: SAGE.
- Foucault, M. (1988) *Politics, Philosophy, Culture : Interviews and Other Writings, 1977-1984* (translated by A. Sheridan and others/ edited with an introduction by L. Kritzman) London, New York : Routledge.
- GIO (Government Information Office) (2005) *Taiwan Six-Year National Development Plan*. Available at <http://www.washingtonstate.org.tw/English/taiwan-trade-economy/6-year-plan/index.htm>, accessed April 20, 2005.
- Halliday, M.A.K. & Matthiessen, C.M.I.M. (2004) *An Introduction to Functional Grammar*. (3rd edition) London: Hodder Arnold.
- Ho, W. K. & Wong, R. Y. L. (eds.) (2003) *English Language Teaching in East Asia Today: Changing Policies and Practices*. Singapore: Times Academic Press.
- Hodge, R. & Kress, G. (1993) *Language as Ideology*. (2nd edition) London: Routledge.
- Jørgensen, M. & Phillips, L. (2002) *Discourse Analysis: As Theory and Method*. London: SAGE.
- Kubota, R. (1998) Ideologies of English in Japan. *World Englishes* 17, 3, 259-306.
- Lull, J. (2000) *Media, Communication, Culture: A Global Approach* (2nd edition). Cambridge: Polity Press.
- MOE (the Ministry of Education) (2005) *Challenge 2008: Key National Development Plan – Project of Cultivating Talent for E-Generation*. Taipei: MOE.
- Nunan, D. (2003) The impact of English as a global language on educational policies and practices in the Asia-Pacific region. *TESOL Quarterly* 37, 4, 589-613.
- Pennycook, A. (1994) *The Cultural Politics of English as an International Language*. London: Longman.
- Phillipson, R. (1992) *Linguistic Imperialism*. Oxford: Oxford University Press.
- Richardson, J.E. (2007) *Analysing Newspapers: An Approach from Critical Discourse Analysis*. Basingstoke : Palgrave Macmillan.
- Seidlhofer, B. (2003) *Controversies in Applied Linguistics*. Oxford: Oxford University Press.
- Tollefson, J.W. (2000) Policy and ideology in the spread of English. In Kelly Hall, J. and Egginton, W. (eds.) *The Sociopolitics of English Language Teaching*. Clevedon: Multilingual Matters.
- Tsui, A.B.M. & Tollefson, J.W. (eds.) (2007) *Language Policy, Culture, and Identity in Asian Contexts*. Mahwah: Lawrence Erlbaum.
- Weiss, G. & Wodak, R. (2003) Introduction: Theory, interdisciplinary and critical discourse analysis. In Weiss, G. & Wodak, R. (eds.) *Critical Discourse Analysis: Theory and Interdisciplinary*. Basingstone: Palgrave Macmillan.

## Appendix A Cultivate Talent for the E-Generation- An Extract of Taiwan Six-Year National Development Plan

### 1. Cultivate Talent for the E-Generation

(1) The first project of the 6-year national development plan is the cultivation of talent for the e-generation, since manpower is the basis of all development. (2) To meet the future challenges of globalization and internationalization, the ROC should first enhance the abilities of its people. (3) Concurrently, the government will establish an environment for internationalizing learning.

(4) This project emphasizes the ability to master foreign languages, especially English, and the use of Internet. (5) Since English is the language that links the world, the government should designate English as a quasi-official language and actively expand the use of English as a part of daily life.

(6) In addition, physical health and culture is the foundation for the next generation of society. (7) Therefore, education in culture, art, sports, and civility is an important goal for the new government. (8) Since it is necessary for the entire society to enter the new era with the e-generation, we must make concerted efforts to establish a comprehensive life-long learning system, promote voluntary social services, and integrate learning resources, including those in the civil service system, to immerse ourselves in cultural, social transformation, and reconstruction.

Text 1 Extracted from the English version of 'Taiwan Six-Year National Development Plan' (available at <http://www.washingtonstate.org.tw/English/taiwan-trade-economy/6-year-plan/talent-e-generation.htm>, accessed 21 September 2005)

# Student Presentation as a Means of Learning English for Upper Intermediate to Advanced Level Students

Eunpyo Lee

School of Medicine, Eulji University

elee@eulji.ac.kr

## Abstract

This study observes and examines how college students perform and perceive one-topic-for-each student presentation as a means of learning English. The 49 second year medical students enrolled in English reading and writing class in 2008 were given a single topic for each presentation in class as part of writing assignment. They were encouraged to be creative, informative and fun to induce class participation. A questionnaire was given to learn of students' opinion on the presentation. Based on these, the data were analyzed to see if preparing and giving presentation was conducive for medical students. The results revealed that students wrote double length of the previous year writing assignment and 92% of them felt presentation on different topics was interesting and helpful for learning English. More than two-thirds of the students (69%) thought that preparing for their class presentation provided them with a chance to learn new vocabulary, expressions as well as knowledge in English. The results show that presentation on various topics can be used as a means of learning English for upper intermediate to advanced level students.

## Keywords

One-topic-for-each presentation, Power Point, UCC

## 1 Introduction

Before the demand of the English language swept the nation as a requirement for globalization, having an ability to speak English well was often considered as higher status and elite group as stated in Shim's (1994) study. However, as the world has become a big global village, competence in the English language seems like a must acquired skill regardless of individual's major or interest.

Being financially hit hard during the IMF era, the Korean government lead by former president Kim Dae-jung emphasized importance of the language. A decade later, the incumbent president Lee Myung-bak and his cabinet members seemed to be driven by reformative measures of accelerating the speed of the Korean people's language competence, thus poured out plans like immersion program, more support for the native and bilingual teachers into teaching, and utilizing the military service period of the men, educated in the English speaking countries, as substitution for English teaching. Such measures have caused a national ripple and resistance from the people due to lack of thorough preparation by the experts and professionals in the field. Any kind of reform in

English education seemed to cause more money spending on the private education market due to parents' fear that their children might not have the edge in the language.

At the upper level of education, English for Specific Purposes (ESP) has become a popular theme for the people who learn and teach majors other than English. Teaching at a medical college, ESP has been one of the research topics for the researchers. Medical students need to get by with the language as a lot of their textbooks are written in English, and they are demanded to publish and present medical papers in the language even after they graduate from medical college. Such phenomenon triggered the demand for English competence greater than ever. Accordingly, teaching the language in the way they are expected to use may serve the purpose of medical college English. As pointed out in Kim's (2008) study, English ability is no longer a choice but a tool for survival. This seems so true especially for these professionals because of their Science Citation Index (SCI) research paper publication in English.

This paper observes and examines an English reading and writing course of second year medical students. All students were asked to prepare a presentation on different topics similar to a research paper and present it in class. It was also intended to induce students' participation in the English class. The student presentation turned out very interesting as each student created a praiseworthy one with unique Power Point materials and intriguing audio-visual clips relevant to his/her topic. The students agreed that the class was enjoyable and active. Presentation class provided them with a chance to learn English on their own to explain about the topic and be able to answer the questions their peers had with a close interaction between the presenter and the classmates.

## 2 Literature Review

### 2.1 Effective ways of English teaching

As the demand and emphasis on communicative competence grew, fluency has become the objective of the English education. Once prevalent grammar-translation method, in some cases still used in classroom, was replaced with learner-centered way of communicative method partly because of the change in the national curriculum policies and

fulfilling the demand from the language users. Such communicative method included audio-lingual method in which students were exposed to the listening materials to improve listening and speaking skills.

Toward the end of the 20<sup>th</sup> and beginning of the 21<sup>st</sup> century, computer technology became part of the English education. Kelm (1992) and Warschauer & Kern (2000) stated that since computer technology helps students to generate their own discourse in the target language, such as in English, with much more freedom than ever before, computer assisted English language instruction could be more learner-centered, allowing students to have multiple opportunities to participate in real-life communication in English. In recent years, many instructors rely on the visual materials as educational technology is highly desirable because it can help improve the effectiveness of English teaching (Park, 2004). A growing number of studies have explored ways to incorporate technology into language classrooms (e.g., Kim, 2002; Kim, 2003; Hong, 2003), especially by comparing two delivery modes: computer-mediated communication (CMC) and face-to-face instruction. Lee & Chang (2001) presented a model of CMC-based English teaching materials intended to improve adult learners' communicative interaction skills. They claimed that CMC could provide an ideal learning environment for Korean learners who have some affective barriers to English language learning.

As more native speakers of English were put into the English classrooms, more awareness for producing English similar to native speakers has grown among the non-native English teachers. To encourage using the target language in the classroom, teaching English through English (TETE) has become popular mode of teaching. Although TETE has been verified as effective, some Korean students' affective barriers have not been overlooked. There are also some studies on TETE conducted in higher education setting (Jong-Bai Hwang, 2002; Eun-Ju Kim, 2002, 2003; Yo-An Lee, 2006). Jong-Bai Hwang (2002) looked into anxiety and achievement motivation of Korean college students enrolled in a content-based course for a period of semester. Hwang's (2002) study examined how the students' anxiety and motivation would change over time by using anxiety and motivation questionnaires (FLCAS & AMQ, respectively). The study found that the students' anxiety decreased and that their utilitarian goal orientation increased over time.

Chomsky (1988) stresses the importance of activating learners' motivation and claims that "the truth of the matter is that about 99% of teaching is making the students feel interested in the material" (p.181).

Schumann (1997) also asserts in his book, *The Neurobiology of Affect in Language*, that second language acquisition is primarily emotionally driven

and emotion underlies most, if not all, cognition. Research on motivation (Covington, 1999; Heckhausen, 1991; Heckhausen & Kuhl, 1985; Husman & Lens, 1999; Karniol & Ross, 1996) has incorporated time elements into the motivational research paradigm.

There are numerous factors involved in the second (foreign) language acquisition. The effective methods that take affective barriers of the EFL learners into account may be desirable for language teaching.

As a measure for improving English competence of Korean teachers of English so that they can be fluent English teachers with all four skills, the Ministry of Education plans to include an intensive oral interview and writing test as part of the teacher certification exam beginning in 2009.

## 2.2 Presentation in English

Even though efficient ways of teaching the language are adopted in the classroom, learners still need to produce their language (Swain, 1985, 1995; Swain & Lapkin, 1995). It obviously means that learners ought to use the language in the target form. Once they do, as pointed out in Izumi & Bigelow (2000), Pica, Holliday, Lewis, & Morgenthaler's (1989) study, they become more observant not only in meaning of language but also in forms.

While presentation may be a stressful task for learners who are not yet readily equipped with the skills of preparing, it can be an effective tool for upper intermediate to advanced level learners.

As shown in Choi et al's (2008) study that was done on developing English discourse competence through self-directed practices, the results showed the subjects became more confident in the language in the process of preparing for the weekly presentation session.

The study explores the possibility of utilizing student presentation in classroom whether the upper intermediate to advanced level students benefit from preparing for class presentation and learn on their own in terms of new vocabulary, expressions, and knowledge in English as well as practice pronouncing unfamiliar words, where to pause, and even memorize the content of their presentation.

## 3 Method

### 3.1 Participants

For the study a second-year English reading and writing class was observed. The total of 49 medical students were given a single topic for each at least two weeks prior to their presentation to prepare and present in class. The rationale for the two-week time was to let the students have sufficient time to do research to write a research paper and be ready for the class presentation. To encourage the students' active participation, the presentation was given 15 percent of the grade as medical students tended to

perform better when points or credit were given. Grading was based on creativity, content information and interesting presentation to attract the class members' attention. Also as for an incentive, excellent presenters were awarded at the end of each month, March through May, as well as the ones with outstanding content.

The length of the presentation was set for 10 minutes each for 4 students a week. At the beginning of the semester, the length of the presentation was less than 10 minutes. But in a couple of weeks, the students seemed more absorbed in showing various visual materials from the Internet, UCC, and the Power Point, thus presenting in a more interesting way. Consequently, the presentation time gradually became longer. Some were enthusiastically applauded by their peers after the presentation and the class atmosphere was lively with more questions and answers afterwards. A few students mimicked the instructor's lecturing style that brought a good laugh and some came up with witty remarks and answers at the end of their session. By the end of each presentation, the class seemed relaxed and contented.

### 3.2 Data Analysis Procedure

The topics were selected by the instructor based on the reading materials and current issues. Eighteen topics were given in March and the lists are as follows: Picasso, Taean oil spillage, great CEOs, earthquake, pets, global warming, newspaper, broadcasting, psychology, Sigmund Freud, Louvre Museum, Da Vinci Code, real estate speculation, credit card, Oprah Winfrey, Nobel prize, autism and great American anchors. In April, 16 topics were given such as materialism, Socrates, cross-over music, Ecstasy (drug), foreign legendary musician, hip-hop music and young generation, recycling, Korean legendary musician, famous Korean anchor, great musician and actor, religion & science, stocks & funds, pesticide, FTA, great inventors, and ballet. Then in May, 15 topics were given and the lists are as follows: 3Ds, social belief, weather forecasting, Asperger's syndrome, Grameen Bank & Yunus, pneumonia, ideal government, genetics, great choreographer, impressionism, obsession, maniac, diet, and pollution. From these lists, the students selected the one they were interested in. After the presentation, the students' papers were collected and reviewed. Several noticeable things were the marks indicating where to pause, stress, and how to pronounce certain words on the paper. It showed that the students did research, wrote papers in accordance with a format, and practiced to present it in class.

During the month of March, 18 students presented their papers, and 16 in April and 15 in May, respectively. They were also given a questionnaire on the presentation at the end of the semester. Based on the content of the presentation and how well they presented it, two presenters were selected each

month to be praised with an award written in English. Each presentation paper was then reviewed and counted how many words and sentences were used. The results of the questionnaire were also analyzed.

## 4 Results and Discussions

There were 20(40%) female and 29(60%) male students. They took the TOEIC by the end of the first year and their results showed that they were categorized as upper intermediate to advanced students with the mean score of 793 (806 for female and 780 for male students). The lowest was 715 and the highest 990. Table 1 and 2 are omitted.

Answers to the questionnaire on whether the presentation was interesting and useful and whether the presentation on different topic rather than the same-topic-for-all student presentation was more interesting, 40 students (92%) gave favorable answers and only 3 students (6%) were not sure. As shown in Table 4, majority of the students (80%) agreed that it was more effective than face-to-face lecture-only class.

Upon asked whether preparing for the presentation helps acquiring new vocabulary, expressions and knowledge in English, 30 students (69%) answered favorably.

## 5 Conclusion and Implications

Acquiring the English language takes time and effort on both teachers and learners. At the same time, students need to be motivated to learn the language. With an ever-more emphasis on the communicative aspect and other skills in the language, college English should include what students can benefit from.

By observing a class conducted with student presentation on a single topic for each student, it was learned that the students could outperform in preparing their own presentation using the Power Point and UCC to create an interesting, creative, and informative presentation. For example, one student with a topic of global warming approached a rather heavy topic lightly showing cartoon clips, comic strips, and pictures to get the message across, and therefore, her presentation was easily connected to the class. Other topics such as the Korean B-Boys, and famous entertainers prepared the UCC clips of actual performance and such various audio-visual materials were indeed entertaining. The students actively participated by asking numerous questions. Also a student whose topic was Oprah Winfrey did a thorough research and delivered a very strong message on poverty, determination, and fame.

However, the study has some limitations: The students' improvement was not tested or measured in the current study. It was a descriptive study based on the presentation paper and student questionnaire. Also since students' presentation skills are

considered a prerequisite for such an English presentation, implementing it may be limited. Some students were extremely shy and their eye contact and/or interaction with other students were non-existent. Even though the substance of the presentation was good, such a presenter made the class boring and students not pay attention.

Overall, the reading and writing class otherwise can be somewhat boring, turned out to be a very lively and fun activity having student presentation. It was learned that upper intermediate to advanced level students benefit from such a class. These students felt that learning through presentation is effective and helpful.

## References

- Choi, S., Joh, J., & Lee, Y. (2008). Developing English discourse competence through self-directed practices of non-native English teachers. *English Language & Literature Teaching*, 14(2), 25-46.
- Chomsky, N. (1988). *Language and problems of knowledge*. Cambridge, MA: MIT Press.
- Covington, M. (1999). Caring about learning: The nature and nurturing of subject-matter appreciation. *Educational Psychologist*, 34, 127-136.
- Heckhausen, H. (1991). *Motivation and Action*. New York: Springer.
- Heckhausen H., & Kuhl, J. (1985). From wishes to action: The dead ends and short cuts on the long way to action. In M. Frese & Sabini (Eds.), *Goal-directed behavior: The concept of action in psychology*. Hillsdale, NJ: Lawrence Erlbaum.
- Hong, S. (2003). Discussion ability development between CMC and on CMC group. *English Teaching*, 58(2), 227-248.
- Husman, H. & Lens, W. (1999). The role of the future in students motivation. *Educational Psychologist*, 34, 113-125.
- Hwang, Jong-Bai. (2002). L2 learners' anxiety and motivation in an English-only content-based class. *English Teaching*, 57(1), 193-211.
- Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition? *TESOL Quarterly*, 34(2), 239-278.
- Karniol, R. & Ross, M. (1996). The motivational impact of temporal focus: Thinking about the future and the past. *Annual Review of Psychology*, 47, 593-620.
- Kelm, O.R. (1992). The use of synchronous computer networks in second language instruction: A preliminary report. *Foreign Language Annuals*, 25(5), 441-454.
- Kim, Eun-Ju. (2002). Voices of learner from English-medium lectures. *Journal of the Applied Linguistics Association of Korea*, 18(1), 145-163.
- Kim, Eun-Ju. (2003). A comparative study of academic oral interaction in English-medium lectures and Korean-medium lectures. *English Teaching*, 58(3), 3-20.
- Kim, S. (2002). Korean college students' reflections of English language learning via CMC and FFC. *Multimedia-Assisted Language Learning*, 5(2), 9-28.
- Kim, Sung-Yeon. (2008). Five Year of Teaching English Through English: Responses from Teachers and Prospects for Learners. *English Teaching*, 63(1), 51-70
- Lee, Yo-An. (2006). Speaking proficiency of non-native English teachers: Construction multiple utterances. *English Teaching*, 61(3), 189-212.
- Lee, Y., & Chang, J. (2001). CMC-based materials design to improve learners' communicative interaction skills. *Multimedia-Assisted Language Learning*, 4(1), 20-56.
- Park, J. K. (2004). Needs analysis for developing a culture-focused teacher training course. *English Language Teaching* 16(2), 73-100.
- Pica, T., Holliday, L., Lewis, N., & Morgenthaler, L. (1989). Comprehensible output as an outcome of linguistic demands on the learner. *Studies in Second Language Acquisition*, 11(1), 63-90.
- Schumann, J.H. (1997). *The neurobiology of affect in language*. Malden, MA: Blackwell.
- Shim, R. (1994). Englishized Korean: Structure, status, and attitudes. *World Englishes*, 13(2), 225-244.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principles and practice in applied linguistics: Studies in honour of H.G. Widdowson* (pp. 125-144). Oxford: Oxford University Press.
- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16(3), 371-391.
- Warschauer, M., & Kern, R. (2000). *Network-based language teaching: Concepts and practice*. Cambridge: Cambridge University Press.

# A Japan Appropriate Teaching Model for Adult Learners

William W. Baber<sup>1</sup>

<sup>1</sup>Ritsumeikan University

willbaber@yahoo.com

## Abstract

In this paper I describe the pilot implementation of a classroom approach that is familiar to daily Japanese culture, unlike practices common to Communicative Approach (CA) teaching. This approach, drawn from Japan's business world, may provide an antidote to the difficulty CA has faced in Japan. The pilot implementation involved small groups with a clear boss (*kachou*) responsible to the teacher for class functions as well as a short presentation and discussion. The model differs from CA group activities in that students are not peers – the *kachou* is hierarchically above the group members and the teacher above the boss. The resulting in-group pressure and group desire to achieve as a group are integral to the classroom's motivation and management. The aspects in which this approach fits and conflicts with CA are discussed.

## Keywords

Teaching approach; Japan higher education; TEFL; cross cultural; *kachou*

## Introduction

Japan is home to a significant population of English teachers, about 5,000 in the Japan Exchange and Teaching program and almost 5,000 in the recently collapsed NOVA school and untold thousands more working in higher education, industry and other settings throughout Japan. These teachers often employ Communicative Language Teaching, or Communicative Approach (CA), in which students are encouraged to vocalize and interact in student centered situations (Ellis, 1996; McDonough and Shaw, 2003). In this approach, students vocalize target language as they are presented with it and in practices thereafter. As the students progress, they are expected to refresh and hone their abilities through repeated use, including especially verbal efforts. Yet students in Japanese classrooms generally avoid risky actions such as vocalizing unfamiliar phonemes, words, and phrases (Collins, 1999).

Despite the establishment of CA practices in

around the world, Japanese students do not appear to be excelling and indeed may be falling behind their counterparts in neighboring countries according to widespread perceptions.

Following this apparent failure, some consideration of alternatives to CA is in order. This report asks if there are Japanese teaching idioms that are more appropriate and culturally familiar to Japanese students and explores the application of one approach widely used in Japan business world.

## 1 Literature review

### 1.1 An inherent weakness?

CA often founders in the classrooms of Japan (McVeigh, 2002; Sato, 2002; Usuki 2007). The cultural appropriateness of CA has been called into question by Ellis (1996) who found that CA must be adjusted to Far Eastern classrooms, and additionally, accepted. The issues to be overcome are not only for learners, but also teachers as pointed out by Kojima and Kojima (2005). In their writing, Kojima and Kojima conclude variously that “teachers should...” and “students will have to...” however they do not provide a cross cultural comparison or guide to how these shifts in behavior might occur. Collins (1999) suggests that the difficulties with CA in Japan arise from incompatibility with usual teacher roles, cultural values and other key points. Specifically, Collins refers to cultural barriers towards risk taking, in-class verbalization, challenging teachers and other elements vital to CA. Nonetheless, it seems that students and educators cannot reliably overcome these barriers. Other barriers to CA arise from the overpowering effect of the university entrance exams as students and teachers conform to that variety of English (Wada, 2002) as well as insufficient teacher training (Sato).

In summary, it seems that CA is inherently flawed due to low cross cultural compatibility, relying on deep reaching mediation (Ellis, 1996) and the limited ability of students and teachers to negotiate acceptance of the approach in the classroom.

## 1.2 Corrections, alternatives to CA

It may be that as Ellis (1996) and Bax (2003) suggest, CA can be improved, corrected or otherwise fixed. On the other hand, there may be alternatives that can be fruitfully employed with some classes and groups in higher education or other settings.

Corrections and mediations may bear significant fruit in the long term, but are not easily attainable in the short term. Therefore this paper will examine other alternatives.

In their 2005 article on indigenous traditions suitable for TEFL, Shin and Crookes suggested a search through historical and recent training patterns indigenous to Korea. Importing their idea, the subsequent discussion will consider a few Japanese teaching approaches found in modern day Japan. The first goal is to consider approaches that are familiar to learners, culturally accepted, and generally compatible with TEFL teaching goals and practices. Thus, we will avoid any discussion of teaching practices found in ancient Japan as these could be as alien to contemporary students as an approach imported from Mars. Further, this paper will not discuss the traditional training of *deshi*, sumo, or the sports world from all of which examples are to be learned or avoided.

### 1.2.1 The Japanese business world

The business world of Japan, however, provides a widely known and broadly experienced training process. After an initial large-group training period, new staff are absorbed into small groups led by an experienced *kachou*, or boss. The *kachou* and experienced workers as well as new staff sit in a compact group with clear sight and sound contact. Mistakes and successes are clear to all in the group. The new person observes closely the detailed activities of others and is guided as needed by the boss. New group members feel pressure from the boss and from peers to learn quickly and perform well. In common with other indigenous Japanese training models is the clear hierarchy among learners and distance to the trainer.

With some modification, this model is the subject of this study.

## 2 Methodology

For this study, the *kachou* approach was piloted with four classes, two large and two small, of business students at a private university in western Japan. There were some variations in application among the classes, but the overall experience was similar for all.

Outcomes of the classes were collected based on student feedback in questionnaires and open

ended writing, teacher observations, and teacher self-reflection.

## 2.1 Implementation

The structure of the course was explained to students and the daily pattern continued through eight weeks. Anonymous feedback was collected at mid point and end of course from the 2 smaller classes. Non anonymous written comments were collected from the larger and smaller classes at various points throughout the semester. The smaller classes received instruction in use of English for managers including handouts to guide their speaking. Bosses were assessed for presentation to their group as well as for managing discussion with the group. Group members in the two larger were not assessed, but group members in the in two smaller classes were assessed.

The classes that included the *kachou* approach repeated the same pattern and timing eight of the 15 weeks in the semester. After warm up activities, the students formed the same groups with a different boss each week. The *kachou* delivered a short presentation (2-4 minutes), being graded by the teacher on the second delivery. Thereafter the *kachou* lead the group in discussion about the presentation topic for several minutes, also being graded. The subsequent phase saw the *kachou* directing the students in various practices from pronunciation to writing to workbook materials. Between each phase, the bosses were called out of the groups to form a huddle with the teacher. In the huddle, bosses received brief evaluations, instructions, and materials for the students. The huddles were a time of intensive teacher-student interaction.

Between huddles, the students worked as directed by the bosses with the teacher walking the room to see that work was progressing. Where there was a problem, the teacher did not interact with the small group members, but only the boss. After the problem was identified and briefly discussed with the boss in front of the group, the boss was left to resolve it. Bosses could resolve misunderstandings, errors or even student misbehavior such as chatting as expedient by their own devices.

Boss directed activities usually took one half to three-quarters of class time. After completing activities directed by the boss, the class pursued other activities from independent work to whole group activities. The boss position was transferred weekly and each student was a boss two times.

## 2.2 Student data

Data collected from students included open ended non-anonymous responses to the teacher made in all four classes in this study. Typically these comments ranged from a few words to two sentences and covered all aspects of the classroom from content to points unrelated to the classroom work. The teacher responded to these comments briefly and individually. The two smaller classes completed anonymous questionnaires in addition to the above mentioned open ended comments.

Data include grades given for group discussions led by the student boss. The discussion grades reflect the quantity of speaking in the group discussions as well as the quality of those discussions in terms of target language, interaction among participants, and the boss's management of the discussion. Discussions were graded by the teacher using scorecards and the grades were then transferred to a database. These grades indicate improvement over time during the semester. The improvement included quality of speaking by all participants and use of target language by boss and participants.

These data are necessarily subjective assessments by the teacher, despite being based on a rubric and thus represent a flaw in this study.

## 2.3 Teacher data

The data originating from the course teacher are self-reflective impressions. As such these impressions cannot be rigorously verified or compared with impressions or information from other staff or researchers. Future study should undertake analysis based on a multiple class and multiple teacher approach in order to successfully compare and interpret teacher impressions.

# 3 Discussion

The following discussion covers the outcomes of the *kachou* approach, then compares the *kachou* and CA approaches and finally considers their compatibility.

## 3.1 Outcomes based on student data

The outcomes of the approach as implemented were satisfactory to the teacher and to students in several respects, but lacking in other areas. Satisfactory results included the perception by students that the boss job was tough but doable. These sentiments are reflected in un-elicited written comments like these: "Be boss was tough!" and "I work so hard." as well as "I talk a lot in English." These comments were collected from non-anonymous open ended written communications with the teacher. Further outcomes based on anonymous student

questionnaires will test the following null hypotheses:

- students had no clear opinion positive or negative about the *kachou*-approach as suitable for Japanese students;
- being boss was neither a positive nor negative experience;
- group members did not provide or feel pressure on their members to perform;
- group members did not support their boss's efforts to be successful.

Other questionnaire items explore impact of the *kachou*-approach on student motivation, autonomy, quality of classroom experience, and success in learning from other students. Additionally, the questionnaire explores student impressions of acting as a boss, a hierarchical position, in relation to their fellow students.

## 3.2 Outcomes based on teacher observations

Grades, collected chronologically, showed improvement on the part of all students, whether acting as bosses or participants. Grading was done by the teacher according to a rubric which the students had in hand, however grading is inseparable from the teacher's perceptions of the moment, and necessarily was not objective.

More general teacher observations provided the following picture of authority and dynamics in the group. The bosses were supported in their efforts to direct and correct group members partly by the authority of simply being named the boss that week. Moreover, even shy and ineffective bosses were supported in their efforts by the small group members. The very members who were directed by the boss often took upon themselves the effort of ensuring that group members participated actively to complete their tasks. In short, the group developed its own pressure to succeed even if the boss was weak and applied this pressure to group members to continue activities, change behavior, or grapple with assignments.

## 3.3 Compatibility of *kachou* and CA

Understandably, teachers steeped in the ideas and use of CA will be cautious about reaching for a new system. However, the two approaches are not mutually exclusive and indeed overlap in many of the goals and actualities of the classroom. Overlap occurs as both approaches strive for meaningful communication, real life context, sense of achievement, student centered activities, intense s-t

interaction, enjoyment, effort to achieve task, sense of group, and socio-linguistic competence.

One key omission in the *kachou* approach is the direct pressure from the teacher toward students. This absence is because the effort to have students complete assignments, stay on task and the like is transferred to the boss and as well as the group itself.

Further there are four points that are more evident in the *kachou* approach including the pressure among students alluded to above.

Non-verbal communication is not encouraged in standard CA classrooms and, in a verbalizing environment, may be discouraged. The *kachou* approach, however, allows this kind of knowledge transmission. Tacit knowledge (Polyani, 1967) about how to complete tasks, behave, and interact thus are transferred without the direct intervention of the teacher and without codification.

While close observation skills are not discouraged in the CA classroom, the *kachou* seating plan makes the most of this behavior which seems omnipresent in Japanese classrooms, offices, and even public spaces.

Finally, the group develops a sense of responsibility towards each other. As suggested previously, the group develops a desire to manage the activities successfully together. The result is not only pressure on group members who underperform or misbehave, but also mutual supportiveness that results in learners helping each other to build competence and communication skill.

## 4 Conclusions

Comparison suggests there is a strong affinity between the Japanese business world's training model and CA as it is generally practiced. The *kachou* model as presented here is modified somewhat to fit classroom needs in terms of space and number of learners. Because the group dynamic is based on pressures and desires for achievement common to groups in Japan, the approach does not stumble against disinclination towards verbal performance as does CA. Additionally, the teacher is less burdened with motivating the students because the group dynamic and hierarchy play to motivation patterns already established in Japan.

### 4.1 Critique of *kachou* approach

While the *kachou* approach provides motivation, structure, and familiarity, it does not strongly support the language and behaviors associated with the culture of the English language. To some extent this weakness is ameliorated by discussion format and language content. The discussion

practice requires frank verbal interaction in a non-hierarchical and therefore non-Japanese business setting. The phrases and language provided for use in the discussion, especially by the boss, are typical of an idealized supportive and constructive English speaking boss. Ultimately these changes represent a modification of the *kachou* model to fit the needs of English speakers.

Next steps in researching and adapting the *kachou* model to the EFL classroom include more detailed review of existing knowledge about the *kachou* model in businesses. Furthermore, a study with two groups of classes, *kachou* and non-*kachou*, covering the same material and taught by different teachers remains to be designed and undertaken.

## References

- Bax, S. (2003). The end of CLT: a context approach to language teaching. *ELT Journal*. 57,3, 278-287.
- Collins, S. (1999). Communicative methodology: A health warning? *Explorations in Teacher Education*. 7, 1.
- Ellis, G. (1996). How culturally appropriate is the communicative approach? *ELT Journal*. 50,3, 213-218.
- Hall, J. (2005). What are the obstacles to communicative language teaching in Japan and what can we do to remove them? 岩手大学英語教育論集. 7.
- Kojima H. & Kojima, Y. (2005). Teacher Roles in Learner-Centered Communicative EFL Instruction. *Bulletin of Faculty Education Hirosaki University*. 94, 59-72.
- McVeigh, *Japanese higher education as myth*. Armonk, NY: M.E. Sharpe.
- Polyani, M. (1967). *The tacit dimension*. Gloucester, MA: Peter Smith.
- Sato, K. (2002). *Practical understandings of communicative language teaching and teacher development*. In *Interpreting Communicative Language Teaching*. Ed. S. Savignon. New Haven: Yale University Press.
- Shin, H. & Crookes, G. (2005). Indigenous critical traditions for TEFL? A historical and comparative perspective in the case of Korea. *Critical inquiry in language studies*. 2,2, 95-112.
- Usuki, M. (2001). *The learning process model and autonomy*. (ERIC Document Reproduction Service no. ED 455693).
- Wada, M. (2002). *Curricular innovation in Japan*. In *Interpreting Communicative Language Teaching*. Ed. S. Savignon. New Haven: Yale University Press.

# Construction and implementation of automatic L2 speech evaluation system

Yusuke Kondo<sup>1</sup> and Michiko Nakano<sup>2</sup>

<sup>1</sup>Language Education Center, Ritsumeikan University, <sup>2</sup>School of Education, Waseda University

ykondo@fc.ritsumeai.ac.jp, nakanom@waseda.jp

## Abstract

This study reports the construction, the implementation, and the evaluation of a web-based L2 speech evaluation system. The system implements the automatic speech evaluation in the following procedure; examinees read out the text, “The North Wind and the Sun” on their client computer, and their audio files are sent to the server computer which analyzes the audio files and replicates the evaluation scores given by human raters based on a prediction formula, and then, examinees receive their scores on their client computer. Because the system recognizes speech data based on Hidden Markov Model, certain amounts of speech data are required to train the model and obtain accurate results of speech recognition. Furthermore, learners’ speech data need to be evaluated by human raters. Therefore, the speech data of 101 Asian learners of English were collected, and their speeches were evaluated by human raters, and the correlations were examined between the examinees’ speech characteristics and their evaluation scores (Nakano, Kondo, and Tsutsui; 2008), because the system replicates the raters’ evaluation using learners’ speech characteristics. The read-out speeches given by the Asian learners were divided into three groups; high, mid, and low, according to the evaluation scores analyzed based on Neural Test Theory (Shojima, 2007). In this system, new examinees are given their evaluations based on the two speech characteristics of theirs. Their evaluations were done according to the average distance to the prototype of each category. In the evaluation of the automatic evaluation system, the degree of rating agreement among human raters and the automatic evaluation system was examined, and substantial agreement was obtained.

## Keywords

Second language speech, Reading aloud, automatic evaluation system

## Introduction

This paper introduces the construction, the implementation, and the evaluation of automatic Second Language (L2) speech evaluation system. In this system, the evaluations to be predicted were given by the trained raters, the raters and the evaluation items were scrutinized based on Multifaceted Rasch Analysis, and the learners’ speech characteristics to be analyzed were selected from the measures adopted in Applied Linguistics. The text which examinees read out was a fable from Aesop, “The North Wind and the Sun”, which consists of 113 words, five sentences. The evaluations to be given to examinees are categorical evaluation scores: A, B, and C. Therefore, the rankings were estimated based on NTT instead of the estimated ability in Multifaceted Rasch Analysis. As Shojima (2008) mentioned, a test is not an accurate tool to measure examinees ability. It is not capable of measuring examinees’ ability in interval scale, but it is capable of measuring examinees’ ability in ordinal scale at its best. For these reasons, the understandability of examinees and the capability of the test, categorical evaluations were adopted in this system.

In this paper, firstly the data used to build the system are described. Secondly, the scoring procedure, the speech recognition technology adopted in this system, the structure of the system, and the test-taking procedure are described. Lastly, the evaluation of the system is discussed.

## 1 Data

101 Asian learners of English participated in the recording of reading-aloud passage. They were forty Japanese, seventeen Chinese, nineteen Korean, six Filipino, ten Thai, four Vietnamese, and one Indonesian. They were either undergraduate or graduate students. Table 1 shows the key information of the participants. All the recordings were made in soundproof rooms in the universities which the participants belonged to. The participants were called in the room and given the instruction of

recording individually. In the recording, the participants gave their self-introduction to an interviewer and read out the text. After the recording, the participants were given a small gift for their participation. It took about ten minutes for each participant to complete the recording.

The reading text, was a fable from Aesop, “The North Wind and the Sun”, which is so famous that the students at university level must know it. This passage was also used in the NIE corpus (Deterding and Ling, 2005), and is used in the phonetic description of the International Phonetic Association.

Table 1: Key information of the participants in reading-aloud speech data

	<i>M</i>	<i>SD</i>	Range
Age	23.46	4.42	20
Study of English (year)	11.88	5.41	29

Six raters joined this evaluation, who participated in the rater training where they discussed the speech characteristics of L2 learners and watched the video of learners categorized based on Common European Frameworks of References (CEFR). The reliability of the raters was examined in the self-introduction speech evaluation data (Nakano, Kondo, Tsubaki, and Sagisaka: 2008).

To evaluate the read-aloud passage, ten evaluation items were selected from the self-introduction speech evaluation. The six raters evaluated 101 learners’ speech with the fourteen items shown in Table 2. The raters evaluated the learners’ speech individually on an evaluation website.

Table 2: Evaluation items in read-aloud passage

Loudness	Sound pitch
Quality of vowels	Quality of consonants
Epenthesis	Elision
Word stress	Sentence stress
Rhythm	Intonation
Speech rate	Fluency
Place of pause	Frequency of pause

The items are originally from Yashiro, Araki, Higuchi, Yamamoto, and Komissarov. (2001), and each item was thoroughly reviewed in order to make the items suitable in the evaluation of read-out speech. Unreliable raters and items were deleted based on Multifaceted Rasch Analysis. To delete unreliable raters and items, the scores of “infit” was adopted as the index. The score of infit “provides the size of the residuals, the differences between predicted and observed scores (McNamara, 1996: 172). The acceptable range of infit is “the mean  $\pm$  twice the standard deviation of the mean score statistics” in the case where the population

exceeded thirty (McNamara, *ibid*: 182). In this analysis, neither raters nor items was not considered to be unreliable according to the criterion.

We investigated the relationship between the evaluation scores estimated by Multifaceted Rasch Analysis and the speech characteristics realized in read-aloud speech (Kondo, Tsutusi, Nakano, Tsubaki, Nakamura, and Sagisaka, 2007; Kitagawa, Kondo, and Nakano, 2007; Kondo, Tsutsui, Tsubaki, Nakamura, Sagisaka, and Nakano, 2007; Nakano, Kondo, and Tsutsui, 2008). The examined speech characteristics were number of nonlexical pause, number of silent pause, duration of nonlexical pause, duration of silent pause, mean length of run, number of syllable unneeded, pruned syllable per second, and the ratio of weak syllable to strong syllable. Through three pilot studies of correlation, we found the two independent predictors of the evaluation scores: the pruned syllable per second and the ratio of weak syllable to strong syllable. In the present study, the analysis was done by multi-regression analysis (stepwise method): the criterion variable is the evaluation score; and the predictor variables, the pruned syllable per second and the ratio of weak syllable to strong syllable. The significance of the model was verified ( $F_{(2, 98)} = 44.57, p < .01$ , adjusted  $R^2 = .47$ ). The correlation between the observed values and the predicted values is .69. Figure 1 is the scatter graph of the observed and predicted value: y-axis is the observed value and x-axis, the predicted value.

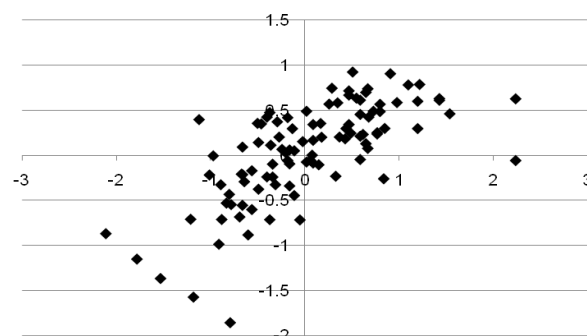


Figure 1. The Correlation Between the Observed Score and the Predicted Score

## 2 Automatic evaluation system

### 2.1 Introduction

Through the correlation studies between the evaluation scores and the L2 learners’ speech characteristics realized in read-out speech, we found two independent predictors of the evaluation score: the pruned syllable per second and the ratio of weak syllable to strong syllable. In the correlation analysis, the evaluation score we adopted was an interval scale, which was calculated based on Multifaceted Rasch Analysis. An interval

scale was easy to handle for researchers in the analysis of the test score, but it is difficult to interpret the score of interval scale for examinees and, as Shojima (2008) mentioned, a test is not an accurate tool to measure examinee's ability with interval scale. Therefore, in this system, we adopted categorical scores with three levels: A, B, and C, which were calculated based on Neural Test Theory (NTT).

## 2.2 Analysis of evaluation scores

The evaluation scores were analyzed based on NTT to estimate the examinees' levels. The automatic speech evaluation system to be constructed here is the system to replicate the evaluations given by the human raters. Considering the reliability in the human rating and the accuracy of replication by the system, it is reasonable and proper that examinees are grouped into three levels which correspond to the criterion given by CEFR: basic user, independent user, and proficient user in terms of pronunciation and prosody. In this analysis, the levels are set up to three, and the fit of the data to the model is examined.

The examinees were group into three groups: thirty six proficient users, thirty one independent users and forty four basic users. Table 3 shows the test fit indices in NTT. The indices shown below all indicates the good fit of the data to the model in NTT.

Table 3: Test fit indices in NTT

Index	Value
$\chi^2_{df=116}$	237.655
NFI	0.859
RFI	0.859
IFI	0.994
TLI	0.993
CFI	0.994
RMSEA	0.021
AIC	-146.323
CAIC	-731.250
BIC	-574.376

To examine the fit of data to the model in NTT, in addition to  $\chi^2$  statistics, several indices are available: the normed fit index (NFI), the relative fit index (RFI), the incremental fit index (IFI), the Tucker-Lewis index (TLI), the comparative fit index (CFI), the root mean square error of approximation (RMSEA), the Akaike information criterion (AIC), the consistent AIC (CAIC), and the Bayes information criterion (BIC). The model fit statistics depend on the number of levels which test developers set up. Hence, if we set up many levels such as thirty levels, the model fit statistics indicates better fit of the data to the model

compared to the case where we set up small number of levels such as three. However, Shojima (2008) mentioned that because a test is not reliable enough to detect fine differences among the abilities of examinees', the realistic range of levels is from three to twenty. Shojima (ibid) also mentioned that levels should be set up not only by the model fit statistics, but also based on test developers' experience and the practicality of the test.

## 2.3 The scoring procedure

In Figure 2, the ranked speech data of the learners examined are identified with the values of pruned syllable per second and the average ratio of weak syllable to the strong syllable. The x-axis indicates the value of the average ratio of weak syllable to the strong syllable; and the y-axis, the values of pruned syllable per second. The values of the average ratio of weak syllable to the strong syllable are inverted (the plotted values are 1 minus the original values) for the clear picture. Although some outliers and a multi-occupied area by all the three ranks are found, the areas of each rank can be specified to some extent. The averages of the two values were calculated in each category, and plotted. The bigger indicators are the averages (prototypes) of the two values in each category.

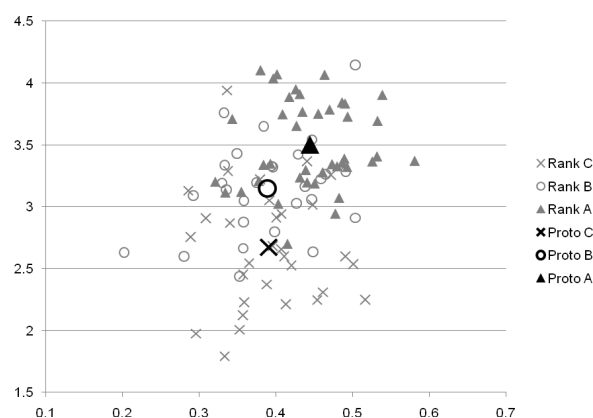


Figure 2: Scatter Graph for the Values of Pruned Syllable per Second and the Average Ratio of Weak Syllable to the Strong Syllable in Each Category

A new examinee's category is determined based on the Euclidean distances to the prototypes in each category. The distance to each prototype are calculated in the equation below:

$$D(x,p) = \sqrt{(p_1 - x_1)^2 + (p_2 - x_2)^2}$$

where  $p_1$  is the average of pruned syllable per second in a category, and  $x_1$ , the pruned syllable per second of a new examinee's; and  $p_2$  is the average of the ratio of the weak syllable to the strong syllable, and  $x_2$ , the ratio of the weak syllable to the

strong syllable of a new examinee's. Comparing the three distances of the new examinee's values to each prototype: Rank A, B, and C, the examinee is given the category of the nearest distance.

## 2.4 Speech recognition

To measure the speech features of examinees', the system adopts Hidden Markov Model Toolkit (HTK). HTK is a tool for Hidden Markov Modeling optimized for speech recognition. The procedure of the model training in HTK is depicted in Figure 3. Firstly, HTK codes the raw speech waveforms into sequences of feature. In this study, Mel Frequency Cepstral Coefficients were used.

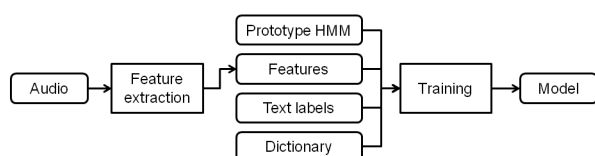


Figure 3: The Procedure of Model Training in HTK

In the model training, because HTK requires prototype Hidden Markov Model (HMM), text labels for the speech data, and the pronunciation dictionary, the text labels were created. The phonetic description was done based on Jones (2003). For the initial training, as the speech data TIMIT Acoustic-Phonetic Continuous Speech Corpus (Garofolo, Lamel, Fisher, Fiscus, Pallett, Dahlgren, and Zue: 1993) were used to train the HMM. For adapting the model to L2 speech of Asian learners' were used. In the process of the model training and adopting, HTK phone-aligns the target speech data based on the order of occurrence of the phones referring the text labels and the pronunciation dictionary. HTK runs through the model training several times to create robust HMMs.

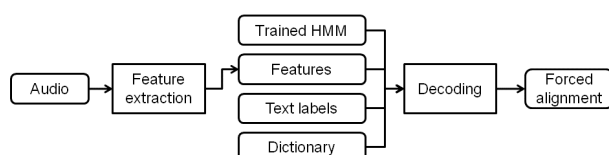


Figure 4: The Procedure of Forced Alignment in HTK

Figure 4 shows the procedure of the forced alignment. In this procedure, HTK phone-aligns the speech data based on the corresponding text labels.

## 2.5 Structure of the system

The automatic L2 speech evaluation system to be implemented here is a web-based system where examinees read out the passage, "The North Wind

and the Sun" on their client computers, the recorded speech data are transferred to the server computer where the data are analyzed, and the examinees receive the feedback from the server computer on their client computer. Figure 5 shows the procedure of this automatic evaluation.

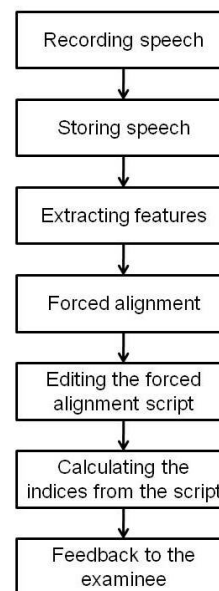


Figure 5: Procedure of Automatic evaluation

The system records the examinees' speeches using Java applet, JavaSonics ListenUp (Mobileer, Inc, 2008). The recorded speeches are transferred to the sever computer and are stored. Then the speeches are converted to the HTK format and are analyzed. The results of forced alignment are edited to calculate the two indices, pruned syllable per second and the average ratios of weak syllable to the strong syllable. Then, based on the two indices, the examinees' scores are calculated, and the feedback is sent to the examinees computer. All the processes are controlled by Perl scripts, including calling JavaSonics ListenUp and HTK processes. The processes of examinees' side are done with a web browser (e.g. Internet Explorer, FireFox, and Google Chrome).

Figure 6 shows the structure of the evaluation system. The recording procedure is done by Java applet stored in the folder, `codebase`, and the transfer and the retention of the speech file process is controlled by `upload_x.cgi`. The examinees record their speech sentence by sentence, and this process repeats five times. In each process, `upload_x.cgi` calls HTK to convert the speech file, phone-align the speech using the files stored in `hmm`. The converted speech file is stored in the directory `mfcc`; and the output file of the phone alignment, in the directory `out`. The output file of the phone-aligned speech is edited by `upload_x.cgi` and stored in the directory `lab`.

Lastly, `eva.cgi` calculates the two indices of speech characteristics and the distances to the prototypes stored in the information in the edited output file in the directory `out`. The program produces the evaluation and feedback to the screen on the client computer of the examinees'. The directory `img` contains the image file used in the webpage for examinees. The file `instruction.cgi` creates the instruction page; `recx.cgi`, each recording page; and `testrec.cgi` controls the test recording in the instruction page. The files `data.dat` and `ques.dat` store the information the examinees enter in the initial page.

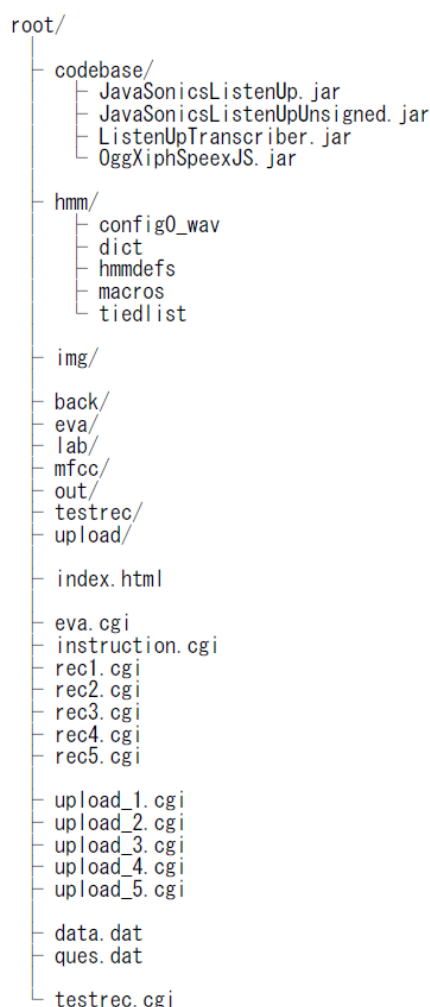


Figure 6: The structure of the evaluation website

## 2.6 Test-taking procedure

This section introduces the procedure of the automatic evaluation. Firstly, examinees access the evaluation website, enter their names, and answer the questionnaire. They submit their answers and go to the instruction page.

Secondly, in the instruction page, the examinees receive the instruction on how to take the test, and practice using the application for recording. The

whole passage to be read and the translation into Japanese are provided in this page. After the practice, they go to the recording page. In this test, they read out the passage, "The North Wind and the Sun" and record and submit their speech sentence by sentence. They record and submit their speech five times in total. Figure 7 shows the screenshot of the recording page.

Lastly, after the examinees complete their recordings of the five sentences, they receive their own evaluation and their own feedback according to their level which the system estimates.

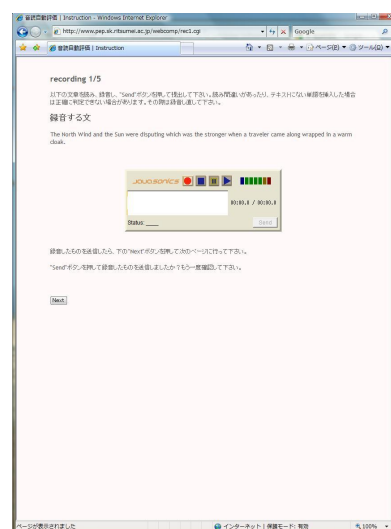


Figure 7: Recording Page

## 2.7 Evaluation of the system

### 2.7.1 Participants

Participants of this study were twenty Japanese learners of English and three raters. The raters were the Japanese language teachers of English who received rater training according to CEFR. The learners were Japanese undergraduate students.

### 2.7.2 Methods

The raters evaluated the twenty learners' speeches according to the criterion, CEFR and gave the categorical evaluations: A, B, and C. We investigated the degree of agreement in the evaluation among the human raters and the automatic evaluation system based on Fleiss' kappa.

Fleiss' kappa is a measure of inter-rater reliability for assessing the degree of agreement when more than three raters evaluate performance using evaluation items with a fixed number of categories (Gwet, 2001). The interpretation of this index is somewhat controversial, because it depends on the number of raters, categories, and examinees. Table 4 shows the interpretation of Fleiss' kappa provided by Landis and Koch (1977). Though this table is not widely accepted, it can be adopted as one of the criteria in the examination of the degree of

agreement.

Table 4: Interpretation of Fleiss' kappa

$\kappa$	Interpretation
< 0	Poor agreement
.21 - .40	Fair agreement
.41 - .60	Moderate agreement
.61 - .80	Substantial agreement
.81 - 1.0	Almost perfect agreement

### 2.7.3 Results

Table 5 shows the Fleiss' kappa of this evaluation. We calculated the indices four times. In each time, one of the raters was excluded. By comparing these indices, the rater who lowers the degree of agreement can be detected. For example, the kappa in the second row indicates the rater agreement among Rater 1, 2, and the automatic evaluation system, excluding Rater 3. The kappa in the lowest row indicates the rater agreement among all the raters: Rater 1, 2, 3 and the automatic evaluation system.

Table 5: Fleiss' kappa

Raters	$\kappa$
Rater 1, 2, and AES	.70
Rater 1, 3 and AES	.60
Rater 2, 3, and AES	.60
Rater 1, 2, and 3	.75
ALL	.66

Note. AES stands for automatic evaluation system.

## 3 Discussion and conclusion

In this study, we reported the construction, the implementation, and the evaluation of the automatic L2 speech evaluation system. 101 speeches given by Asian learners of English were analyzed, and the trained raters evaluated their speeches based on CEFR. Then, the relationship between the speech characteristics and the evaluation scores was investigated. The system was built based on the relationship revealed in this study.

As for the evaluation of the system, Fleiss' kappa was adopted as the index of rater agreement. Though the agreement was the highest when the system was excluded from the raters, we obtained substantial agreement between the human raters and the system.

Perfect agreement is difficult to achieve in the performance assessment, as was indicated the Fleiss' kappa among the three human raters (.75), but the scoring procedure, the test-taking procedure, and the correlation studies should be reviewed to improve the system.

## References

Deterding, D., & Ling, L. E. (2005). The NIE

corpus of spoken Singapore English. Deterding, In D., Brown, A., & Ling, L. E. (Eds). *English in Singapore*. Singapore: McGraw-Hill Education.

Garofolo, J. S., Lamel, L. F., Fisher, W. M., Fiscus, J. G., Pallett, D. D., Dahlgren, N. L., & Zue, V. (1993). *IMIT Acoustic-Phonetic Continuous Speech Corpus*. Philadelphia: Linguistic Data Consortium.

Gwet, K. (2001) *Statistical Tables for Inter-Rater Agreement*. (Gaithersburg : StatAxis Publishing)

Kitagawa, A., Kondo, Y., & Nakano, M. (2007). Does vowel quality matter? *Proceedings of 12th Conference of Pan-Pacific Association of Applied Linguistics*, 224-227.

Kondo, Y., Tsutsui, E., Tsubaki, H., Nakamura, S., Sagisaka, Y., & Nakano, M. (2007). Examining predictors of second language speech evaluation. *Proceedings of 12th Conference of Pan-Pacific Association of Applied Linguistics*, 176-179.

Kondo, Y., Tsutsui, E., Tsubaki, H., Nakamura, S., Sagisaka, Y., & Nakano, M. (2007). Examining predictors of second language speech evaluation. *Proceedings of 12th Conference of Pan-Pacific Association of Applied Linguistics*, 176-179.

Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics* Vol. 33, 159-174.

McNamara, T. F. (1996). *Measuring second language performance*. Essex: Pearson Education Limited.

Nakano, M., Kondo, Y., & Tsutsui, E. (2008). Fundamental Research on Automatic Speech Evaluation. *9th APRU Distance Learning and the Internet Conference--New Directions for Inter-institutional Collaboration: Assessment & Evaluation in Cyber Learning*. 207-212.

Nakano, M., Kondo, Y., Tsubaki, H., & Sagisaka, Y. (2008). Rater Training Effect in L2 and EFL Speech Evaluation. *The 8th Phonetics Conference of China and the International Symposium on Phonetic Frontiers*. 8 pages in CD-ROM Proceedings.

Shojima, K. (2008). Neural test theory: A latent rank theory for analyzing test data. *DNC Research Note*. 08-01.

Yashiro, K., Araki, A., Higuchi, Y., Yamamoto, S., & Komissarov, K. (2001). *Ibunka communication workbook*. [A workbook for cross-cultural communication]. Tokyo: Sanshusha.

# "Lion Heart's Legacy": Personification of the Japanese Prime Minister in Cabinet E-mail Magazines

Riikka Länsisalmi

Leiden University Institute for Area Studies & Centre for Linguistics

r.j.lansisalmi@hum.leidenuniv.nl

## Abstract

The purpose of this presentation is to examine overt 1<sup>st</sup> person reference in the English translations of these Cabinet E-mail Magazines and compare this phenomenon of "personification" with the original Japanese versions. It is common knowledge that Japanese subjects can often be ellipted and therefore no overt term is needed for example to refer to the 1<sup>st</sup> person. In English, however, overt reference usually becomes obligatory.

## Keywords

Overt reference, personification, ellipsis

## Introduction

Natsume Soseki 1907 (Ikegami 2008):

I have to confess that I find the device of personification simply unnatural and intolerable. (...) It is as if a monkey is trying to pretend to be a lord by putting a crown on its head."

It is common knowledge that Japanese agents can often be ellipted and therefore no overt term is needed for example to refer to the 1<sup>st</sup> person. In English, however, overt reference usually becomes obligatory. While in English elliptical styles in general seem to be linked to sociolinguistic values such as brusqueness, shared knowledge (solidarity), emphasis (power) and intimacy (Fowler 1985: 71), in Japanese, ellipsis usually occurs in what is referred to as unmarked situations. Ellipsis is likely to take place, when the referent can be identified through a grammatical construction, inferred from honorific mode or deictic expressions, understood through contextual cues or represents the (paragraph) topic. Person terms in Japanese comprise not only (so-called) personal pronouns, but also various other categories, such as (professional) titles, kinship terms, proper names, status terms, and so forth.

Against the background described above, this presentation presents a pilot study, which examines overt 1<sup>st</sup> person reference in English translations of a Japanese Prime Minister's Cabinet E-mail Magazine launched by Prime Minister Koizumi, also known as "Lion Heart", and

compares this phenomenon of "personification" with the original Japanese texts. I pay particular attention to overt 1<sup>st</sup> person reference terms in the original Japanese E-mail Magazines and examine how their English translations accommodate intended meanings. I intend to demonstrate that 1<sup>st</sup> person references may be extremely complex and political terms, sometimes raising interpretive issues of who or what is being represented. There are, therefore, puzzling cases of 'I' or – in particular – 'we'.

## 1 Data

In 2001 Japanese Prime Minister Koizumi launched a "Cabinet E-mail Magazine" to "express (his) thoughts and learn what (his) compatriots are thinking". This electronic service continues today as Prime Minister Aso's Cabinet E-mail Magazine and was adopted also by Koizumi's successors preceding Aso, namely Prime Ministers Abe and Fukuda. In addition to a Japanese version, the Abe Cabinet E-mail Magazine is currently available in English, Korean and Chinese. Interested readers can subscribe to the Magazine and have it delivered directly to their email addresses or read current and back number issues online.

The data of this pilot study consists of twenty issues of the Cabinet E-mail Magazine, including five issues from each Prime Minister. Character and word counts are indicated in Table 1.

**Table 1: Analyzed data**

Prime Minister	Analyzed Cabinet E-mail Magazine issue numbers	Character count in Japanese original	Word count in English transl.	Average character/ word count per issue
Aso	26-30	4,515	2,772	903/554
Fukuda	42-46	6,937	2,975	1,387/595
Abe	42-46	5,282	3,033	1,056/607
Koizumi	246-250	7,805	4,101	1,561/820
<b>Total</b>	20 issues	<b>24,539</b>	<b>12,881</b>	1,227/644

## 2 Analysis of data

Yamamoto (1999: 67) argues that the concept of animacy has a strong influence on the mind of any speaker in the process of language use and can

therefore be linked to varying “mind-styles”, a concept originally introduced by Fowler. Such mind-styles would characterize distinctive linguistic presentations of individual mental selves, hence allowing for example examination of the extent to which animacy influences the mind of English and Japanese speakers in their process of using these particular languages as native speakers. In this presentation the focus will be upon referential expressions in Japanese and English, two languages which are “expected to contrast strikingly in terms of epistemic attitude toward the expression of animacy and hence to represent two distinctively different types of ‘mind-style’ (Yamamoto 1999: 71)”.

The forms which animate expressions can take vary considerably, but this paper targets (ellipses and) overt expressions referring to the 1<sup>st</sup> person, including 1<sup>st</sup> person pronominal forms and common noun phrases. Analyzed forms are listed in Table 2. As case relations are expressed by postpositions in Japanese, only the encountered basic terms referring to the 1<sup>st</sup> person, 私 *wata(ku)shi* ‘I’ and 私たち *wata(ku)shitachi* and 我々 *wareware* ‘we’, as well as the archaic genitive *waga* (occurring only in わが国 *wagakuni* ‘our country’) are listed for Japanese.

**Table 2: Overt reference to the 1<sup>st</sup> person**

Prime Minister	<i>I</i>	<i>m y</i>	<i>m e</i>	私	<i>w e</i>	<i>ou r</i>	<i>u s</i>	私たち	我々	わが
Aso	34	7	0	9	16	1	1	0	1	6
Fukuda	32	8	4	6	14	3	2	2	0	2
Abe	78	39	2	14	13	9	1	3	0	1
Koizumi	93	39	6	15	26	15	1	1	0	0
<b>Total</b>	<b>237</b>	<b>93</b>	<b>12</b>	<b>45</b>	<b>69</b>	<b>28</b>	<b>5</b>	<b>6</b>	<b>1</b>	<b>9</b>
<b>Tot. 1<sup>st</sup> pers. English</b>	<b>44</b>									
<b>Tot. 1<sup>st</sup> pers. Japanese</b>	<b>61</b>									

As Table 2 demonstrates, Japanese allows predominant use of elliptical reference, whilst an exact overt 1<sup>st</sup> person term is usually required in English (61:444). Clearly this observation is nothing new, as the phenomenon of ellipsis – or non-occurrence – of overt reference is deeply engrained in the structure of Japanese. However, a closer look at the constructions, contexts and types of expressions in which these terms occur may prove to be more interesting. Since counting the exact number of ellipses vs. overt reference in Japanese texts or in comparison with English counterparts would contribute little to the discussion, it seems more illuminating to compare *who* or *what* is actually represented by terms

referring to the 1<sup>st</sup> person in the examined texts. In the context of political discourse, representing a nation and targeted at a wide domestic vs. foreign audience, these choices are likely to shed light on how national identities are constructed – or, at least, how attempts to do that may be undertaken. For example, the exclusive vs. inclusive nature of 1<sup>st</sup> person expressions can obviously be strategically employed to enhance national sameness or uniqueness, or, in other words, differences or similarities with other national collectives.

Slight differences in self-reference can be detected between the examined Prime Ministers. Proportionally, the English versions of Abe’s and Koizumi’s Email Magazines seem to employ slightly more singular 1<sup>st</sup> person agents (*I*) than those of Aso and Fukuda. Aso, on the other hand, appears to have a fondness for the fixed Japanese expression ‘our country’ (わが国 *wagakuni*), including a genitive marker, while other Prime Ministers would use the country name ‘Japan’ (日本 *Nippon/Nihon*), for example. In the Japanese originals, Fukuda tends to employ singular overt 1<sup>st</sup> person reference proportionally even less frequently than the others, but it should of course be kept in mind that such overt self-personification is, overall, rather limited (45 occurrences in total). It must also be noted that, due to the limited number of Email Magazines analyzed in this pilot study, the claims offered here should not be considered to be fully representative of the respective Prime Ministers.

To elaborate, the following section of this presentation concentrates on a comparison of overt 1<sup>st</sup> person referents in the original Japanese texts and their English translations.

## 2.1 Overt 1<sup>st</sup> person reference in original Japanese texts

Although the characteristic of 1<sup>st</sup> person reference to strongly encode sentiency and potential intentionality of the writer generally triggers ellipsis/non-occurrence in Japanese, certain constructions and expressions in the analyzed Email Magazines do retain an overt term, as became evident in the preceding section.

A first distinction could be made between clear structural contexts, which require the use of an overt term of reference, and those contexts, where the use of such terms is not (structurally) obligatory. Possessive constructions represent one type of grammatical context, where overt reference becomes compulsory. The analysis of our data further pinpoints the constructions displayed in Table 3 below, illustrated by the examples that follow. In these examples it is necessary to fill the ‘slot’ preceding the constructions represented in the table with some kind of overt term of reference.

**Table 3: Constructions with overt 1<sup>st</sup> person referents in Japanese texts**

Rank	Definition	n
1.	Topic marker は <i>wa</i>	25
2.	Genitive marker の <i>no</i>	13
3.	IO marker に <i>ni</i>	3
4.	Particle 'for' にとって <i>nitotte</i>	2
5.	Subject marker が <i>ga</i>	1
5.	Focus marker も <i>mo</i>	1
5.	Conjoining particle と <i>to</i>	1
5.	'Source' marker から <i>kara</i>	1
5.	Particle 'only' だけ <i>dake</i>	1
5.	'Status' particle として <i>toshite</i>	1
5.	Pron.+ refl. pron 自身 <i>jishin</i>	1
<b>Examples</b>		
国民の安全を守るため、 <b>私</b> は、世界各国と連携を密にしながら、引き続き、全力を尽くしてまいります。 <b>I</b> will continue to make my utmost efforts, while working in close collaboration with countries around the world, in order to defend the people's safety. (Abe27)		
<b>私</b> の好きな言葉に「龍蛇無鱗（りゅうにへびのいろこなし）」があります。 There is a saying that <b>I</b> like very much: "Dragons do not have snake scales." (Fukuda45)		
局面を転換しなければならぬ。これが <b>私</b> に課せられた責任であると考えました。 There has to be a change in the current situation, and <b>I</b> realized this was the responsibility that had been given to <b>me</b> . (Abe46)		
(...) 私の最後の外遊の地フィンランドの世界遺産でひろったこの石は、 <b>私</b> にとって思い出深い「宝石」になりました。 For <b>me</b> , this stone, the one I picked up at this World Heritage site in Finland, the place of my last official overseas visit, has become a deeply memorable gem. (Koizumi247)		
今振り返れば、厳しいご批判も、温かい励ましも、毎週寄せられるご意見は、 <b>私</b> が政策を進めるための大きな原動力でありました。 When I look back now, all the opinions I received each week -- whether they offered severe criticism or warm encouragement -- added up to a major driving force for <b>me</b> to promote policy. (Fukuda46)		
<b>私</b> も1日の静岡県伊豆市での防災訓練を皮切りに、3日には来日されたチリのパチェレ大統領と会談するなど、気持ちを新たに仕事に取り組んでいます。 <b>I</b> myself am undertaking my work on a new note too, beginning with participation in the disaster prevention drills held in Izu City, Shizuoka Prefecture on the very first day of this month and a meeting with visiting Chilean President Michelle Bachelet Jeria on the third. (Aso45)		

<p>ヴァンハネン首相は、<b>私</b>との会談の後、ヘルシンキ郊外にある、シベリウスが過ごし、数々の名曲を作曲した邸宅「アイノラ」を、自ら案内していただきました。</p> <p>After Prime Minister Vanhanen and <b>I</b> had our meeting, the Prime Minister himself kindly showed me around Ainola, Sibelius' home in the suburbs of Helsinki where he lived and composed countless masterpieces. (Koizumi249)</p>
<p>日曜日（10日）の昼過ぎから始まった会議の冒頭、<b>私</b>から、テロや核の不拡散、海賊などの国際的な広がりをもつ非人道的な問題に対して、国際社会が一致して取り組むことが必要であると訴えました。</p> <p>At the outset of the summit which started on the afternoon of Sunday, September 10, <b>I</b> stressed that the international community must engage in concerted efforts to tackle internationally widespread issues irreconcilable with the principles of humanity, notably terrorism, nuclear proliferation, and piracy. (Koizumi249)</p>
<p>山のように残された宿題と必死になって格闘した思い出は、<b>私</b>だけではないでしょう。</p> <p>I am sure <b>I</b> am not the only one who has memories of desperately struggling to complete a pile of homework at the vacation's end. (Abe42)</p>
<p>この枠組みには、すべての主要国が参加することが必要であり、<b>私</b>としても積極的に働きかけていきます。</p> <p>It is necessary that every major nation participate in the new framework. For <b>my</b> part, <b>I</b> will actively call on nations for their participation. (Aso28)</p>
<p>(...)などを最優先として、<b>私自身</b>や、中曽根外務大臣を含め、あらゆるレベルで外交努力を傾注し、行動しました。</p> <p>In response, <b>Japan</b>, including <b>I myself</b> and Foreign Minister Hirofumi Nakasone, took actions by making diplomatic efforts at all levels. (Abe27)</p>

As Table 3 illustrates, the topic construction, marked mainly by the topic particle *wa*, is the most likely candidate for an overt 1<sup>st</sup> person reference in the original Japanese E-mail Magazines with twenty-five occurrences. The topic-marker particle is generally used to indicate that the latter part of the sentence serves as a comment on the (sentence-initial) topicalized element, but it is also explained to reflect speaker involvement in organizing a narrative by highlighting a main character (Iwasaki 1993: 5).

The following table of examples hopefully sheds further light on the contexts in which the writer – supposedly the Prime Minister of Japan – of the examined E-mail Magazines has felt the need to make referential personification information overt as a 1<sup>st</sup> person topic, despite, in some cases, already being activated in the reader's mind.

**Table 4: Overt 1<sup>st</sup> person topic construction types in Japanese texts**

Category (representative predicates) + example	n
<b>1. To act/give instructions/make decisions (past/present [future] tense)</b> (するよう指示しました、決意をいたしました、参列いたしました、判断しました、訴えました、全力を尽くしてまいります、去ります) <u>私は</u> 、月曜日、総理大臣の職を辞する決意をいたしました。それは、国民の皆さんのための政策をより力強く進めていくためには、新しい体制を整えるべきであると考えたからにほかなりません。 On Monday of this week, <u>I</u> made the decision to resign from the post of Prime Minister. <u>I</u> did so because <u>I</u> believe that a new system should be put in place in order to proceed even more powerfully with policy for the people. (Fukuda46)	7
<b>2. To have/had a habit of doing/learn(ed) something (paragraph-initial)</b> (思うことがありました、思うときがあります、知りました、過ごしたものです) <u>私は</u> 、政策を立案する際、この「永遠の今」を思うことがありました。 <u>I</u> often used to think of "the eternal now" when drawing up policy. (Fukuda46)	4
<b>3. To place priority, emphasis (past tense)</b> (最優先として、優先しました、強調しました) 昨日、インドの国会で「二つの海の交わり」と題する講演を行いました。N「二つの海」とは、太平洋とインド洋のことです。講演の中で <u>私は</u> 、日本とインドは世界で最も可能性を秘めた二国間関係であることを強調しました。 Yesterday, <u>I</u> delivered a speech to the Parliament of India entitled "Confluence of the Two Seas." The two seas <u>I</u> referred to are the Pacific and the Indian Oceans. In the course of <u>my</u> speech, <u>I</u> stressed that the "Japan-India relationship is blessed with the largest potential for development of any bilateral relationship anywhere in the world." (Abe43)	3
<b>4. To think/believe (often expressing contrast with another entity)</b> (思っています) 環境対策、というとか「制約」というイメージを持つ人がいますが、 <u>私は</u> 、将来の飛躍に向けた大きなチャンスだと思っています。 Some people associate environmental measures with restrictions. However, <u>I</u> think of them as a great opportunity for <u>us</u> to leap forward in the future. (Aso28)	3
<b>5. Must/will (never) (+ neg.)</b> (決して忘れません、忘れてはなりません) この間にいただいた、みなさんの忌憚のないご意見、心温まる激励を、 <u>私は</u> 決して忘れません。 <u>I</u> will never forget the unreserved views and the heart-warming encouragement that were so generously extended to <u>me</u> during that period. (Abe46)	3
<b>6. Must/have obligation to/should (neg.)</b> (傾けなければならない、歩んでいかねばなりません、すべき) <u>私たちは</u> 、こうした地方の声に耳を傾けなければならない。そして、謙虚に受け止め、政策で対応していかねばなりません。今回、地方で知事を経験した増田さんに総務大臣として加わっていただきました。地方の生の声を、政策に反映していただくよう期待しています。 <u>We</u> must heed the voices of the people in local regions, humbly accept their opinions and respond to them with	3
policies. <u>My</u> expectation for Mr. Hiroya Masuda, who has joined <u>my</u> Cabinet as the Minister of Internal Affairs and Communications and has had experience as a prefectural governor, is to reflect these voices directly in Government policies. (Abe44)	
<b>7. To have responsibility/respect (progressive tense)</b> (責任を持っています、尊敬の念を抱いています) <u>私たちは</u> 、将来の世代に対して責任を持っています。人口が減少し、経済がグローバル化する中でも、日本が今後とも発展を続けるためには、厳しくとも、今、改革を前進させなければなりません。 <u>We</u> have a responsibility to future generations. As Japan's population declines and the globalization of the economy continues, <u>we</u> must resolutely push reforms ahead now, difficult though it may be, to ensure <u>our</u> nation's continued development in the future. (Abe44)	2
<b>Total</b>	25

Paragraph-initial topic constructions referring to past or habitual actions performed by a 1<sup>st</sup> person agent (category 2) typically introduce new discourse topics related to the protagonist and could be characterized as more “neutral” than other encountered topics. In fact, in most cases, overt 1<sup>st</sup> person topics appear to be linked to emphatic expressions; the protagonist is portrayed as a determined and strong-willed action-taker, who firmly moves forward taking difficult decisions, placing priority on matters of national interest, while remembering to remind his compatriots of national responsibilities in a manner appropriate to a true reliable statesman. In a number of contexts, a clear nuance of contrast with another entity can also be detected in the analyzed examples.

As Yamamoto (1999: 96) observes, when overt 1<sup>st</sup> person reference is used instead of ellipsis in Japanese, thus rendering the writer stand out as a highly ‘animated’ entity, utterances will possibly sound persistent or bearing a heavier responsibility for his/her claim. In other words, as indexicals, these pronouns can be interpreted to perform the function of linking speech acts to “those (...) responsible for their illocutionary force” (Mühlhäusler and Harré 1990: 25). Furthermore, in speech act theory, overt 1<sup>st</sup> person pronouns are usually considered to represent key elements for an utterance to be “performative” rather than “constative” (Iwasaki 1993: 1).

## 2.2 Plurality in 1<sup>st</sup> person expressions

Plurality can be employed to blur and impersonalize the identity of referents, thereby weakening the sense of animacy and directness encoded in overt reference (Yamamoto 1999: 99). As Table 2 above illustrates, in the examined E-mail Magazines English plural 1<sup>st</sup> person referents outnumber Japanese plural expressions by a ratio of 102:16 (including the archaic genitive *waga* わが). In this

section, let us first examine plurality in the Japanese texts.

## 2.2.1 Manifestations of 1<sup>st</sup> person plural reference in Japanese texts and their English counterparts

All seven occurrences of the Japanese plural 1<sup>st</sup> person referent ‘we’, (*watashitachi* 私たち 6 occurrences, *wareware* 我々 1 occurrence) have been rendered into English by a corresponding plural pronoun: *we* (5 occurrences) or *us* (2 occurrences). In three cases, ‘Japan’ (*Nippon* 日本) is overtly mentioned later in the same paragraph, thereby making it clear to the reader that *watashitachi* in that particular context represents the Japanese people. Another example, although without overt reference to ‘government’ (*seifu* 政府) or similar entities in Japanese, frames *watashitachi* as representatives of the government, further clarifying the referent for a non-Japanese readership by having recourse to terms such as *Cabinet* and *Government policies* in the English version. In an additional example, the original text, following an initial overt 1<sup>st</sup> person reference *watashitachi*, has ellipsis, which in turn is encoded as *Japan* in the English translation:

### Example (1)

しかし、**私たち**は、内向きな志向のとりこになることなく、心を開き、そのまなざしをしっかりと世界に向けながら、歩んでいかねばなりません。これから、平和協力国家として、世界の恒久平和の確立に向け積極的に活動します。

Nonetheless, **we** should move forward, indifferent to inward orientation and keeping **our** hearts open and eyes firmly focused on the world. As a Peace Fostering Nation, **Japan** will continue to actively work toward the establishment of lasting world peace. (Fukuda44)

All nine occurrences of the genitive *waga* (わが) in the Japanese texts are used in the fixed expression ‘our country’ (*wagakuni* わが国) and thus refer to Japan. Apart from one example, these expressions have been rendered into English using *Japan*. Interestingly, only in one case does the English version use a different expression, namely *the Government of Japan*:

### Example (2)

その深い反省の上に立って、戦後、**わが国**は、一貫して平和国家としての道を歩んできました。今日の平和と繁栄は、戦争によって命を落とされた方々の尊い犠牲の上に築かれたものにほかなりません。

Out of profound remorse, **the Government of Japan** has been committed to a path as a peaceful nation throughout its post-war history. The present peace and prosperity of **Japan** are built squarely on the precious sacrifices made by those who lost their lives to the war. (Fukuda44)

In this politically extremely sensitive context, the English translation conveniently encodes the entity experiencing remorse as the *Government* of Japan, rather than the Japanese nation.

## 2.2.2 Manifestations of 1<sup>st</sup> person plural reference in English texts and their Japanese counterparts

Our observation that overt plural 1<sup>st</sup> person pronouns occur frequently in the English E-mail Magazines invites a further investigation of the original Japanese texts as contrasted with their translations: Who or what does the English ‘we’ represent and how is this interpretive entity encoded in the original texts?

Obvious candidates for the English *we* include Japan or the Japanese people as well as the Japanese Government and Cabinet members, and the majority of English plural pronouns indeed seem to represent these parties. The Japanese originals, again, often have recourse to ellipsis, but later references in the texts for example to the ‘government’ (*seifu* 政府), ‘Japan’ (*Nippon* 日本) or other similar overt referents serve to elucidate the identity of the missing actors (or experiencers). Other occurrences of *we* depict contexts in which a difference has to be made between ‘I’ as the writer/Prime Minister and ‘we’ for example as Cabinet members. When describing the signing of treaties, on the other hand, an inclusive *we* can be employed to include both the writer/Prime Minister himself and his signing partner, usually another head of state, or Japan and another – friendly – nation. In its most generic function, *we* appears to refer not only to the Japanese people, but perhaps widely to the people on this planet, leaving again the final interpretation of whom or what to include to the reader him/herself.

The contrast between Japanese elliptical nothingness and English *we*-agents is striking and must sometimes turn into a cause of headache for translators of these, at times, sensitive political texts. As some ellipses leave room for varying interpretations, not always clearly identifiable from anaphoric (or exophoric) contexts, trying to find the best possible – or politically correct – solutions for the English versions can be a daunting task. In this sense, due to its potential impersonality and identity-effacing nature, the plural *we* serves as a handy device to fill in the necessary grammatical slots in the English texts without compromising too much the actor-dissolving “mind-style” embedded in the Japanese originals.

It would perhaps be appealing to claim that, with obligatory actor *we*’s, the English versions seem to reflect a more involvement-enhancing

character than their Japanese counterparts. It has frequently been suggested that personal referents can be examined as markers of involvement (Chafe 1982), that is, of “emotional connection and engagement that the speakers feel in the situation towards the subject matter and/or the other participants in interaction” (Nikula 1996). Thus, in our data, singular 1<sup>st</sup> person referents could be interpreted to mark the speaker’s involvement with the subject matter (or himself). However, whether such “iconic” interpretations do in fact work well for pro-drop languages such as Japanese, would, in my view, require more rigorous research. As Yamamoto (1999: 101) puts it, in Japanese “explicit entities tend to be expressed implicitly and implicit entities tend to stay implicit”.

### 3 Discussion: Agency, “subjective-frames” and subjectivity

It has been argued that the notion of animacy hierarchy – and related agency – bears relevance to the concept of perspective (Iwasaki 1993: 79) or “mind-styles” in differing languages. This pilot study has not intended to explore the absence or existence of a writer (or speaker) systematically in the examined languages, Japanese and English, but rather to offer a glimpse at the varying degrees in which the 1<sup>st</sup> person may be used strategically to represent not only the writer himself, but his entourage, (occasional) liaisons, government or an entire nation in these languages.

Japanese is often referred to as a subjective-frame language (vs. objective-frame language such as English) and Japanese speakers and writers as subjectively oriented, having a tendency to encode situations as if being involved in them as experiencers, rather than describing them as detached objective on-lookers. A classic written example is the opening of Kawabata Yasunari’s novel *Yukiguni* (*Snow Country*). The original version has no mention of a 3<sup>rd</sup> person referent, while the English translation starts off with a definite noun, *the train*:

#### Example (3)

*Kunizakai no nagai tonneru o nukeru to yukiguni deatta. Yoru no soko ga shiroku natta. Shingoojo ni kisha ga tomatta.*

*The train came out of the long tunnel into the snow country. The earth lay white under the night sky. The train pulled up at a signal stop.*

How or if this tendency is to be linked with cognitive constructs of speakers and writers of a particular language – and therefore for example with intent, strategies or even social agency – remains under discussion. Additional studies would be needed to develop this topic and invite linguists

working on “subjective-frame languages” to further explore meanings and ‘fashions’ of speaking (allusion to Whorfianism intended).

Japanese (socio)linguists are sometimes keen on comparing Japanese with English (for them the representative “Western” language) and link typological characteristics of Japanese with traditional “ways of thinking”, for example interpreting or explaining the ellipsis/non-appearance of subjects as a manifestation of cultural traits like “modesty” (*kenkyo*) and “consideration (for others)” (*sasshi*) (Matsuo & Watanabe 2000: 103). In these formulations structural aspects of ‘personification’ receive iconic pragmatic interpretations based on the theory of negative politeness (Laitinen 1995: 341): no reference equals avoidance of subjectivity (to be differentiated here from “subjective-frames” à la Ikegami). This pilot study has attempted to show that, as Ono-Premper (2000: 63) ingeniously puts it, the role of ‘person’ (*ninshoo*) in Japanese is like that of “*kuroko*”, stagehands dressed in black, assisting *kabuki* theatre actors in various ways during performances. Although avoidance of explicit reference to animate human agents is deeply ingrained in the structure of Japanese, it is a matter of degree, with a reverse side which can be associated with textual strategies discussed in this presentation. I believe that future analyses of natural data will show that ‘person’ (*ninshoo*) and how it is communicated in Japanese interaction is an area in need of more rigorous research.

#### References

- Chafe, W. 1982. Integration and Involvement in Speaking, Writing, and Oral Literature. In: Tannen, D. *Spoken and Written Language: Exploring Orality and Literacy*. Norwood: Ablex, 35-53.
- Fowler, R. 1985. Power. In: van Dijk, T.A. (ed.), *Handbook of Discourse Analysis*. London: Academic Press.
- Ikegami, Y. 2008. *Linguistics and Poetics of ‘Ego as Zero’*. Paper presented at the European Association for Japanese Studies conference, 20.09.2008, Lecce.
- Iwasaki, S. 1993. *Subjectivity in Grammar and Discourse*. Amsterdam: John Benjamins.
- Laitinen, Lea. 1995. Nollapersoon. *Virittäjä*, 3: 337-358.
- Matsuo, K. & K. Watanabe. 2000. Nihongo ni okeru nihon bunka. *Yooroppa nihongo kyooiku*, 5: 103-109. Espoo: AJE.
- Mühlhäusler, P. & Harré, R. 1990. *Pronouns and People*. Oxford: Blackwell.
- Nikula, T. 1996. *Pragmatic Force Modifiers: A Study in Interlanguage Pragmatics*. *Studia Philologica Jyväskyläensia*, 39. Jyväskylä: University of Jyväskylä.
- Ono-Premper, Y. 2000. Daikushisu yooso no shidoohoo ni tsuite 2: Ninshoo no kakureta hyoogen. *Yooroppa nihongo kyooiku*, 5: 63-74. Espoo: AJE.
- Shimojo, M. 2005. *Argument Encoding in Japanese Conversation*. New York: Palgrave Macmillan.
- Yamaguchi, T. 2007. *Japanese Language in Use*. London: Continuum.
- Yamamoto, M. 1999. *Animacy and Reference*. Amsterdam: John Benjamins.

# Movie Reviews in Japanese and English Newspapers Published in Japan: Rhetorical Preferences in Each Discourse Community

Tama Kumamoto

Nagoya University of Foreign Studies

kumamoto@nufs.ac.jp

## Abstract

This article aims to disclose the different purposes the movie review genre has in the Japanese and English media. Ten movie reviews are analyzed, five in Japanese and five in English that appeared in newspapers published in Japan. The study has three phases: First, the comparison of the physical appearance of these articles is made, together with the examination of the placement of the thesis statement. Second, they are read and evaluated by four native speaker teachers of English, who are considered to be sharing the writing norms of the English discourse community, on the following three points: how much the review is informative, critical, and persuasive. Their free comments are also gathered. The final phase is to present the testimony of a Japanese movie critic on her experience of writing articles in newspapers, both in Japanese and in English, in order to reveal the norms of the Japanese discourse community in this genre. The results show that the focus of the English reviews is to critically evaluate the movies, while that of the Japanese ones is mainly to give information about the movie.

## Keywords

Intercultural rhetoric, genre analysis, movie reviews, language use in the media, teaching English writing

## Introduction

Ever since the seminal work set forth by Kaplan (1966), contrastive rhetoric has played a significant role in teaching English to speakers of other languages. However, there has been strong criticism against it: that it captures culture statically (Atkinson, 2004); that it stereotypes cultural conventions of writing (Kubota, 1997; Kubota and Lehner, 2004); and that it lacks consideration of the social and educational backgrounds of students (Connor, 1996). It has also been pointed out that errors students make can be accounted for by developmental problems, and not by the negative transfer from their first language (Mohan & Lo, 1985). These criticisms led Connor (2004a, 2004b)

to revamp this disciplinary area as intercultural rhetoric, by reexamining research methodology, and by redefining culture. She proposes that IR (intercultural rhetoric) research should go beyond text-level analysis, reflect the writing context, and include not only student essays but writing in many disciplines and genres.

There have been three schools of genre analysis according to Johns (2003), i.e., Sydney school, ESL, and the New Rhetoric in the USA. In the field of ESL, much work has been done in professional writing: business request letters, sales letters, grant proposals, application letters, letters of recommendation (Bhatia, 1993, for example), and in academic writing (Swales, 1990).

However, these works are all text-based, detached from wider social contexts. The new concept of genre does not limit its scope to business or academic written form, but expands to everyday social interactions, both in written and non-written forms, as suggested by Bakhtin (1986). Incorporating this new concept, it becomes necessary for a genre analyst to find out “the overall purpose or function for the creation of a text. The purpose will determine the internal generic structure of the text as well as its key grammatical features. (Nunan, 2007, p. 136)”

Movie review as a genre, when looked at with this new concept of genre in mind, may have different purposes (or functions) in different cultures. An English definition of “movie review” as “a report or essay giving a critical estimate of a work or performance” (The American Heritage Dictionary, fourth edition) may not always be true in other cultures. Beside the purpose of criticizing, movie reviews must have an information giving purpose, and the emphasis may shift between these two elements depending on the speech community.

The ability that a writer must have may vary in different communities. For the information giving purpose, the writer must have a descriptive writing ability to introduce the production of the movie, the plot, the actors, their acting, etc.; and for fulfilling the critical function, the writer must excel in an

expository writing ability. In order to critically review the movie, the writer has to present a judgment (a thesis) and give reasons (arguments) to support his/her judgment, which is exactly the definition of expository writing given by Martin (1985, p. 14).

In this paper, the purpose (or function) of movie reviews in two speech communities is examined by comparing reviews written in the Japanese and English newspapers. The comparison is made first in their physical properties (the generic and key grammatical structures), and then by eliciting comments by language experts both in the English and Japanese language communities to shed more light on the social context of this genre. In the final section of this paper, suggestions are made to improve teaching English writing in Japan.

## 1 Methodology

### 1.1 Procedure (Materials and participants)

Two sets of movie reviews on the same five movies, the total of ten reviews, are selected. One set is taken from *The Asahi Shimbun*, a Japanese newspaper, and another set from *The Daily Yomiuri*, an English newspaper published in Japan during the years 2006 and 2007. The five movies are *Bushi no Ichibun*, *Bobby*, *Happy Feet*, *The King of Scotland*, and *Flags of our Fathers*, which will be hereafter referred to as *Bushi*, *Bobby*, *Feet*, *King*, and *Flags* respectively. These movies were open to the public from November 2006 through March 2007 in Japan. (See Appendix.) The selection of the movies is made with care so that various types of movies are included, such as an animation (*Feet*), and a Japanese period drama (*Bushi*). The five writers in each newspaper are professional reviewers who were writing in each newspaper regularly at the time of selection<sup>1</sup>.

The five Japanese articles from *The Asahi Shimbun* are translated into English by this author, sentence by sentence without changing the order of them.

In the first phase of the study, the comparison of the physical appearance of these articles in each newspaper is made, together with the examination of the placement of the thesis statement, as has been traditionally done in contrastive rhetoric studies.

In the second phase of the study, the ten articles are read and evaluated by four native speaker teachers of English, who are considered to be sharing the writing norms of the English discourse

community. Two of them are American, one British, and one Australian. Their teaching experience in Japan varies from 5 to 20 years, while their experience in teaching writing varies from 1.5 to 20 years. They are asked to read the translated version of the Japanese articles first, and then a few months later, the English articles. Each article is evaluated on a six-point Likert scale for the following three questions: (1) informative—if they think the review gives a full account of the movie, (2) critical—if they think the review is making careful and fair judgments about the movie with well-reasoned support, and (3) persuasive—if they think the writing style of this review causes the reader to agree with the author. They are also encouraged to write comments freely on the texts and also on the questionnaire. These comments, together with the points given to each question, are analyzed to see the values attached to this genre in the English speaking discourse community.

The third phase of this study presents the testimony of a Japanese movie critic on her experience of writing articles in newspapers. She is a professional translator, currently writing a movie review regularly on a national weekly newspaper, and occasionally writes reviews and articles on movies in English as well. The exchange was done through email on a person to person basis.

### 1.2 Analyses

The data is analyzed along the three phases. First, the comparison in the physical appearance between English and Japanese articles is made, such as the length of the article, the sentence construction in the title, etc. Also, the location of the thesis (overall judgment of the movie) and that of the conclusion are investigated.

Second, the four teachers' evaluations on the Japanese articles and the English ones are compared to locate differences, if any, in regard to the informative, critical, and persuasive quality of the writing. The teachers' comments are also used to illustrate the norms of the English speech community.

Third, the Japanese movie critic's view on this genre is presented to explicate the function of this genre in the Japanese speech community.

## 2 Results and discussion

### 2.1 Physical appearance

#### 2.1.1 The length of the articles

The Japanese articles are generally shorter than the English counterparts, the average number of words in the translated English version in the former being 432, as opposed to 722 in the latter. It is customary

<sup>1</sup> Tadao Sato, for example, who reviewed *Bushi* in Japanese, is active also in the field of education, and published a book, *Ronbun wo do kaku ka (How to Write a Thesis)* from Kodansha in 1980.

that the same space is allotted to the movie review column in Japanese newspapers, whereas the whole page is allotted to movie reviews in English newspapers, resulting in each article having a different length.

### 2.1.2 The titles (headlines)

The majority of the English titles are actually the overall judgment about the movie, given in a full sentence. In fact, its function is like the thesis statement of an essay, with a topic and a controlling idea. The verbs are often transitive verbs, which indicate that the reviewers perceive that people (directors, actors, etc.) actually caused some effect on the movie in question. The Japanese titles are often static, with just a noun phrase to indicate the content of the movie. Compare:

Japanese titles:

Refreshing Laughter and Warm-heartedness  
The USA in '68 in Miniature  
Dizzy with First-rate Images  
Expressing Amin's Two-sidedness Superbly  
Describes the Cruelty of Battlefields and the Emptiness of War

English titles:

Yamada gets manipulative  
'Bobby' an odd mix of fact, fiction, good acting  
'Happy Feet' shakes its tail-feather  
Whitaker's 'Last King' rules the screen  
Army of actors makes up for uneven directing

### 2.1.3 The location of the thesis

Each article is outlined according to the topics of the paragraphs. Two of the Japanese reviews have the thesis (overall judgment of the movie) at the beginning; two at the end; and one at the beginning and at the end, but the second statement is contradictory to the first one (Table 1). Four out of the five English reviews have a thesis at the beginning, with two of them repeating it at the end (Table 2).

Table 1: The Location of the Thesis (Japanese)

	Beginning	Middle	End
<i>Bushi</i>	○		
<i>Bobby</i>			○
<i>Feet</i>	○		○ <sup>1</sup>
<i>King</i>	○		
<i>Flags</i>			○

Note 1: The judgment is contradictory to the first one.

Table 2: The Location of the Thesis (English)

Beginning	Middle	End
-----------	--------	-----

<i>Bushi</i>	○	○ <sup>2</sup>
<i>Bobby</i>		○
<i>Feet</i>	○	
<i>King</i>	○	
<i>Flags</i>	○	○ <sup>2</sup>

Note 2: The judgment is consistent with the first one.

## 2.2 Teachers' evaluations

### 2.2.1 Scores

The results of the questionnaire survey are summarized in Table 3. In spite of their length, Japanese reviews are rated higher in the informative category than English ones. English reviews are rated higher in the other two categories, i.e., critical and persuasive.

Table 3: The Average Scores of the Four Teachers' Evaluations

	Japanese		English	difference
informative	4.45	>	4.35	-0.1
critical	3.45	<	4.05	0.6
persuasive	3.90	<	4.20	0.4

### 2.2.2 Teachers' comments

Comments by native speaker teachers (T1 – T4) on the informative quality of Japanese reviews mention that they give out too much of the storyline. There are also positive comments on their offering some background knowledge that the general public might not know.

(T2) *Yes, pretty much the whole movie is described, even to the point that there are "spoilers"—he basically gives the ending away. (Bushi)*

(T1) *I learn a lot about the storyline, but happily not the way it is resolved in the end. The technical explanation about "motion capture" is also usefully informative. (Feet)*

On the contrary, comments on the informative quality of English reviews often mention that the giving out of the storyline is adequate. Sometimes there are comments that there is not enough information.

(T4) *The review appears to provide sufficient general outline of the events in the movie without providing details that would spoil the storyline for potential viewers. (Flags)*

(T3) *Not such a full account of the movie is given. The introduction comprises one-quarter of the review. The information is very sketchy: I feel that too much (information) is missing. How was Memphis distracted? What happened exactly?(Feet)*

Comments on the critical quality of Japanese reviews often point out that there is almost no

judgment, or that the reviews lack logical support:

(T1) *The judgments are contradictory, so I'm left wondering what the reviewer actually thinks.* (Feet)

(T2) *Well, it is an overview so there are almost no critical comments except one: "...screenplay loses its tension..." There is support and examples about the movie but it's more emotional than logical.* (Flags)

(T3) *Very little judgment: Could not fully understand what the critic thought of this film.* (King)

(T4) *I tend to agree with Sato's judgments but they are simply observations based on opinion and there is no concrete support or evidence provided to prove that those judgments are either careful or fair.* (Bushi)

On the persuasive quality of the Japanese reviews, native speaker teachers often expressed skepticism on the overly positive tones of the Japanese reviewers. T4 wonders if the reason he agrees with the reviewer is because of the shared knowledge that he has of the movie.

(T2) *Yes, with a review that is so positive, it might seem difficult to disagree with the author, however, I am dubious of overly positive reviews which lack any criticism.* (Feet)

(T4) *To some extent; each body paragraph basically consists of describing a section of the movie's storyline and an observation by Sato on its importance. The review appears to be well written but the fact that I am forced to agree with the author may largely be due to having already watched and enjoyed the movie.* (Bushi)

The reasons why the teachers think the English reviews are persuasive seem to lie in the details they read in the passages:

(T1) *I felt convinced that this is a movie I don't need to see, especially because of Par. 11 and 12.* (Bobby)

(T3) *It is very easy to follow the reviewer's line of reasoning in this article. "Whitaker dominates the screen", "riveting, and mesmerizing performance". Garrigan: "We share in his excitement that Amin is going to change Uganda for the better. Also persuades us that the setting, (color/chaos, "leap off the screen") is authentic/realistic ("brings the 70's back to life). It makes me want to watch this movie again!"* (King)

### 2.3 Testimony of a Japanese movie critic

There seems to be some difference in the emphasis that Japanese and English review writers place. The author asked a Japanese movie reviewer (referred to as JMR hereafter) on this point in our email exchanges, and received the following answer (the

author's translation):

(JMR) *To aim for the wide readership of the Japanese newspaper<sup>2</sup>, most reviews just touch upon the story (or content) of the movies without discussing technical issues. In other words, they are "introduction" of movies, rather than "criticism" of them. Actually, I know a noted critic who uses up more than half of the space for writing the outline of the story, which I don't think we can call "criticism." However, it is true that the number of good critical articles is decreasing, not only in newspapers but also in magazines, because of a recent trend in which readers don't want to know about anything but the genre and the storyline of the movie. In my purely "personal view", reviews in Japanese newspapers are treated as passing (light?) entertainment, not as culture. Of course, there are excellent reviews from the point of movie culture, I must add.*

She also has an experience of writing movie reviews in English for *The Asahi Evening News* (presently called *International Herald Tribune / The Asahi Shimbun*) several years ago, and recalls what her editors advised her, which obviously was not part of her training as a movie reviewer in Japanese. (JMR) *I received a lot of advice. The importance of the lead, the writer's personality, humor, the inclusion of footage and lines that show the atmosphere of the movie, etc. In sum, they wanted good readings, but each editor wanted different things....Since movies are composite art, I think reviewers should incorporate not only the storyline, but also various other things, such as the director's background, style, and technique.*

Her testimony reveals that the movie review as a genre has different functions in the Japanese and English speech communities. The expectations of the communities are different; therefore, reviewers write in different manners. Fowler (1991) observes a similar system working in writing news: "News is not a natural phenomenon emerging straight from 'reality', but a *product*....From a broader perspective, it reflects, and in return shapes, the prevailing values of a society in a particular historical context. (p. 223)" Likewise, Bakhtin states as follows:

Both the composition and, particularly, the style of the utterance depend on those to whom the utterance is addressed, how the speaker (or writer) senses and imagines his addressees, and

<sup>2</sup>*Yomiuri Shimbun*, for example, has more than one million subscribers as of January 2009.

the force of their effect on the utterance. Each speech genre in each area of speech communication has its own typical conception of the addressee, and this defines it as a genre (p. 95).

### 3 Conclusion and implications

Several differences are found in the Japanese and English movie reviews:

- (1) English titles (headlines) often are sentences with transitive verbs, which express judgments (thesis) of the reviewer, while Japanese titles are often noun phrases that show the content of the movies without much judgment.
- (2) Four out of five English reviews have the thesis at the beginning (deductive mode), with two of them repeating it at the end, while two Japanese reviews have the thesis at the beginning and two at the end (inductive mode). The last one has two contradictory judgments at the beginning and at the end.
- (3) Native English teachers feel Japanese reviews are informative, sometimes giving out too much, while they feel English reviews are more critical because of factual and logical reasoning, and persuasive because they offer both positive and negative views on the movie.
- (4) A Japanese movie reviewer's testimony confirms that they are in fact writing to cater to the palate of Japanese readers.

The findings of this study suggest the following points to those who teach writing movie reviews in English in Japan.

- (1) Awareness raising  
Students need to be aware of the different expectations that the genre has in the English speaking community. Show them movie reviews written in Japanese and English on the same movie and let them analyze those reviews. The cultural values behind the reviews should also be discussed.
- (2) Writing an outline  
The outline should be kept to a minimum. Students should also be instructed to omit the ending.
- (3) To overcome weak criticism  
It is important to encourage students to form their opinions about the movie, then let them find as many reasons as possible to support their opinions.
- (4) Finally, writing a good title (headline)  
The title should encompass the whole content and express the student's *judgment* about the movie.

### References

- Bakhtin, M. M. (1986). The problem of Speech Genres. In *Speech Genres and Other Late Essays* (pp. 60-102). Austin, TX: U of Texas Press.
- Bhatia, V. K. (1993). *Analyzing Genre: Language Use in Professional Settings*. London: Longman.
- Connor, U. (2004a). "From contrastive to intercultural rhetoric." Handout for Distinguished Lecture, Graduate TESOL Program, Temple University Japan, November 2004.
- Connor, U. (2004b). Intercultural rhetoric research: Beyond texts. *Journal of English for Academic Purposes*, 3, 291-304.
- Fowler, R. (1991). *Language in the News*. London: Routledge.
- Johns, A. M. (2003). Genre and ESL/EFL composition instruction. In Kroll, B. (ed.), *Exploring the Dynamics of Second Language Writing*. Cambridge: Cambridge UP, 195-217.
- Kubota, R. (1997). A reevaluation of the uniqueness of Japanese written discourse: Implications for contrastive rhetoric. *Written Communication*, 14 (4), 460-480.
- Kubota, R. and Lehner, A. (2004). Toward critical contrastive rhetoric. *Journal of Second Language Writing*, 13, 7-27.
- Martin, J. R. (1985). *Factual Writing: Exploring and Challenging Social Reality*. Oxford: Oxford UP.
- Mohan, B. and Lo, W. A. (1985). Academic writing and Chinese students: Transfer and developmental factors. *TESOL Quarterly*, 19, 515-534.
- Nunan, D. (2007). *What is This Thing Called Language?* New York: Palgrave MacMillan.
- Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge, UK: Cambridge University Press.

### Appendix. Movie Reviews Used in This Study.

#### *Bushi no Ichibun*

- Sato, T. "Bushi no ichibun--Warai to jo no sawayakasa (The social position of samurai: Refreshing laughter and warm-heartedness)," *Asahi Shimbun*, evening edition, November 30, 2006.
- Gerow, A. "Yamada gets manipulative," *The Daily Yomiuri*, December 2, 2006.

#### *Bobby*

- Inagaki, T. "Bobi--68 nen, Amerika no shukuzu (The USA in '68 in Miniature)," *Asahi Shimbun*,

evening edition, March 1, 2007.

Baker, T. "'Bobby' an odd mix of fact, fiction, good acting," *The Daily Yomiuri*, February 17, 2007.

*Happy Feet*

Akiyama, N. "Happi fito—Gokujo no eizo ni atama ga kurakura ('Happy Feet'--dizzy with first-rate images)," *Asahi Shimbun*, evening edition, March 22, 2007.

Bergen, A. "'Happy Feet' shakes its tail-feather" *The Daily Yomiuri*, March 24, 2007.

*The Last King of Scotland*

Shinada, Y. "Rasuto kingu obu sukottorando—Amin no hyori migoto ni hyogen ('The Last King of Scotland'--expressing Amin's two-sidedness superbly)," *Asahi Shimbun*, evening edition, March 15, 2007.

Satterthwaite, J. "Whitaker's 'Last King' rules the screen," *The Daily Yomiuri*, March 10, 2007.

*Flags of Our Fathers*

Sawaki, K. "'Chichioyatachi no seijoki'—Senjo no muzansa to senso no munashisa egaku ('Flags of Our Fathers'--describes the cruelty of battlefields and the emptiness of war)," *Asahi Shimbun*, evening edition, November 6, 2006.

Chapman, B. "Army of actors makes up for uneven directing," *The Daily Yomiuri*, October 21, 2006.

## Teaching Business Japanese through Internet: Importance of Teacher Training

Remi Kakiyama<sup>1</sup> and Yoko Okita<sup>2</sup>

<sup>1</sup>Faculty of Liberal Arts, Sophia University, <sup>2</sup>International Student Center,

Tokyo Medical and Dental University

remithk@gmail.com, yokita.isc@ymd.ac.jp

### Abstract

This paper reports about an experimental project of teaching business Japanese in a distance language learning context using the “LiveOn” ®(Japan Media System) web-conference system. The instructional purpose of this project is to provide opportunities to speak business Japanese through the internet, for learners of Japanese as a second language, who have little opportunity to practice Japanese in their daily life. The experiment was conducted from November 2008 to January 2009. Participants of the classes were two coordinators acting as class observers, five Japanese tutors, two teaching assistants and 14 graduate students in Graduate School of Information, Production, and Systems, Waseda University in Kitakyushu. The students’ Japanese proficiency levels varied from novice-high to intermediate-high level. This project explored main aspects of the distant course such as the use of LiveOn, teaching materials and future expectation of on-line learners, focusing on the participants’ feedback. Although students provided positive comments on the course, our observation of classes revealed that tutors are a key factor in the successful implementation of distance course through the Internet. Findings indicated that the importance of tutor training which can be set up in an appropriate timing so as to facilitate the classes taught entirely at a distance.

### Keywords

e-learning, distance language learning, teacher training

### Introduction

Crossing borders is a feature of distance learning courses and in recent years various distance educational courses are provided by distance learning in a wide range of fields such as school, company, local community and government. In Japan, in the late 1990’s, as ICT (Information and Communication Technology) advanced, e-learning

as a representative of distance learning through the Internet has slowly been established mainly in universities (see, for example, ALIC 2002; e-Learning Consortium 2008) making the most of its characteristics of combinations of time and place in learning context (White, 2003; Jung and Kubota, 2006). In fact, this experimental project was promoted by Waseda digital campus consortium with the aim of providing a good educational service tool towards Japanese language learning.

With regards to the effectiveness of the distance course, it is considered that a number of factors might influence on course management in relation to choice of media (e.g. TV conference, video conferencing system, PC) and software, audio equipment, course contents, teacher training, etc., which can affect quality in the learning context. In this project, however, researchers assumed tutors’ role is a key to success to students’ learning with interactive communication ongoing classes.

With respect to activities during the course, feedback is regarded as one of the important resources for the development of the course. And also, the LiveOn which is called synchronous system, permits immediate feedback that can be triggered by tutors’ views and experience, as well as providing opportunities for the development of learning environment within the learning group (White, 2003). Thus, it is worthwhile that the instructional issues concerned would be investigated improving quality of the distance learning and teaching, focusing on both positive and negative comments by the participants.

## 1 Method

### 1.1 Participants

Participants of classes were five native tutors of Japanese, two coordinators as class observers on Waseda University Campus, Tokyo, two teaching assistants and 14 graduate students, all of whom were involved in research at Graduate School of Information, Production, and Systems, Waseda

University in Kitakyushu. The students' Japanese proficiency levels were from novice-high to intermediate-high level (61.9%: elementary, 20.3%: intermediate; in total 320 students), depending on their background of Japanese language learning before and/or after their exchange students' life in Japan. The participants can access the course which is a component of an on-line Business Japanese Language program and Learning Management System (LMS) of courseN@vi.

## 1.2 Procedure

The instructional purpose of this project is to provide speaking opportunities of business Japanese through the internet, for learners of Japanese as a second language, who are looking for job opportunities in Japan after graduation (67.1%: 47 out of 71 students). However, an obvious problem is that the students have little opportunity to practice Japanese in their daily life either due to their research or housing environment, where English and/or their mother tongue are mainly used. Course objectives are a) to develop basic communication skills for use in business contexts, b) to learn how to participate in both casual and formal situation with colleagues and/or clients in a workplace, and c) to acquire knowledge and skills for fulfilling personal needs and aspirations through exercises.

The focus of this project was a six-week on-line course conducted at a distance from November 2008 to January 2009. Four classes were opened and each class consists of three to four students and one tutor. Each class met 18 times (three times per week and 45 minutes per class) for six weeks. The course was delivered to students using the "LiveOn" web-conference system, which allows at least five people to communicate with each other audio-visually through Internet. Each student and a tutor wearing a headset with a microphone sitting in front of a PC equipped with a web camera. After logging into a room of LiveOn, all participants' faces are displayed individually in a PC monitor with small windows. The tutor is able to observe four students' participation individually and all members in the room can share the same practical conversation. The alternative format of face-to-face classroom was set through interactive communication in the use of sound and visual of Internet.

In this project, lesson plans were designed by coordinators and a textbook for intermediate/advanced *Japanese: Practical Business Japanese* by TOP Language was adopted as a sub-material. Classes were conducted in Japanese by native tutors and mainly focused on model conversation and role plays consisting of

useful vocabularies and expressions that can be applied for students' job seeking activities.

This experimental project explored main aspects of the distant course such as the use of LiveOn, teaching materials and future expectation of on-line learners, focusing on feedback given by the participants. The regular observation of classes explored aspects of the course advantages and disadvantages for its implementation. The classes were monitored by the coordinators in order to facilitate the course assisting class activities and interactions between tutors and students, and tutors' teaching with tracing the progress of their technology use as well. Nevertheless, as Blake (2007) emphasized, how to obtain the needs of an institution's students is greatly influenced by quality that pursuing effectiveness of the course.

Through the course, it was considered that obtaining timely and informative feedback data is also very important so as to assess the process of the course as well. Thus, reports of lessons taught by the tutors were submitted through the mailing list after each lesson. At the end of course, feedback from questionnaire survey regarding the whole course was collected by tutors and students. Furthermore, reacting and evaluating opportunities were given to students, teaching assistants and tutors during feedback meeting, which was video recorded. The data collected was qualitatively reviewed and considered to shed light on possible strengths and limitations of the distance course in terms of students' satisfaction and instructional issues.

## 2 Results

### 2.1 Student feedback

The feedback was collected to elicit the students' view on experience of Business Japanese distant course using "LiveOn". As students' comments showed in questionnaire and feedback meeting, both positive and negative comments present that the course provided a good opportunity for them in learning Business Japanese. Significant determinants of students' satisfaction presented in their positive comments under LiveOn, future distant course, after graduation and class content. Overall, the findings indicate that the students did not have any serious problems with the use of technology in the process of language learning, and the course fairly met their needs. Additionally, although the critical comments show that some students have individual needs in language learning, it also represents students' positive attitude toward their language learning. According to their statement, they are eager to participate in Japanese language course whenever there are any

opportunities. It can be said the distant course greatly supported students' language learning for their purpose.

## 2.2 Tutor reports of lessons

Tutors reports were focused on two main points. One is related to technology problems can be caused by the system. The other is about how well the tutor could manage each class following the lesson plans provided. Tutors reported that the technology can interfere with the progress of the class as they often had to confirm sound and screen condition, and it did cause an anxiety to some extent. Another example was shown in the use of whiteboard. Although most of tutors understand it is useful for students to confirm the correct Kanji reading and pronunciation, the speed of the whiteboard function was slower than tutors expected and it did not allow smooth use for teaching. In terms of instructional issues, tutor reported that many students did not do pre-study and some students' language level have not reached the level of role play.

## 2.3 Tutor feedback

To follow up the tutors' view on the distant course, feed back (both questionnaire and feedback meeting) was collected and the summary below shows main aspects based on categories. Overall, comments submitted by tutors provided evidence that the distant course promoted students' Business Japanese language learning, although system problems occurred and disturbed class management, particularly, at the beginning stage. It is important to enhance tutors' teaching experience and their use of technology. Moreover, demonstration is strongly suggested so as to understand well the real intention of lesson plans that should be shared through the course.

## 2.4 Coordinators' observation

Coordinators' observation was mainly focused on the following matters: a) how useful the course materials were, b) how well tutors taught using technology, and c) what students could learn and whether they were satisfied with the course, etc.

Although the tutors were experienced in teaching courses in a face-to-face classroom, they had difficulties in teaching the course at a distance, including use of the LiveOn, e-learning system.

Firstly, with respect to lesson plans, both students and tutors' feedback was resulted in positive comments such as "well designed" and "succinct and adequate". However the issue arose that some tutors could not follow the lesson content properly because sometimes they could not understand the

real intention of content designed by coordinators, or probably students' levels in a class varied. Secondly, an example was that an experienced tutor seemed to have difficulties in gauging students' understanding of the lessons and/or the tutor's questions. It can be interpreted that e-learning would hide students' reactions, for example, to say "No. I don't understand," and it can be caused by no direct interaction. On the other hand, it might indicate that managing the classes and manipulating computer monitor require skills that can take time to get accustomed oneself to the environment. In fact, the five tutors' experiences of teaching Japanese at a distance or relevant experience and tutor training were varied. Lastly, through observation, tutors' class report, students and tutors' feedback of the whole course, it is clear that students were satisfied with the course. It is also revealed that the internet use can facilitate language learning and focusing on conversation practice. However, problems were brought to the attention of the observers and/or tutors. Observers and tutors realized that the abilities to use technology in conducting classes are greatly required and pedagogical skills need to be trained as it is different from face to face classroom.

To sum up, it can be considered the development of the distant course would rely on design and evaluation in order to set up an effective language learning environment. However the findings through the observation strongly suggests that teacher training in using technology would be focal point to guide and assist students' class activities successfully in an art of language learning through the Internet.

## 3 Conclusions

This experimental project raised many issues and suggested further areas of study in the development of distant course through Internet technology.

Comments showed by the students and tutors in this project appeared to be positive for future distance course with regard to students' needs in their language learning environment. It also provided the evidence that the combination of factors such as use of system in good conditions, teaching styles and tutors immediacy can result in an effective virtual learning environment. As Blake (2007) stated that language tutors prioritize collaborative interactions providing their students with opportunities to be engaged in. Researchers believe that the keys to an effective language teaching at a distance heavily depend on individual tutor's abilities to integrate new technology into a use of computer based environment.

Further research might examine issues of training tutors within distance learning contexts and

to gain some insight into tutors' views of the group-based distance learning.

#### 4 References

- ALIC (Senshin Gakushu Kiban Kyogikai). (2002). *E-learning Hakusho 2002-2003 Nendoban* [e-learning whitepaper 2002-2003]. Ohmsha
- Blake, R. J. (2007). New Trends in Using Technology in the Language Curriculum. *Annual Review of Applied Linguistics*, 27, 76-97.
- Japan e-learning consosiam (eLC). (2008). *e-learning bijinesu chosa* [A business survey of e-learning]. Tokyo Denki Daigaku Shuppanyoku.
- La, I. and Terashima, K. (2008). e-learning Hakusho. Jun, I. and K. Kubota (Eds). Enkakukyoiku to e-learning [Long distance education and e-learning]. Kitaoji-shobo
- White, C. (2003). *Language Learning in Distance Education*. Cambridge: Cambridge University Press.

# Intercultural Communication through Distance Learning among Korean and Japanese University Students

Bok-Myung Chang

Department of English, Namseoul University

bmchang@nsu.ac.kr

## Abstract

The purpose of this study is first to prove that the intercultural communication is important in understanding NNS-NNS interaction in EFL context. Second, this study focuses the fact that the cultural communication among university students in EFL contexts is effective in English acquisition and understanding intercultural differences. So the cultural theme is very important in constructing the distance learning syllabus. Third purpose is to prove that it is urgent to construct a new environment for acquiring cultural intelligence.

This study consists on two parts: the first is to survey the new model of NWCCDLP of the second semester, 2008, which focuses on intercultural discussions organized on the preparation class and joint class. The second part is to survey the cultural themes of out-of class chatting in the first semester, 2009 in NWCCDLP. The data was collected from the students' reflection papers and interviews. Qualitative methodology is used in analyzing the data from the students who participated in the course.

## Keywords

Communicative Competence, Cultural Competence, Distance Learning, Intercultural Communication, On-line Chatting

## 1. Introduction

This study focuses on achieving the followings:

- 1) To recognize the importance of English as an International Language (EIL) and the concept of World English
- 2) To develop the best way to help English learners to acquire English as an International language in the countries of outer/ expanding circles
- 3) To prove that interactions between non-native speakers through international distance learning are very effective in intercultural communication and cultivating the cultural intelligence among NNS.

## 2. Literature Review

### 2.1 English as an International Language

#### 2.1.1 Kachru's model of world English

Kachru(1992) defined World Englishes as three concentric circles, the Inner Circle, the outer Circle, and the Expanding Circle. Each circle represents the types of spread, the patterns of acquisition, and the functional allocation of English in diverse cultural contexts. The Inner Circle contains the countries in which English is used as a national language: USA, UK, Canada, Australia, and New Zealand. The Outer circle contains the nations in which English is used as a second language; Bangladesh, Ghana, Kenya, Malaysia, Singapore, etc. The Expanding Circle includes the nations in which English is being used as a foreign language: China, Egypt, Indonesia, Japan, Korea, etc. Jenkins(2003) said that the English spoken in the Inner Circle is said to be 'norm-providing', that in the Outer Circle to be 'norm-developing' and that in the Expanding Circle to be 'norm-dependent'. In other words, English-language standards are determined by speakers of English as a national language.

#### 2.1.2 Modiano's model of world English

A much more recent attempt to take account of developments in the spread of World Englishes is that of Modiano (1999). He breaks completely with historical and geographical concerns and bases the first of his two models, 'The centripetal circles of international English', on what is mutually comprehensible to the majority of proficient speakers of English, be they native or non-native. The centre is made up of these who are proficient in international English. That is, these speakers function well in cross-cultural communication where English is the lingua franca. They are just as likely to be non-native as native speakers of English. The main criterion, other than proficiency itself, is that they have no strong regional accent or dialect. Modiano's next band consists of those who have proficiency in English as either a first or second language rather than as an international language.

## 2.2 Intercultural Communication

### 2.2.1 Emotional Intelligence (EQ)

- Salovey and Mayer(1990)

The subset of social Intelligence that involves the ability to monitor one's own and others' feelings and emotions, to discriminate among them and to use this information to guide one's thinking and actions.

Goleman's Framework for Emotional Competence (1995, 2000)

Self-awareness, Social awareness,  
Self-management, Relationship Management

### 2.2.2 Cultural Intelligence(CQ)

- Social and emotional intelligence + cultural awareness
- To develop students' awareness of culture-specific aspects of communication that may combine with more personal styles and responses
- Cross-Cultural Aspects of Communication
  - 1) Sharing Information
  - 2) Receiving Information
  - 3) Noticing similarities and differences

## 3. Research Methodology

The data was collected from the students' reflection papers and interviews. Qualitative methodology is used in analyzing the data from the students who participated in the course.

Subjects : The university students from Japan and Korea who participated in NWCCDLP new model during the second semester, 2008.

## 4. Course Description

### 4.1 New Model of the 2<sup>nd</sup> semester, 2008

#### 4.1.1 Class Model

- Preparation Class – 1 week  
(Reading + Research +BBS)
- Joint Class – 1week  
(Voice Chatting + Text Chatting + BBS)
- Reflection Paper

#### 4.1.2 Social and Global Issues

Topic 1 : Introduction

Facilitation Skills, Intercultural translation

Topic 2 : Happiness Factors

Individualism/Collectivism

Topic 3 : Family Roles

High/Low context communication styles

Topic 4 : Neighborhood

Relationship Development

Topic 5 : Climate Change : Agreement Styles

### 4.1.3 Learning points

The new model of NWCCDLP focuses on cultivating the learning points based on the intercultural communication.

Raising awareness of reasoning behind others' happiness factors.

Identifying individual and group related sources of happiness.

Analyzing society in terms of group membership.

Understanding the difference between individualism and collectivism.

Raising awareness of personal value orientation related to individualism and collectivism.

Avoiding over-generalization and appreciating the difference between personal, cultural and national values.

Considering the connection between personal and cultural values in terms of happiness.

### 4.2 Cultural themes of out-of-class chatting of the 1<sup>st</sup> semester, 2009

The students from Namseoul and Waseda Universities were matched as partners and had to meet in cyberspace at the appointed time and date through the "Live On" system. The students from Namseoul and Waseda Universities could exchange the voice chat and text chat, and had to record all the processes of their chatting and must write their reports based on their records and submit them as a part of their course work.

In On-line chatting activities, the students participated in the form of small groups consisting of 3-5 students from Japan and Korea. Namseoul University students have to submit their report about On-line chatting activities as a part of course work.

#### 4.2.1 Discover common heritage

1) Language and Society: popular proverbs and metaphors. This theme includes some discussions of the meaning of silence in Asian cultures and to what extent we can be talkative or out-spoken and

sincere in English.

- 2) Traditional arts and folk tales
- 3) Food culture: natto, tofu, miso, soy source, cooking methods, eating and drinking manners
- 4) Social mores: Dependency, vertical vs. horizontal society, male and female relationship, good wivew and wise mothers,
- 5) Common grammatical or pragmatic features in Asian Englishes

#### 4.2.2. Current Society

- 1) Campus life and students Life: majors, graduation theses, human relationships, extra-curricular activities, love and marriage, etc.
- 2) Current Issues: NEET, Unemployment rate, environment issues, Modern history and School Textbooks, Child Abuse, Gender Issues, Social security, etc.
- 3) Media studies: newspaper, TV commercials, cellular phones, Internet, TV Dramas, movies, etc.
- 4) Social systems: Presidential Election, Group consciousness, funerals, wedding ceremony, World peace, crime rate, etc.
- 5) Globalization Issues: Convenience stores, multi-national corporation, European Union, East Asian Union, etc.

### 5. Conclusions

This study focuses on achieving these purposes.

- 1) to recognize the importance of English as an International Language(EIL) and the concept of World English, 2) to develop the best way to help English learners to acquire English as a foreign language in the expanding circle countries, 3) to devise a model of international distance learning and adapt it to the English learning environment of Japan and Korea. 4) to prove the intercultural communications between non-native speakers in EFL context are very effective in English acquisition.

Until recent times, English was used as a communicative tool among native English speakers, but recently English was used as a tool among native English speakers (henceforth NS) and nonnative English speakers (henceforth NNS). Even more the communication between NNSs and NNSs is delivered through English, so it is urgent to set up a new model of English as an "International Language" or a "World English." It is urgent to construct a new environment for learning and practicing English as an International language, and one of the possible models is to construct the international distance learning, which can be used as an adequate environment for interaction among non-native speakers in Asian countries.

For this research study, the new model of NWCCDLP in the 2<sup>nd</sup>, semester has been adopted and the out-of-class chatting data was collected in the 1<sup>st</sup> semester, 2009. The results of data collection and analysis through the quantitative and qualitative research methodology proved NWCCDLP was very help in improving the communicative abilities of the students who participated in this project, so that this research study can improve that the interactions among non-native speakers in EFL contexts are very effective in English acquisition process.

The results for this new model for CCDL were analyzed through both qualitative method and quantitative method, and the results can be summarized into several categories like these:

- 1) Most students could develop their English proficiency through the NNS-NNS interactions.
- 2) The international distance learning model can be suggested as English acquisition model for outer / expanding- circle countries.
- 3) The international distance learning model is very effective for understanding the intercultural communication and cultivating the cultural Intelligence among NNS.

In conclusions, NWCCDLP can be proved as a good model for cultivating English proficiency of the students in EFL contexts. From the results of this research, the conclusions can be suggested as follows. The first thing is that most students could develop their English proficiency through the NNS-NNS interaction. So it is very helpful in English acquisition to communicate in English among non-native speakers, especially focusing on the intercultural communication. The second thing is that the international distance learning model can be suggested as English learning model in EFL context. The third thing is that one of the distance learning models, NWCCDLP was very impressed with the possibility of being able to get in touch with students and cultures in foreign countries.

Most of the students who practiced English through NWCCDLP wished to participate in this CCDL model again if possible, and they think this project can help themselves to activate the interests in cultural understanding, get their self-confidence in using English, and improve their English skills.

## References

- Alptekin, C. (1993). Target-language culture in EFL materials. *ELT Journal*, 47(2), 136-143.
- Chang, B. M. (2005). World English acquisition through CCDL project. *Journal of Pan-Pacific Association of Applied Linguistics*, 9(2), 241-261.
- Chang, B. M. (2008). New environment for acquiring English as an International Language. *Proceedings of the 13th Conference of Pan-Pacific Association of Applied Linguistics*, 129-132.
- Chapelle, C. A. (2001). *Computer applications in second language acquisition: Foundations for teaching, testing and research*. Calbridge: Cambridge University Press.
- Crystal, D. (1997). *English as a global language*. Cambridge: Cambridge University Press.
- Goleman, D. (1995). *Emotional Intelligence*. New York: Bantam Books.
- Goldman, D. (2006). *Social Intelligence*. New York: Bantam Books.
- Jenkins, J. (2003). *World Englishes*. New York: Routledge.
- Kachuru, B. (1992). *The other tongue: English across cultures* (2<sup>nd</sup> ed.). Urbana, IL: University of Illinois Press.
- Klopf, D.W. and McCroskey, J.C. (2007). *Intercultural communication encounters*. MA: Allyn & Bacon
- Nakano, M. (Ed.) (2007). *World Englishes and Miscommunications*. Tokyo: Waseda University International.
- Nakano, M. and CCDL Research Center (2006). *Cross-Cultural Distance Learning Teacher's Manual Vol.1*. Tokyo:Media Mix, Waseda University.
- Nunan, D. (1999). *Second language teaching and learning*. Boston: Heinle & Heinle Publishers.
- Warschauer, M. (1997). Computer-mediated collaborative learning. *Modern Language Journal*, 81(4), 470-481

## University students' recognition and use of strings in L2 writing

Victoria Muehleisen<sup>1</sup> and Sean Wray<sup>2</sup>

<sup>1</sup>School of International Liberal Studies, Waseda University, <sup>2</sup>Waseda University International

vicky@waseda.jp, swray@kurenai.waseda.jp

### Abstract

Paraphrasing is an essential writing skill tool in composing academic essays. Many studies of L2 writing, however, primarily research the issue of plagiarism, only touching on paraphrasing as a way of avoiding plagiarism. This study examines how L2 students actually render text from outside sources in their academic essays; specifically, how students deal with unique lexical strings they encounter in the writing process. We use a corpus of nearly 700 student essays (totaling more than 450,000 words) written for two specific assignments (one for a basic level writing class, and one for an advanced level class), together with concordancing software to find repeated four-word strings. We compared these to the strings explicitly taught in the writing course, as well as to the strings which occur in the articles which the students were assigned to use as sources for their essays. We find that while even the basic level students are fairly skilled at identifying unique language, the ability to appropriately integrate these strings into their own writing is something that only develops over time. We also find that the advanced students have a greater command of lexical strings, which could be a result of their longer exposure to the conventions of English writing.

### Keywords

Academic writing; bundles; clusters; concordance; corpus; lexical strings; paraphrasing

### Introduction

Corpus-based research has established that a sizeable portion of language is built upon predictable word sequences. These lexical clusters, bundles or strings, such as “in order to,” or “on the other hand,” are characterized by their lexical commonness (Sinclair, 1991), their high-frequency, three- and four-word sequencing (Biber, Conrad, & Cortes, 2004), and by their distribution according to discipline (Hyland, 2008). Along with these formulaic strings, idiomatic language, such as “over the top” also occur as fixed-meaning phrases, and

are equally essential in communication. Indeed, Ellis, Simpson-Vlach, and Maynard (2008) claim that the presence or absence of a sufficient amount of certain lexical strings help distinguish native-like discourse from non-native. Yet, for the L2 student, particularly those learning academic writing, an additional burden lies with multi-word sequences found in sources that are not formulaic, but rather word choices unique to the author.

While students can be taught strategies to deal with words and phrases considered “shared” by particular discourse communities, i.e. language that is not unique to the author and thus does not have to be reworded, paraphrasing words and phrases that are not shared remains the initial hurdle for most second-language learners. The first stage of our investigation into how students deal with lexical strings they encounter seeks to identify and describe the strategies they deploy in incorporating non-formulaic four-word strings into their academic essays.

For this study, we use two learners' corpora of essays gathered from basic and advanced level English-language students in the academic writing program at the School of International Studies, Waseda University. The basic-level essay corpus contains 306 essays (about 166,000 words) written as first drafts for an assignment which required students to make an argument on the ethics of keeping animals in zoos. The advanced-level essay corpus contains 387 essays (about 296,500 words) written as first drafts for an assignment on the dangers and benefits of genetically-engineered (GM) food. For both assignments, students were required to choose sources from a set of provided readings with which to support their arguments.

Using the concordance application *AntConc*, we compared the strings most commonly found in students' essays to the strings used in the readings and the descriptions of the assignments. *AntConc*'s N-gram tool allowed us identify strings which were used repeated in the corpus (that is, strings which were used in the essays of many different students), while the File View tool allowed us to zoom in on the use of a string in the context of a particular essay. In this paper, we present a

preliminary analysis of four-word clusters found in the two student essay corpora.

## **1 The most common four-word strings and their sources**

### **1.1 Strings in the basic level essays on zoos**

There were 433 four-word strings which occurred more than 10 times in the corpus. The top 25 of these are shown in Table 1 in the Appendix. Looking at just the top 25, we immediately see that the strings do not occur independently of each other. For example, the top two strings “zoos are cruel and” (175 occurrences) and “cruel and unnecessary” (154 occurrences) occur together 150 times in the longer string “zoos are cruel and unnecessary.” (Other ways of continuing the string “zoos are cruel and” include “zoos are cruel and provide no significant benefits”, and “zoos are cruel and costly”.) Similarly, of the 71 occurrences of “reason why zoos are”, 56 are part of the longer string “another reason why zoos are”. (Other longer strings include “the reason why zoos are” and “another common reason why zoos are”.) The total figure of 433 four-word strings thus includes many strings which overlap in the text.

Table 1 also shows the probable sources of the top 25 strings. In many cases, we can be sure through the use of quotation marks or attribution verbs and phrases that the writer intentionally took the string directly from a particular reading. In other cases, even when such markings are absent, our experience as writer teachers makes us think that students have learned the use of a word or phrase from the reading; one example of this is “play an important role”, with a sense of the word play not often encountered in the writing of students at this level. However, there are still other strings, such as “animals in zoos are”, which occur in at least one of the readings but which a student might also independently create, so the sources can only be tentatively identified.

Six of the top 25 strings seem to have been taken directly from the description of the task given to the students. Another eight of the top 25 include parts of strings which were explicitly taught in class, and which the students are using to structure their essays. (In all, 39 of the top 100 strings were of this type). Four of the top 25 strings are strings which occur in the same exact form in one or more of the readings, and which the students may have taken from these sources (Of the top 100 students, 11 were of this type). Four of the strings are what we are calling “topic-related”. They are not found word-for-word in the task description or the readings, but many students have come up with

them because they are closely related to the topic and therefore an efficient way to package information. Finally, the asterisks mark strings which seem to come from the students themselves: students are specifically urged not to use absolute quantifiers such as the “all” found “all zoos are cruel” and “that all zoos are,” and the strings “over the world” and “are a lot of” are both strings we have often seen in essays written for other assignments. (In fact, the former is also one of the top 25 strings in the essays of the advanced students, which are described in the next section.)

### **1.2 Strings in the advanced level essays on GM food**

The essays written by the advanced level students for this assignment were required to be longer than those written for the basic level assignment on zoos. This is probably one reason for the greater number of strings in the GM Food essay corpus: there were 792 four-word strings which occurred 10 or more times. The top 25 strings are listed in Table 2 in the Appendix.

Looking at sources of the top 25 strings, we see that unlike the basic students, the advanced students did not rely so heavily on strings that had been used in the task description or explicitly taught in class. Three of the top 25 strings are ones that students may have been taught explicitly in previous writing classes, namely “on the other hand,” “can be said that”, and “it can be said”, and altogether the top 100 strings include 13 of this type. However, the strings most often used by advanced students come directly from the readings (10 of the top 25 strings, and 35 of the top 100) or are closely related to the topic (9 of the top 25, and 46 of the top 100). The nature of the strings taken from the readings seems to be somewhat different as well, with several of these strings being from the titles of the readings or proper names from the texts, in other words, strings that cannot be paraphrased. The topic-related strings, however, are good candidates for stretches of text in which the students have tried to paraphrase the reading.

## **2 Closer examination of two strings taken from the readings**

In this section, we look more carefully at how students are using strings taken from the readings in quotations and in paraphrases. Due to space limitations, we will focus on just one interesting string from each of the corpora.

## 2.1 “deprived animals of their” in the zoo essays

This string occurred 39 times in the basic-level corpus. We'll first look at the 28 times it occurred as part of the longer string “zoos deprive animals of their freedom”, which occurred in the following sentence from one of the readings: “Since zoos deprive animals of their freedom, it is difficult for species such as elephants, bears and lions to be content in the limited condition of a zoo.” Only one student included the 6 word string in a direct quotation: “According to Naturewatch, ‘zoos deprive animals of their freedom.’” One other student used the whole sentence and attributed it to a particular reading, but without quotation marks: “In article2, it is said that since zoos deprive animals of their freedom, it is difficult for species such as elephants, bears and lions to be content in the limited condition of a zoos.” Seven students embedded the string in a sentence after a phrase such as “some people may say” or “some claim that”, while the others simply incorporated the string (in one case, the whole sentence from the reading) into their own essays without any attribution: “First, zoos deprive animals of their freedom.”

In the 11 cases in which the shorter four-word string was used by itself, we can see some interesting attempts at using the phrase creatively. Some of these are successful (“Although zoos deprive animals of their freedom to a certain degree...”) but some are less so (“This means that zoos limit animals’ growth and deprive animals of their lifetime.”)

## 2.2 “single case of harm” in the advanced-level GM food essays

Next, we examined the use of the string “single case of harm”, which occurred 30 times in the advanced-level corpus. This string also occurs in the following sentence from one of the readings: “Further, Americans have collectively eaten over a trillion servings of food containing one or more GM ingredients, without a single case of harm.” We found that fifteen of the students quoted the sentence in its entirety within quotation marks (although two slightly changed the wording!) with an attempt to indicate the source. Ten students did not use quotation marks, but correctly attributed the idea to the authors of the reading, either in the same sentence or in a paragraph summarizing the reading. Only five students used the string “single case of harm” with no attribution at all, one in the introductory paragraph and two in the conclusion.

Looking at paraphrasing, we found several different ways that the four word string was taken

up: “not a single case of harm”, “without one single case of harm”, “there was no single case of harm”, and “no single case of harm is reported”. These short paraphrases are fairly successful, but we see that a few students run into difficulty when trying to paraphrase the whole sentence: “...the overall US people have eaten over a trillion serving of food added to at least one genetically modified ingredients and so far not a single case of harm has been found.”

## 3 Discussion and conclusions

Although in its early stages, our initial findings do suggest that students’ source integration strategies of unique lexical strings share some over-arching characteristics, but also differ according to the stage of English development. While the advanced students relied much less on intact strings from the task description and taught structures, both basic and advanced level students tended to avoid paraphrasing source-based strings, but did use them in either a cited quotation, near copy or copy. For L2 students, this strategy of imitation is indicative of their inexperience. Keck (2006) found that when given the same summarizing task, most L1 students attempted to substantially reword the text while most L2 students did not. Similarly, in a study of Chinese students, Shi (2004) reported similar findings, but also emphasized that the students’ borrowing of large chunks of language was not evidence of intentional plagiarism, but rather of learning. This is likewise true of our students, since the avoidance of plagiarism and the correct use of sources was an explicit focus of instruction in these classes. Rather, we would argue, until students gain both a breadth and depth of lexical sophistication, they have little recourse but to blend their voice with another to create both coherent and intended meaning. Howard (2001) calls this process of learning to paraphrase patchwriting; Pecorari (2003) further adds that patchwriting is a process whereby students borrow without malice to supplement their output to achieve a goal, gradually reducing their dependency as they gain competency; our early results seem to support this notion.

The results of our corpus enquiry indicate that students at the basic level reproduced strings from the task description more often whereas the advanced students, perhaps judging the task description an easy or obvious structure to reword, attempted significantly more changes. This tendency is also in line with Day and Brown’s (1983) research on L2 summarizing ability; at the novice level, students rely on a basic copy-delete strategy in which a few words are removed or changed, but a large chunk of the text is copied

almost verbatim, while at the invention stage students are able to reproduce meaning faithful to the original text, but independent of the source language. Our findings also suggest a relationship with Li and Schmitt's (2009) research into how lexical strings are learned. Their study suggests that acquisition seems optimal when strings are encountered in context, such as in sources, and when lexical strings are explicitly taught, as when presented in textbooks, but less effective via instructor feedback. With the basic students in our study, the higher frequency of strings traced back to sources and the course textbook may indicate that these novice writers are simply using pre-structured strings relevant to the topic, but new to them.

Future analysis of the learners' corpus hopes to yield a more complete description of student strategies when confronted with both unique lexical strings and formulaic strings. Other areas inviting research are also whether an intuitive word limit exists which acts a line of demarcation between opting to paraphrase, or to quote or summarize. In addition, a study to determine if the grades awarded to L2 essays which include these "native-like" strings bear a relationship may reveal more about 'the other half' of writing instruction – the instructor.

## References

- Anthony, L. (2008). *AntConc 3.2.2w for Windows*. Waseda, Japan: Retrieved from <http://www.antlab.sci.waseda.ac.jp/>
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at... Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25L, 371–405.
- Brown, A., & Day, J. (1983). Macrorules for Summarizing Texts: The Development of Expertise. *Journal of Verbal learning and Verbal Behavior*. 22, 1-14.
- Ellis, N., Simpson-Vlach, R., & Maynard, C., (2008). Formulaic Language in Native and Second Language Speakers: Psycholinguistics, Corpus Linguistics, and TESOL. *TESOL Quarterly*. 42(3), 375-396.
- Howard, R. (2001). Plagiarism: What Should a Teacher Do? Paper presented at the Conference on College Composition and Communication, Denver, Colorado. Retrieved from <http://wrt-howard.syr.edu/Papers/CCCC2001.html>
- Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*. 27, 4-21.
- Keck, C. (2006). The use of paraphrase in summary writing: A comparison of L1 and L2 writers. *Journal of Second Language Writing*. 15, 261-278.
- Li, J. & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*. 18 (2), 85-102.
- Pecorari, D. (2003). Good and original: Plagiarism and patchwriting in academic second-language writing. *Journal of Second Language Writing*. 12, 317-345.
- Shi, L. (2004). Textual Borrowing in Second-language Writing. *Written Communication*. 21, 171-200.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

# Appendix. Tables of the most frequently occurring clusters.

Table 1. Top 25 four-word clusters in the basic level essays on zoos

Rank	Number of occurrences	4 word cluster	Probable source
1	175	zoos are cruel and	task description
2	154	are cruel and unnecessary	topic-related
3	133	zoos are necessary for	task description
4	106	all over the world	*
5	103	for animals and people	task description
6	102	zoos are not necessary	reading
7	92	necessary for animals and	task description
8	80	all zoos are cruel	**
9	72	<i>these reasons</i> zoos are	taught in class
10	71	<i>It is true that</i>	taught in class
11	71	<i>reason why</i> zoos are	taught in class
12	68	<i>Another reason why</i> zoos	taught in class
13	67	<i>First of all</i> zoos	taught in class
14	65	that zoos are cruel	task description
15	63	are necessary for animals	task description
16	63	Fact Sheet on Zoos	reading
17	62	an important role in	reading
18	60	play an important role	reading
19	59	animals in zoos are	topic-related
20	59	are a lot of	*
21	55	<i>say that</i> zoos are	taught in class
22	54	<i>argue that</i> zoos are	taught in class
23	52	<i>for these reasons</i> zoos	taught in class
24	50	for animals to live	topic-related
25	48	that all zoos are	**

Table 2. Top 25 four-word clusters in the advanced level essays on GM food

Rank	Number of occurrences	4 word cluster	Probable source
1	177	Union of Concerned Scientists	reading
2	123	that GM food is	topic
3	93	all over the world	*
4	90	On the other hand	*
5	89	of genetically modified food	reading
6	82	genetically modified food is	topic
7	78	that genetically modified food	topic
8	73	can be said that	*
9	73	genetically modified foods are	topic
10	72	Risks of Genetic Engineering	reading
11	68	Paul Driessen and Cyril	reading
12	67	of GM food is	topic
13	64	Driessen and Cyril Boynes	reading
14	64	of genetically modified foods	topic
15	64	that genetically modified foods	topic
16	63	it can be said	*
17	62	and Cyril Boynes Jr	reading
18	59	the use of pesticides	reading
19	58	it is possible to	reading
20	58	that GM foods are	topic
21	56	GM food can be	topic
22	54	humans and the environment	topic
23	50	the safety of GM	topic
24	49	GM Facts versus Fears	reading
25	48	at the same time	reading

# Links between research on L2 phonological variation and the pedagogy of L2 phonology: A study of phonological variation in the English interlanguage of Cantonese speakers

Yuet Hung Cecilia Chan

Department of Chinese, Translation and Linguistics, City University of Hong Kong  
ctcyhc@cityu.edu.hk

## Abstract

This study focused on the phonological variation of initials /n-/ and /l-/ in the English interlanguage of Cantonese speakers in Hong Kong. This study adopted Labov's (1972) and Tarone's (1983, 1988) research methodology. It included elementary, intermediate and post-intermediate English learners from primary school, secondary school and university respectively. Six English tasks were designed for collecting data: free conversation, informal interview, passage reading, word reading, minimal pair reading and minimal pair repetition. The results of the study show that (1) L1 negative transfer plays a role in L2 phonological variation; (2) L2 phonological fossilization occurs; (3) The frequency of phonological substitution varies with speech styles; (4) A low level of phonological awareness leads to phonological substitution. There are links between the research findings and the pedagogy of L2 phonology. The pedagogy of L2 phonology should aim at enhancing learners' phonological awareness in L2, enriching them with knowledge of L2 phonology and phonetics, training them to analyse the existing variation patterns in their L2 production, and using different speech styles for promoting pronunciation accuracy.

## Keywords

Phonological transfer, interlanguage phonology, phonological variation

## Introduction

This study investigated the English interlanguage of Hong Kong Cantonese speakers, with a focus on the production of syllable initials /n-/ and /l-/. Previous studies found that Hong Kong Cantonese speakers do not make a distinction between /n-/ and /l-/, and initial /n-/ is replaced by /l-/ in their speech (Yeung, 1980, Bourgerie, 1990, Bauer, 1982, 1997, and Ho 2004). Studies on Hong Kong students learning English indicate that due to L1 phonological transfer, students mispronounced English words with /n-/ as /l-/ (Bolton and Kwok, 1990, Chan, 2000, Hung, 2000, 2002, and Au, 2002). This study, unlike previous studies, not only investigated the role of L1 in L2 interlanguage of Hong Kong students, but also closely examined the interaction

of L1 phonological transfer and phonological awareness in L2 interlanguage phonology. On the basis of the results of this study, the pedagogical implications are discussed.

## 1 Research background

Recent studies have started paying attention to the merging of /n-/ with /l-/ in Hong Kong English. The findings of these studies will be briefly reviewed, followed by a short description of the issues of second language acquisition (SLA) related to our study, including L1 transfer in SLA and the variability of interlanguage.

### 1.1 Previous studies on the merging of /n-/ with /l-/ in Hong Kong English

The phonological feature of the merging of /n-/ with /l-/ in Hong Kong English received attention in a small number of studies (Bolton and Kwok, 1990, Chan, 2000, Hung, 2000, 2002, and Au, 2002). The phonological variation of /n-/ and /l-/ described in Bolton and Kwok (1990) and Chan (2000) was based on a small quantity of incidental observation data. This phonological variation was more closely investigated in Hung (2000, 2002) and Au (2002). These two studies exclusively used unnatural data collected from the reading of wordlists and controlled presentations. All the students in the above research were studying at university. Among these studies, Bolton and Kwok (1990), Chan (2000), and Hung (2000, 2002) pointed out that the two initials, /n-/ and /l-/, were used interchangeably by students due to the influence of their mother tongue, Cantonese, in which /n-/ and /l-/ are merged. In Au's study (2002), among 74 replacement patterns of single onset consonants shown in the data collected, the phenomenon of the loss of /n-/ in a number of lexical items demonstrates that the students had difficulty in producing the onset /n-/ in English. Au suggested that the replacement of /n-/ by /l-/ in English is influenced by non-distinction of /n-/ and /l-/ in Cantonese. All the above studies claimed that mother tongue interference plays a significant role in the variation of /n-/ and /l-/ in Hong Kong English. This argument would be more convincing if the language data were elicited from learners of different levels of L2 proficiency, and

from different stylistic contexts.

## 1.2 SLA issues related to our study

L1 transfer is always an influencing factor which receives serious attention in L2 studies. In most of the traditional studies, interference and developmental errors were collected from L2 learners of different native language backgrounds and proficiency levels, and the role of the L1 in L2 acquisition was examined (Ellis 1994). Following the traditional studies, this study investigated the pronunciation errors of the replacement of /n-/ by /l-/ in the English interlanguage of Hong Kong students.

As described in the SLA literature, variations appear in L2 interlanguage. This study focused on the systematic interlanguage variations associated with the variation of stylistic contexts. Previous studies suggested that the variations of contexts are associated with variations of learners' attention to language form or propositional content, and with the availability of planning time. Ochs (1979) and Tarone (1982, 1983, 1988) adapted Labov's (1972) methodological framework and captured L2 learners' interlanguage variability along with the shifting of styles. The systematic variability in interlanguage was analysed as on an unplanned/planned continuum (Ochs 1979), or on a vernacular/careful stylistic continuum (Tarone 1982, 1983, 1988). The planned and careful styles, which involve much attention to speech, manifest more correct and target-like language forms than do unplanned and vernacular styles. The effects of planning time on L2 production were examined and mixed results obtained. Hulstijn and Hulstijn (1984) found that learners' focusing attention on form, rather than planning time, increased language accuracy. Ellis' (1987) study suggested that the availability of planning time systematically influenced the accuracy of using particular linguistic forms. Ellis (1994) explained that if planning time was used to focus on informational content but not on language form, language accuracy was not likely to increase.

Within Labov's (1972) and Tarone's (1982, 1983, 1988) theoretical and methodological research framework, this study adopted a quantitative approach to analyse the merging of /n-/ with /l-/ in different contexts by Hong Kong students learning English. We further propose not only that the amount of attention to language form varies in different contexts as suggested by Tarone (1983, 1988), but also that the degree and level of language awareness differ in different stylistic contexts. In L2 acquisition studies, language awareness is commonly equated with language consciousness (Schmidt 1990). It has been

recognised that there are various degrees or levels of awareness. Schmidt (1990) summarises three crucial levels: (1) perception, including conscious and subliminal perception, (2) noticing (or focal awareness) and (3) understanding, which are from a low to high order. Most L2 studies were interested in examining the roles of conscious and unconscious processes in second language learning, and investigating the models and methods of consciousness-raising for facilitating language learning (Bialystok, 1978, Rutherford and Sharwood Smith, 1988, Schmidt 1990, 1995). In those studies, attention was paid to the awareness of syntactic properties of the target language by L2 learners. In our study, we focused on the phonological awareness, specifically the awareness of the distinction between /n-/ and /l-/ by Hong Kong students learning English.

## 2 Aims of the study

This study aimed at investigating the phonological substitution of /l-/ for /n-/ by Cantonese speakers learning English.

The following research questions were addressed:

- (1) Does L1 phonological transfer in respect of the replacement of /n-/ by /l-/ occur in the English production of Cantonese speakers learning English? If so, does the occurrence of phonological transfer negatively correlate with L2 proficiency?
- (2) Does L1 phonological transfer vary in different speech contexts which require different amounts of attention to pronunciation, and raise different degrees and levels of phonological awareness?
- (3) Is there any link between our research on L2 phonological variation and the pedagogy of L2 phonology?

## 3 Research methods

### 3.1 Participants

There were three groups of English learners: elementary, intermediate and post-intermediate. Twenty participants were in each group, with equal numbers of females and males. The elementary learners were fifth or sixth year primary school students aged 10-12, the intermediate learners, fourth or fifth year secondary school students aged 16-18, the post-intermediate learners, from year one to year three university students aged 20-23. The mother tongue of the students was Cantonese, a Chinese dialect. All the students learnt English as a second language since studying in primary schools from around 6 years old. They spoke English at school, but seldom did so outside school. They did not receive any formal systematic training in English phonics, and learnt word pronunciation from English

teachers who may not have been native English speakers.

### 3.2 Research instruments

#### 3.2.1 Oral tasks

Six tasks were designed for collecting participants' speech data: (1) conversation in pairs, (2) informal interview, (3) passage reading, (4) word reading, (5) reading of minimal pairs and (6) repetition of minimal pairs.

##### (1) Conversation in pairs:

Two participants who knew each other in the same student group formed a pair. Ten questions were given to them, each question including some words with a syllable-initial /n-/ or /l-/. However, the participants did not know that their pronunciation of /n-/ and /l-/ was assessed. They took turns to ask questions and give responses to their partner's answers. They were also allowed to talk about something else which they were interested in. Each pair was given 20 minutes for casual conversation. Examples of questions follow:

- (a) Do you have an English name? Would you like your friends to call you by your Chinese name or your English name?
- (b) If you had a chance to travel around the world, who would you choose as your partner on the journey?

##### (2) Informal interview:

Each participant was interviewed by the investigator or a trained research assistant in a casual way. Ten simple questions were asked by the investigator. The questions included some words with a syllable-initial /n-/ or /l-/. The participants did not receive any clues that we were focusing on the sounds /n-/ and /l-/. They were asked to answer the questions and talk freely with the investigator. Each interview lasted about 20 minutes. Examples of questions are given below:

- (a) What do you eat for lunch and for dinner?
- (b) Do you like to study in the morning, or at night? Why?

##### (3) Passage reading:

A passage in simple English of around 200 words was designed. About 40 words in the passage had syllable initial /n-/ and five words had syllable initial /l-/. Students read the passages at their own pace.

##### (4) Word reading:

A wordlist with 53 words was designed. In the wordlist, the initial /n-/ was in the 1st, 2nd, 3rd or 4th syllable of the word, and in some words, initial /n-/ was in more than one syllable. For example: number, nation (an initial /n-/ in the first syllable) morning, penny (an initial /n-/ in the second syllable)

National, Panasonic (initials /n-/ in more than one

syllable)

##### (5) Minimal pair reading:

There were 26 minimal pairs in the list. In each minimal pair, one word had an initial /n-/ and one an initial /l-/.

Examples of minimal pairs follow:

- (a) line                      nine
- (b) low                      no

##### (6) Minimal pair repetition:

The minimal pair list designed for the reading task was used again for the repetition task. Students listened to the pronunciation of each minimal pair recorded by a native English speaker, and then read each pair aloud during the pause.

#### 3.2.2 Structured interviews

A structured interview was conducted with each participant after they had completed all the oral tasks. The interview aimed to know: (1) whether the participants learnt to make a distinction between /n-/ and /l-/ in English lessons at school; (2) whether and when they are aware of making a distinction between /n-/ and /l-/ when they do the oral tasks.

#### 3.3 Administration procedure

All participants followed the same sequence to complete the six tasks in one day: (1) conversation in pairs, (2) informal interview, (3) passage reading, (4) word reading, (5) reading of minimal pairs, and (6) repetition of minimal pairs. Students were given breaks between tasks for resting and preparation. All tasks were administered in a soundproof recording room and recorded by a professional MD. Apart for the task of conversation in pairs, all the tasks were administered to students individually. After the completion of all the oral tasks, a structured interview was conducted with each participant.

#### 3.4 Methods of data transcription and analysis

The scripts of conversations and interviews were worked out and all the words which had initials /n-/ and /l-/ were identified. The pronunciation of words with /n-/ and /l-/ in each task by each participant was transcribed. All data were transcribed by three research assistants. The transcriptions were then randomly checked by the investigator and a phonetician. Spectrographic analyses were required when the acoustic properties of /n-/ and /l-/ produced by the participants were difficult to identify. In a few cases, the initials /n-/ and /l-/ were not clearly pronounced, and those data were not included in the analysis.

After the sound identification, the mean percentage of the production of /n-/ words in each task by each group was worked out. One-way

ANOVA and post-hoc Scheffé test was used to compare the results of four reading tasks (i.e. passage reading, word reading, minimal pair reading and minimal pair repetition tasks) of the three proficiency groups. The Proportion Test, which is a significance test of two proportions of two independent samples, was used to compare the results of two spontaneous conversational tasks (i.e. conversation and interview) between each pair of proficiency groups.

## 4 Results and discussion

### 4.1 The replacement of /n-/ by /l-/ by the three proficiency groups

Figure 1 shows that all the three groups substituted /l-/ for /n-/ and this result indicates that L1 phonological transfer takes place.

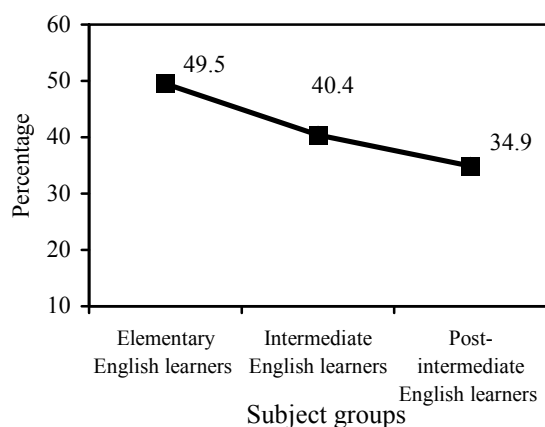


Figure 1: Mean Percentage of words with replacement of /n-/ by /l-/ by the three English proficiency groups

The results of one-way ANOVA show that there were significant differences between the results of the four reading tasks of the three proficiency groups ( $F(2, 58) = 3.72, p < 0.05$ ). The post-hoc Scheffé tests indicate that there was a significant difference between the elementary group and the intermediate group, and between the elementary and post-intermediate group, in respect of the replacement of /n-/ by /l-/ in the four reading tasks, while a significant difference was not found between the intermediate group and post-intermediate group. The same significance test results were obtained from the Proportion Test on the two casual conversational tasks between groups. All the significance test results show that L1 interference in terms of the replacement of /n-/ by /l-/ was less evident in the intermediate and post-intermediate groups than in the elementary group. But the post-intermediate learners did not produce more accurate pronunciation than intermediate learners, indicating that phonological fossilization occurred.

### 4.2 Replacement of /n-/ by /l-/ in different tasks

As shown in Figure 2, in the two casual conversational tasks (i.e. conversation and interview) which required students to attend more to informational content than to accurate pronunciation, a high percentage of /n-/ words (about 65%) was replaced by /l-/.

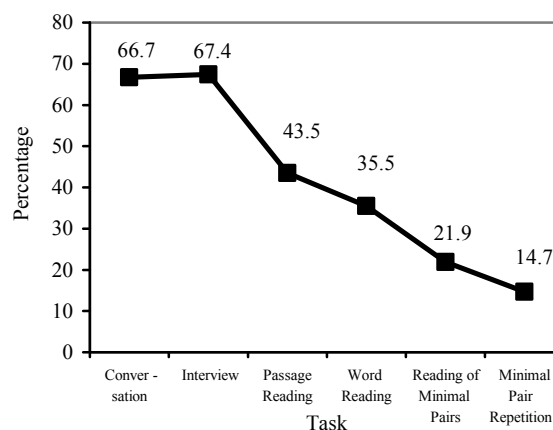


Figure 2: Mean percentage of words with replacement of /n-/ by /l-/ in each oral task

In contrast with the two casual conversational tasks, in the passage reading task, the inaccuracy rate of pronunciation of /n-/ words decreased sharply to 43%. This is due to the fact that the passage reading task of careful and formal style does not require participants' attention to the speech content, but only to pronunciation. In turn, the substitution of /n-/ by /l-/ was relatively less frequent in the word reading task than in the passage reading task. This is because when learners read the word list, they paid very close attention to the pronunciation of each word. The minimal pair reading task of much more careful and formal style raised learners' high degree and level of awareness of the distinction between /n-/ and /l-/ and therefore it was not so likely that /n-/ was substituted by /l-/ as shown in Figure 2, only about 20% of words with /n-/ were replaced by /l-/ in the minimal pair reading task. In the minimal pair repetition task of the most careful style, because learners had a chance to listen to the model pronunciation of each pair of words before they read them aloud, it was least likely for learners to mispronounce /n-/ as /l-/.

### 4.3 Learners' awareness of the distinction between /n-/ and /l-/

The results of the structured interview show that more than half of the participants had incidental or planned learning of making a distinction between /n-/ and /l-/ in English lessons. 18% of participants

in intermediate and post-intermediate groups (but none in the elementary group) reported that they were aware of making a distinction between /n-/ and /l-/ in spontaneous speech. About 20%, 35% and 45% of participants in each group started to be aware of making a distinction between /n-/ and /l-/ when doing passage reading, word reading and minimal pair reading tasks respectively. All the learners reflected that their awareness of making a distinction between /n-/ and /l-/ was lowest in spontaneous casual production (i.e. conversation and interviews), and highest in the minimal reading tasks.

Our learners' responses in the structured interviews and the results of the oral tasks indicate that no matter which proficiency group the learners belonged to, they were less likely to suffer L1 interference when their phonological awareness of making a distinction between /n-/ and /l-/ was increasingly raised.

#### 4.4. Replacement of /l-/ by /n-/ by the three proficiency groups

When students were more aware of making a distinction between /n-/ and /l-/, and more aware of not producing /n-/ as /l-/, the phenomenon of hypercorrection of /l-/ appeared (i.e. replacement of /l-/ by /n-/.) As shown in Figure 3, students were most likely to replace /l-/ by /n-/ in the minimal pair reading and repetition tasks which raised the most phonological awareness. But the percentage of words with this hypercorrected replacement in these two tasks was not high - about 10 to 15%.

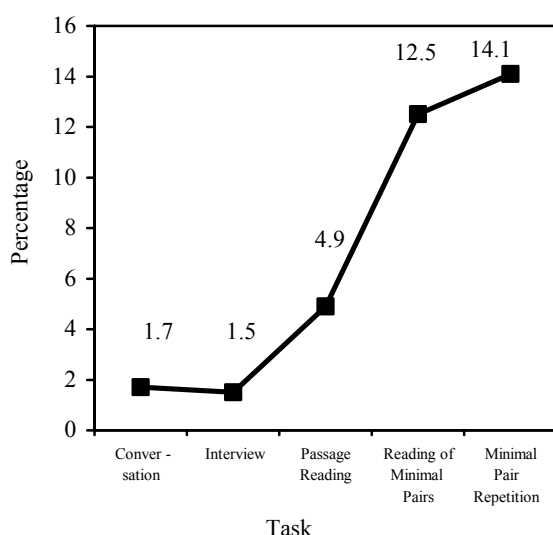


Figure 3: Mean Percentage of words with replacement of /l-/ by /n-/ in each oral task

## 5 Implication

This study has implications for the pedagogy of L2 phonology. The findings of this study highlight the

importance of early development of phonological awareness in L2 learning. As shown in this study, phonological awareness plays an important role in the production of accurate pronunciation in L2. Minimal pair pronunciation practices could be designed to raise learners' phonological awareness between two similar sounds. Focused language input should be provided with learners to arouse their attention to the fossilized sounds.

We propose that subconscious phonological awareness raised by the provision of L2 focused language input and practice, together with conscious learning of L2 phonology and phonetics can further enhance accurate pronunciation in the L2. Using the production of /n-/ and /l-/ in English by Cantonese speakers as an example, if the differences of manner of articulation and phonetic features between these two sounds are explicitly introduced to learners, it will be less likely for them to substitute /l-/ for /n-/. /n/ and /l/ are both alveolar, but /n/ is a nasal sound and /l/ is a lateral. When the /n-/ sound is produced, the soft palate is lowered, which opens a passageway into the nasal cavity and air continually escapes through the nose. When the /l-/ sound is produced, air comes out from the two sides of the tongue and learners could be able to feel the airflow.

Furthermore, as reflected in our study, pronunciation accuracy varies with speech styles. We are more likely to find the occurrence of L1 negative phonological transfer in casual speech than in careful speech. In the training of L2 pronunciation, different speech styles could be used. We could make use of learners' existing phonological variations in different speech styles for sound perception and distinction training. We aim not only at learners' accurate pronunciation in careful reading, but also at high pronunciation accuracy in casual speech, such as casual conversation and interview. The high accuracy in casual speech indicates the complete acquisition of L2 phonology.

## 6 Conclusion

This study shows that L1 phonological transfer in respect of the replacement of /n-/ by /l-/ appeared in the English interlanguage of Hong Kong Cantonese speakers. The occurrence of L1 phonological interference negatively correlated with learners' English proficiency. The replacement of /n-/ by /l-/ might become a natural and habitual formation, and therefore even more advanced learners also incorrectly pronounced /n-/ as /l-/ in casual production. Furthermore, the extent of L1 phonological interference varies with different speech styles which require different amounts of attention to pronunciation, and raise different

degrees and levels of phonological awareness. There is an interaction between L1 phonological interference and L2 phonological awareness. When phonological awareness in the L2 is increasingly raised, L1 phonological interference is less likely to appear. The findings of this study shed light on the pedagogy of L2 phonology.

## References

- Au, Y.A. (2002). *Markedness theory and syllable structure difficulties experienced by Cantonese learners of English*. Unpublished PhD thesis, University of Reading, Reading.
- Bauer, R. S. (1982). *Cantonese sociolinguistic patterns: Correlating social characteristics of speakers with phonological variables in Hong Kong*. PhD dissertation, University of California, Berkeley.
- \_\_\_\_\_. (1997). *Modern Cantonese phonology*. New York: Mouton.
- Bialystok, E. (1978). A theoretical model of second language learning. *Language Learning*. 28, 69-84.
- Bolton, K., & Kwok, H. (1990). The dynamics of the Hong Kong accent: Social identity and sociolinguistic description. *Journal of Asian Pacific Communication*. 1, 147-172.
- Bourgerie, D. S. (1990). *A Quantitative Study of Sociolinguistic Variation in Cantonese*. PhD Dissertation, Ohio State University, Ohio.
- Chan, Y.W. (2000). English and Cantonese Phonology in Contrast: Explaining Cantonese ESL learner's English pronunciation problems. *Language, Culture and Curriculum*. 13, 67-85.
- Ellis, R. (1987). Interlanguage variability in narrative discourse: Style shifting in the use of past tense. *Studies in Second Language Acquisition*. 9, 1-20.
- \_\_\_\_\_. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ho, M. T. (2004). *A sociolinguistic investigation of Cantonese in Hong Kong and Guangzhou*. Unpublished PhD Thesis, University of Essex, Essex.
- Hulstijn, J., & Hulstijn, W. (1984). Grammatical errors as a function of grammatical constraints and explicit knowledge. *Language Learning*. 34, 23-43.
- Hung, T. (2000). Towards a phonology of Hong Kong English. *World Englishes*. 19, 337-356.
- \_\_\_\_\_. (2002). Language in contact: Hong Kong English phonology and the influence of Cantonese. In Kirkpatrick, A. (Ed.). *Englishes in Asia: Communication, identity, power and education*. Melbourne: Language Australia. 191-200.
- Labov, W. (1972). *Sociolinguistic pattern*. Philadelphia: University of Pennsylvania Press.
- Ochs, E. (1979). Planned and unplanned discourse. In Givon, T. (Ed.) *Syntax and semantics, Vol. 12: Discourse and semantics*. New York: Academic Press. 51-80.
- Rutherford, W. & M. Sharwood Smith (1988). Consciousness raising and universal grammar. In Rutherford, W. & M. Sharwood Smith (Eds.) *Grammar and second language teaching*. New York: Newsbury House. 107-116.
- Schmidt, R. W. (1990) The role of consciousness in second language learning. *Applied Linguistics*. 11, 2, 129-158.
- \_\_\_\_\_. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In Schmidt, R. (Ed.) *Attention and awareness in foreign language learning*. Honolulu: University of Hawaii, Second Language Teaching and Curriculum Center. 1-63.
- Tarone, E. (1982). Systematicity and attention in interlanguage. *Language Learning*. 32, 69-82.
- \_\_\_\_\_. (1983). On the variability of interlanguage systems. *Applied Linguistics*. 4, 143-163.
- \_\_\_\_\_. (1988). *Variation in interlanguage*. London: Edward Arnold.
- Yeung, S.W.H. (1980). Some aspects of phonological variation in the Cantonese spoken in Hong Kong. M.A. Thesis. University of Hong Kong.

## Acknowledgement

The work described in this paper was supported by the GRF grant from RGC Hong Kong (Project No. CityU 1438/05H), and by CityU SRG Grant (7001657). The whole research investigated the merging of /n-/ with /l-/ in Hong Kong Cantonese, Putonghua and English. This paper only reported the results regarding Hong Kong English.

# Raising Awareness of English Prosody among Thai University Students

Soisithorn Isarankura

Language Institute, Dhurakij Pundit University

soisithorn@yahoo.com

## Abstract

This paper examined the extent to which the teaching of prosody improved the performance and intelligibility of nonnative speaker's speech. Thirty Thai students' speech samples were drawn from a pool of audio-recordings of 108 first-year English-major students in an English Phonetics course. Based on scores of three pronunciation tests, speech data of fifteen students who scored highest and fifteen students who scored lowest were selected and placed in the High and Low groups. The students recorded their speech of 50 isolated words and a short text before and after the three-week instruction of English prosodic features. The results revealed that both groups showed statistically significant improvements in their pronunciation of prosody. The high group made most improvement in their performance on pausing, whereas the low group showed the greatest degree of improvement in word stress on individual words. The data showed that L1 transfer decreased and intelligibility increased significantly in the high group, but in the low group these results were not statistically significant. The information gained from the study suggests that although some prosodic features may be more challenging than others, it presents an argument in support of incorporating prosody in EFL classes for intelligible pronunciation.

## Keywords

Prosody, pitch, stress, rhythm, pausing, and intonation

## Introduction

Prosody—i.e. stress, rhythm, pausing, intonation, etc.—plays a crucial role in communication (Morley, 1991; Levis, 1999; Hahn, 2004). Anderson-Hsieh et al. (1992) maintains that prosody has a greater influence on comprehensibility and intelligibility than the other elements in pronunciation—i.e. segmentals such as consonants and vowels. A good understanding of prosody is important for nonnative speakers to clearly interpret native speakers' intended messages.

Equally, mastery of prosody is essential if nonnative speakers are to make themselves understood (Taylor, 1993). Errors in prosody can lead to misinterpretation of speaker's intention or even serious miscommunication.

Despite its importance, prosody has been a neglected feature in English language instruction. A few possible reasons that account for the neglect could be that prosody is too complicated and that it appears to be unteachable and thus possibly unlearnable (Greenwood, 2002; Silveira, 2002). The lack of attention to prosodic features has resulted in limited knowledge about how to integrate appropriate instruction and EFL materials into EFL classrooms.

## 1 Background

As prosody includes a number of complicated features, it helps to have some understanding of some of the basic concepts regarding these features, and we shall look briefly at these first.

*Prosody*, also known as suprasegmentals, refers to features that typically extend more than the level of individual sounds like consonants and vowels. English prosodic features mainly include stress, rhythm, pausing and intonation. These features are characterized by the prominent quality of pitch, duration, and loudness.

*Pitch* is determined by the fundamental frequency of the sound made by the vibration of the vocal cords and causes us to hear notes and sounds as relatively 'high' or 'low'.

*Stress* functions at both the word level and the sentence level. *Word stress* refers to the way in which greater prominence in pitch, duration and loudness is placed on a certain syllable than others in a word (Roach, 2000). Different languages have different systems of word stress placement. English, for example, manifests a free stress system, whereby stress is movable and can potentially fall on any syllable of polysyllabic words. Thai is a kinetic tone language and is characterized by a fixed stress system. The primary stress in Thai words almost always falls on the last syllable. This is why most Thai EFL learners have difficulty

pronouncing English words with correct stress placement.

At the *sentence level*, not all words in a sentence receive the same amount of stress. In English, *content words* (words that carry information such as nouns, verbs, adjectives and adverbs) are usually stressed, whereas *function words* (words that carry grammatical functions such as articles, prepositions and auxiliaries) are not, unless marked. In connected speech, sentence stress in English often introduces new information to the discourse context or marks contrastive information. Thai EFL learners often fail to stress new information and de-stress old information. Instead, they tend to stress all words with more or less equal pitch, duration and volume, without one prominent stress to indicate new or contrastive information. This is supported by Wennerstrom's (1994) findings that native speakers of Thai, Japanese and Spanish failed to use pitch movement to signal new or contrastive information in the same manner that native English speakers do.

In English, the combination of word and sentence stress contributes to *rhythm*. The stressed syllables of the content words are aligned regularly into rhythmic beats, which normally occur at regular time intervals. English is thus considered to have a stress-timed rhythm. The problems in Thai speakers of English appear to rise from the stress placement when Thai speakers usually place stress on almost every syllable in English including function words.

*Pausing* is also important in English in that it marks boundaries of tone groups, which serves to indicate syntactic units. A tone group contains a stressed syllable, usually found on the last prominent word assumed to be 'new' information upon which the speaker wishes to draw the hearer's attention. Thus, pausing is important for effective communication as far as the hearer is concerned.

*Intonation* is the melody of pitch changes of the utterance. It is one of the most significant features in English that guides the listener and helps them to follow. Intonation functions as a signal of grammatical structure and marks sentence, clause and other boundaries. Intonation is used to convey contrasts between different question types and the ways in which questions differ from statements. It also conveys distinctive meanings as well as the speaker's attitude, emotion, or even social backgrounds. As Thai is a tone language, intonation in Thai is a complex interplay between tone, word stress and sentence stress. The problems for Thai speakers come from the transfer of constraints of the pitch movement in different types of Thai syllable structures into the pronunciation of English by assigning the Thai

system of intonation contours to English discourse.

Although the important role of prosody in determining perceived comprehensibility and intelligibility has recently been recognized among many scholars (e.g., Munro and Derwin, 1995; Hahn, 2004), much less empirical research has been carried out on second language (L2) production of prosodic features than on other language skills. Evidence has been presented that explicit instruction significantly improved learners' performance on prosodic features. Thus, the objective of the present study was twofold:

(1) to evaluate the extent to which explicit training of prosody improves the speech performance of Thai first-year English-major students at Dhurakij Pundit University; and

(2) to investigate the extent to which improvements in prosodic features had an impact on the reduction of L1 transfer and the increase of intelligibility in L2 speech.

## 2 Method

The method involved selecting speech samples of Thai first-year students, rating the speech samples, and then analyzing the samples through statistical procedures.

### 2.1 Speech Samples

The speech samples were taken from a pool of audio-recordings of 108 first-year English major students in an English Phonetics course at Dhurakij Pundit University. From this original sample, thirty recordings were selected based on the speakers' scores on three previous pronunciation tests conducted earlier in the course. Speech data of 15 students who scored highest and 15 students who scored lowest were selected. They were placed in two ability groups referred to as the High and the Low group, respectively. These students had minimal or no previous linguistic knowledge of English prosody prior to the study.

### 2.2 Native Speaker Raters

Three experienced native English speaking teachers (NESTs), one Australian, one American, and one British, participated as native speaker raters of L2 speech.

### 2.3 Instrument

The test materials used in the study consisted of two parts. The first part contained 50 individual English words drawn from an extended text. The second part was the full text from which the 50 words listed in Part 1 were drawn. The text was retrieved from <http://www.manythings.org/listen/sleep.html>. The test material was used at the pre-training (T1) and post-training (T2).

## 2.4 Data Collection

The students were asked to read and audio-record the two parts of the test in a laboratory without the researcher being present before and after a three-week explicit instruction of English prosodic features. A total of six ninety-minute training sessions of English prosody, in theory and practice, were conducted to raise the students' awareness and help them understand the effective use of word stress, sentence stress, rhythm and intonation. Before the actual recording, the students were told to practice their readings silently, then read the speech material once at normal speaking rate.

## 2.5 Rating of Speech Samples

The two parts of the test were evaluated separately. The first part which contained 50 isolated words was assessed for the performance on word stress by the researcher. Each word pronounced with correct stress placement received 1 point, whereas words produced with stress placed incorrectly or missing entirely received no point. The total score for the first part was 50.

The second part (i.e. the full text) was rated by 3 experienced NESTs. Twenty speech data of T1 and T2 were randomly distributed to each rater. Judgments were made based on a 5-point scale (1 = poor, 5 = excellent) for the performance on 3 prosodic features, referred herein as *Discourse Prosodic Features*—i.e. sentence stress and rhythm, pausing, and intonation. As sentence stress and rhythm contribute closely to each other and are difficult to identify separately, it was decided that both be rated as one feature. The total score for the evaluation of these 3 aspects was 15.

In addition to the discourse prosodic features, the 3 native speaker raters were also asked to judge the students' performance with regard to first language (L1) transfer and intelligibility on a 5-point scale. For intelligibility, 1 is low and 5 is high. For the effect of L1 transfer, 1 denotes poor performance due to excessive transfer of L1, and 5 represents near-native performance as a result of the students' success in eliminating the influence of L1 from their speech production.

Prior to the evaluation, the raters were given an intensive training session, in which they were asked to rate eight speech samples, using headphones in a laboratory. Each rater assigned scores for each of the 5 aspects to the 8 speech data on a separate score sheet. Scores were compared and disagreements over the judgment were resolved through discussion. Then, each rater was asked to individually listen to 20 recordings in their own time, and each completed a separate score sheet.

## 3 Results and Discussions

### 3.1 Word Stress

The group summary of the highest and lowest scores, the mean values and standard deviations of the pre-training (T1) and post-training (T2) data on word stress are presented in Table 1.

Table 1: Score Summary of Word Stress by Group

Word Stress					
Group		Highest Score (50)	Lowest Score (50)	Mean Score	Std. Deviation
High	T1	46	34	42.43	0.975
	T2	50	42	47.07	0.621
	Diff	4	8	<b>4.53</b>	-
Low	T1	40	15	32.6	6.653
	T2	46	23	40	5.819
	Diff	6	8	<b>7.40</b>	-

The broader range of scores and the higher standard deviation in the low group indicate a wider within group variation among the low-ability students. The higher mean score difference in the low group (i.e. 7.40 > 4.53) suggests that the low-group students had a greater degree of improvement in word stress than the students in the high group as a result of the intervention. However, one must caution the possibility of a 'ceiling effect' experienced among the high-ability students as many of them already obtained a high score at pre-training (T1). From the tabulation of scores at the individual level, 12 students in the high group scored more than 40 at T1. Among these, 8 students scored 45 and above. At T2, 2 students obtained the full score of 50, while 5 students made only one error after the training. In the low group, the highest score at post-training (T2) was only 46. It is important to note, however, that many students in the low group, particularly those who obtained a very low score at T1, made a surprisingly large improvement at T2. Seven students placed stress correctly at T2 on at least 8 words more than they did at T1, whereas in the high group, the degree of improvement was lower. The range of improvement in the high group was between 2 and 8, with the mean value of 4.53. The range was wider in the low group—i.e. between 2 and 15, with the mean value of 7.40. The t-test was used to investigate whether the improvement in each group was significant and the results revealed that both groups made a statistically significant improvement at the .05 level.

### 3.2 Discourse Prosodic Features

Table 2 shows the score summary of the discourse prosodic features (sentence stress and rhythm, pausing, and intonation) performed in connected

speech at T1 and T2 by group.

The broader range of scores at T1 in the high group (i.e. scores ranging from 4 to 12, with the mean value of 8.33) and at T2 (i.e. scores ranging from 6 to 14, with the mean value of 9.93) and the higher standard deviation indicate a wider within-group variation among the high-ability students in their pronunciation in connected speech.

Table 2: Score Summary of Discourse Prosodic Features by Group

Discourse Prosodic Features					
Group		Highest Score (15)	Lowest Score (15)	Mean Score	Std. Deviation
High	T1	12	4	8.33	2.769
	T2	14	6	9.93	2.314
	Diff	2	2	<b>1.6</b>	-
Low	T1	9	3	5.53	2.167
	T2	10	3	6.53	1.922
	Diff	1	0	<b>1.0</b>	-

In contrast to word stress, the average improvement in the high group was greater than that in the low group (i.e.  $1.6 > 1$ ). It is important to note that the overall improvement in the performance of the low group on discourse prosodic features in connected speech was lower than their improvement on the pronunciation of word stress. This could suggest that sentence stress, rhythm and intonation at the discourse level are more challenging for the low-ability students than word stress at the level of isolated words. This is in line with Benrabah's (1997) report that word stress seems to be the most accessible feature of all prosodic aspects. He suggested that word stress should therefore receive a high priority and serve as a starting point in pronunciation teaching.

The mean differences of 1.6 and 1 in the high and low groups were tested and the t-test scores revealed that both groups made statistically significant improvement in the performance of discourse prosodic features at the .05 level.

### 3.3 Word Stress and Discourse Prosodic Features

In order to investigate the order of improvements in the students' performance on the four prosodic features, the mean differences between T2 and T1 in both groups on each feature are presented in Table 3. Please note that because the score on word stress was not reported on the same scale as were the three discourse prosodic features rated, each raw score on word stress was divided by 10 to equate the scale rated on each discourse feature. The mean value was then calculated based on the equal score of 5.

As Table 3 indicates, the high-ability students

demonstrated the greatest degree of improvement in the performance on pausing (i.e. mean difference = 0.74) and the lowest degree of improvement in sentence stress and rhythm (i.e. mean difference = 0.27). This suggests that, while pausing appeared to be picked up more easily after the training among the high-ability students, they found stress and rhythm in connected speech to be the most challenging of all discourse prosodic features. In examining T1 and T2 scores in the high group at the individual level, no students deteriorated in their performance on pausing at post-training, but for rhythm and intonation, one student backslid toward non-targetlike pronunciation at T2.

Table 3: Mean Score of Four Prosodic Features by Group

Group	Mean	Word Stress (5)		Stress and Rhythm (5)		Pausing (5)		Intonation (5)	
		T1	T2	T1	T2	T1	T2	T1	T2
High Group	Mean	4.25	4.7	2.93	3.2	2.66	3.4	2.73	3.33
	Mean Diff	0.45		0.27		0.74		0.60	
Low Group	Mean	3.26	4.0	1.8	2.2	1.8	2.33	1.93	2.0
	Mean Diff	0.74		0.40		0.53		0.07	

In the low group, the mean score differences indicate that the improvement was greatest in the performance on word stress (i.e. mean difference = 0.74) and lowest in the pronunciation of intonation (i.e. mean difference = 0.07). This suggests that the low-ability students found word stress to be the most accessible aspect, while intonation appeared to be the most challenging feature for them. The results at the individual level show that 4 students in the low group deteriorated at T2 in their performance on intonation as did 2 students in the pronunciation of sentence stress and rhythm. This is probably due to the students' lack of discourse comprehensibility, resulting in their difficulty understanding the effective use of intonation. Further, insufficient knowledge of word classes in English is likely the source of difficulty for the low-ability students in putting sentence stress on content words which constitutes proper stress-timed rhythmic patterns. As all students made some progress on word stress and no students deteriorated in the pronunciation of pauses, these two features appear to be an area of maximum overlap of teachability and communicative importance.

To summarize the data in Table 3, the ordering of improvements (from highest to lowest) in the pronunciation of prosodic features in each group is shown below:

High: Pausing > Intonation > Word Stress > Sentence Stress and Rhythm

Low: Word Stress > Pausing > Sentence  
Stress and Rhythm > Intonation

Table 4: Score Summary on Four Prosodic Features by Group

Four Prosodic Features					
Group		Highest Score (20)	Lowest Score (20)	Mean Score	Std. Deviation
High	T1	16.5	7.8	12.587	3.0190
	T2	18.5	10.2	14.640	2.4136
	Diff	2.0	2.4	<b>2.053</b>	-
Low	T1	12.4	4.5	8.793	2.4315
	T2	13.9	6.4	10.533	2.1757
	Diff	1.5	1.9	<b>1.740</b>	-

If we add up the scores of all prosodic features to examine the overall progress in the two groups, as shown in Table 4 above, we find that the broader range of score and higher standard deviation in the high group indicate a wider within group variation among the high-ability students in the pronunciation of English prosody. The mean score differences of 2.053 and 1.740 show that the high group made a greater improvement than the low group. The results from the t-test scores revealed that the improvements in both groups were statistically significant at the .05 level.

### 3.4 L1 Transfer

The group summary of scores at T1 and T2 on L1 transfer is presented in Table 5.

Table 5: Score Summary of L1 Transfer by Group

L1 Transfer					
Group		Highest Score (5)	Lowest Score (5)	Mean Score	Std. Deviation
High	T1	4	1	2.47	1.060
	T2	5	2	3.20	1.014
	Diff	1	1	<b>0.73</b>	-
Low	T1	3	1	1.80	0.561
	T2	3	1	1.93	0.799
	Diff	0	0	<b>0.13</b>	-

The data in Table 5 show broader ranges of scores and higher standard deviations at both T1 and T2 in the high group, suggesting that the high-ability students varied more in terms of applying the L1 sound system in their speech production. However, on the whole, the high group demonstrated a greater improvement than the low group in their attempt to eradicate L1 influence from their speech, as shown by the higher mean score difference of 0.73 in the high group as opposed to 0.13 in the low group. The mean differences were tested and the results from the t-test revealed that the high group made statistically significant improvement at the .05 level, but the

improvement in the low group was not significant. This implies that L1 transfer appears to pose a greater impact on the low-ability students. An interesting point to note here is that although the students in the low group improved significantly in their pronunciation of prosody, this does not necessarily mean that they could manage to reduce the influence of the L1 sound system when pronouncing L2 speech. This suggests that, for lower-ability students, more time and effort may be necessary to alleviate the effect of L1 transfer.

### 3.5 Intelligibility

Table 6 presents the group summary of scores on intelligibility at T1 and T2.

Table 6: Score Summary of Intelligibility by Group

Intelligibility					
Group		Highest Score (5)	Lowest Score (5)	Mean Score	Std. Deviation
High	T1	4	1	2.87	1.060
	T2	5	2	3.60	0.986
	Diff	1	1	<b>0.73</b>	-
Low	T1	4	1	2.20	0.941
	T2	4	1	2.27	0.961
	Diff	0	0	<b>0.07</b>	-

The standard deviations displayed in Table 6 show that there was a wider within group variation in the high group than in the low group. The mean score difference between T2 and T1 was higher in the high group (i.e. 0.73 > 0.07), indicating that there was some improvement among the high-ability students, but the improvement was lower in the low group. When examining whether the improvement was significant in each group, the t-test results revealed that the improvement in the high group was statistically significant at the 0.5 level, whereas in the low group the improvement was not significant. As this study investigated the performance of more than one prosodic feature in relation to intelligibility, the results do not allow conclusions to be drawn as to which feature most affects intelligibility and which does not. Because the students in the high group made significant improvements in all prosodic aspects as well as the reduction of L1 transfer effect, this might serve to justify the increase in intelligibility in their speech. In the low group, however, the results showed that the students' performance on prosodic features were judged by native speakers to improve significantly, but their improvement on these features did not help to significantly reduce the influence of L1 transfer on the students' speech, nor did it help to significantly increase intelligibility. As the relationship between each prosodic feature and intelligibility is so complex, further research is

needed which investigates the extent to which each feature has an impact on intelligibility and also on the reduction of L1 transfer.

The findings of this study have shown what aspect of prosody seems to evince the most and least improvements in the high and low groups as a result of the intervention. From the speech data of the individual students in this study, cases were found when no changes occurred at T2, and some students even deteriorated in their performance on some prosodic aspects. Macdonald et al. (1999) argued that these cases are, in fact, not uncommon. Indeed, an initial lack of improvement or even some deterioration in performance does not necessarily indicate a failure of method. The process of L2 learning depends to a large extent on an individual experience and personality. Some students may not perceive a single learning event as beneficial as others; some may not be as fast at learning as their peers. Although the improvement may not occur immediately, the phenomenon may evidence the learning process at work. For some students, the improvement may start to be noticeable at some later time.

Generally, most EFL teachers expect the change in performance to happen immediately following some classroom activity in connection with the language aspects being focused. When no immediate improvement occurs, many teachers often doubt the technique or materials they use. Some may consider the intervention as a failure and want to abandon the activity altogether. This experiment, as well as many others, was conducted to examine the change in the performance immediately following the intervention. Little research has yet been conducted on measuring the delayed effect, whereby favorable changes may occur after an initial lack of improvement or even deterioration in performance. It should be interesting to examine the delayed effect of the feature(s) covered in earlier classes. Also, it should equally be interesting to investigate the retention (long-term improvements) on the students' pronunciation of L2 feature(s) as a result of the training at some later time.

#### 4 Implications of the Study

The findings of this study provide several implications. First, teachers should attempt to show their students that prosody plays an essential role in communication. Second, the teacher's goal should also be to help students select areas for practice based on empirical findings. This study has made a small contribution toward that goal by offering a means for teachers to select areas of prosodic features and set pedagogical priorities for the student to practice. Indeed, it is hoped that

more research will follow which further investigates what aspects of prosody are more critical than others so that teachers can examine existing practices and develop more informed method of teaching pronunciation in EFL contexts. The present article proposes an argument in support of incorporating English prosody into EFL classrooms for intelligible pronunciation.

#### References

- Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning*, 42/4:529-555.
- Benrabah, M. (1997). Word-stress – A source of unintelligibility in English. *International Review of Applied Linguistics in Language Teaching*, 35/3:157-165.
- Greenwood, J. (2002). The role of teaching English pronunciation: Issues and approaches. *International Conference IPBA* 24-26, September 2002.
- Hahn, L. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, 38/2:201-223.
- Levis, J. (1999). Intonation in theory and practice, revisited. *TESOL Quarterly*, 33/1:37-63.
- Macdonald, D., Yule, G. & Powers, M. (1994). Attempts to improve English L2 pronunciation: The variable effects of different types of instruction. *Language Learning*, 44/1:75-100.
- Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly*, 25/3:481-521.
- Munro, M.J. & Derwing, T.M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45/1:73-97.
- Roach, P. (2000). *English Phonetics and Phonology*. Cambridge: Cambridge University Press.
- Silveira, R. (2002). Pronunciation instruction classroom practice and empirical research. *Linguagem & Ensino*, 5/1:93-126.
- Taylor, D.S. (1993). Intonation and accent in English: What teachers need to know. *International Review of Applied Linguistics in Language Teaching*, 31/1:1-21.
- Wennerstrom, A. (1994). Intonation meaning in English discourse: A study of nonnative speakers. *Applied Linguistics*, 15/4:399-420.

# Has English Ability of the Japanese Improved? —Mext's National Curriculum Standards and Embarrassment for English Teachers and Students

Katsura Yuasa

Kyoto Prefectural Kizu High School

kyjapan1961@yahoo.co.jp

## Abstract

It is often said that communicative English ability of the Japanese is inferior to that of other Asians. Although the Ministry of Education, Culture, Sports & Science repeatedly revise guidelines of English education at secondary schools to improve it, preferable result has not been obtained so far, according to scores of Test of English as a Foreign Language (TOEFL). This report discusses the English education at secondary schools in Japan and inquires the reason why their ability has not improved.

## Keywords

Communicative English ability, The National Curriculum Standards, TOEFL, Secondary Schools Education

## Introduction

It has been criticized that communicative English ability of the Japanese is inadequate compared with other non-English-native speakers including Koreans and Chinese. The criticism is mainly directed toward school education and its supervising body, the Ministry of Education, Culture, Sports & Science (MEXT). (All the Japanese school curriculums and education are based on MEXT's teaching guidelines called "the National Curriculum Standards.") Many Japanese people think that they cannot use English even though they have learned English for 10 years: 3 years at junior high school (lower secondary school), another 3 years at senior high school (upper secondary school) and 4 years at college. These critics claim English teachers should emphasize more on communication in English, especially listening and speaking, instead of grammar and translation, as was emphasized in their school days. Accepting the criticism, the MEXT revised the curriculum standards in 1998 to stress communicative activities in English. It does not mean that English education in Japan has dramatically changed but communicative approach is gradually being adopted. Nevertheless, the revision has not necessarily caused preferable result for English learners. On the contrary, there has been a negative effect on students' lack of linguistic knowledge of grammar and vocabulary.

## 1 From Grammar-Translation Method to Communicative Method

In Japan, foreign languages have traditionally been taught through the grammar-translation method. This is because the primary purpose of learning a foreign language was to absorb advanced knowledge and culture from foreign countries, not for mutual communication. The first language that Japanese studied was Chinese. Centuries ago, precursors invented a way to understand written Chinese with some attached signs which help readers modify the word order to make it same as in Japanese. (The word order of Chinese is S+V+O like English while Japanese uses S+O+V, Japanese also has Chinese characters as well as its native characters, *Hiragana* and *Katakana*.) Dutch emerged as another influential foreign language in the 18<sup>th</sup> century but English replaced it in 19<sup>th</sup> century, due to the beginning of the trade with the USA and Britain. Whatever languages they learned, the Japanese have continued to read them as the same way to read Chinese. It is true that many Japanese people still read English by modifying the word order and translate in their minds.

Nevertheless, communication in English has been becoming more and more necessary for Japanese since the globalization after World War II. Those who had keenly realized their lack of communication competence in English blamed it on the grammar and translation method through which they were taught in their school days. This method, which requires students to memorize a large volume of vocabulary and to translate numerous English sentences led to the popular belief that Japanese people are good at reading and writing English but poor at listening and speaking.

The criticism toward the traditional way of teaching English influenced MEXT to begin developing English learner's communicative competence by revising the curriculum standards in 1989 (They were implemented at junior schools in 1993 and senior high schools in 1994). Its main feature was the introduction of three new courses: *Oral Communication A, B, C*. This is not to say that this drastically changed all the English classes

since teachers could not abandon the grammar-translation method, concerned that many colleges entrance examinations still required minute grammatical knowledge and translation skills. It is true, however, that the new guideline led many teachers to adopt a communicative approach together with the traditional method.

MEXT, however, was not satisfied with the outcome since the criticism toward English education had not subsided. In 1999, new national standards were announced for implementation at junior high schools in 2002 and at senior high schools the following year. With each revision, emphasis on communication in English has become stronger. Overall objectives of foreign language education for senior high school excerpted from *the National Curriculum Standards* in 1982-1993 read “To develop student’s abilities to understand and express in foreign languages, and to deepen the interest in language as well as understanding ways of life and thinking of foreign people.” Compared with this, the present one (2003-2012) is more communication oriented :“to develop student’s practical communication abilities such as understanding information and the speaker’s or writer’s intentions, and expressing their own ideas, and deepen the understanding of language and culture, and fostering a positive attitude toward communication through languages.”

## 2 What the New National Curriculum Standard Has Brought

Has the succession of revisions of National Curriculum every decade brought a preferable result for English learners? We rarely hear that Japanese communicative competence in English has improved and on the contrary the criticism toward English education is still rampant. Critics of English education often cite the score of the Test of English as a Foreign Language (TOEFL), claiming that scores of Japanese examinees is near the bottom among Asian countries.

Here is the ranking of TOEFL iBT (Internet-based Test) mean score Asian countries based on January 2008-December 2008 test data. Compared with the traditional paper-based test, which cannot test speaking ability, iBT can test all of four skills: reading, writing, listening and speaking, though it contains ‘integrated tasks’, which combine speaking with reading and listening, and writing with reading and listening.

According to Table 1, the average score of Japanese examinees ranks the 27th out of 30 Asian countries and regions, followed only by Macau, Cambodia and Lao. A close look at the ranking at the top shows that in the top 9 countries and

regions English is either used as a second language or as one of the official languages since they were all once ruled by Britain or the USA. In addition, some countries like Bhutan have few examinees. Considering these conditions, we should compare the score of Japanese with those of examinees from China and the Republic of Korea, both of which are non-English-speaking countries and have a large number of examinees.

Table 1: Top 10 Countries and Regions, China, Republic of Korea and Japan in TOEFL Internet-based Tests

Rank	Country or Region	Rank	Country or Region
1	Singapore(100)	8	Bangladesh(82)
2	Malaysia(88) Philippines(88)	9	Hong Kong(80)
4	Pakistan(87) India (87)	10	Indonesia(79) Kyrgyzstan(79)
		12	Korea, Republic of (78)
6	Bhutan(85 )	13	China(76)
7	Sri Lanka(83)	27	Japan(66)

Note: Figures in ( ) are average total scores

Source: Test and Score Data Summary for TOEFL Internet-based Tests, 2009 (ETS)

Table 2: Scores of Japanese, Koreans and Chinese

	R	L	S	W	Total
Japanese	16	16	16	18	66
Koreans*	20	19	18	20	78
Chinese	20	18	18	20	76

Note: Koreans\*: people from Republic of Korean

R: reading L: listening S: Speaking

W: writing

Source: Test and Score Data Summary for TOEFL Internet-based Tests, 2009 (ETS)

What this comparison clearly shows that the score of the Japanese is behind those of the Koreans and the Chinese in all of the four categories: reading, listening, speaking and writing. This is contrary to the popular belief that Japanese are good at reading and writing English but poor at listening and speaking it.

The section with the widest gap between Japan and the other two counties is not the speaking or the listening section, but the reading one. If we examine the paper-based test, the result is almost the same. It is said that Japanese scores used to be better than Korean score by the 1960s but they fell behind in the 1970s. At present Japanese are behind in all three sections: Listening, Structure & Written Expression and Reading: Chinese and Koreans are far better in Structure & Written expression and Reading than Japanese.

Kumiko Torikai, Professor at graduate school of Rikkyo University analyzed TOEFL scores by the following age groups: 40s, 30s, 20s, 19-22 and 16-18.

Table 3: TOEFL Score Ranking by Age Group

	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>	4 <sup>th</sup>	5 <sup>th</sup>
J*	30's	20's	40's	19-22	16-18
K*	16-18	30	20	19-22	40
C*	20	19-22	30	40	16-18

Note: J\*: Japanese, K\*: Koreans, C\*: Chinese  
Kumiko Torikai, *TOEFL, TOEIC Nihonjin no eigoryoku*(TOEFL, TOEIC& English Ability of Japanese)

She suggested that the two younger Japanese groups are lower in score than the others, while for Koreans the youngest group (16-18 years old) achieves the highest scores. As for Chinese the group in the 20's obtains the highest scores and the second youngest (19-22 years old) earns the second highest. This data shows that despite the fanfare surrounding the introduction of the new *Oral Communication Courses*, the young people who were taught English after the revision of the educational guidelines had the lowest scores of all ages.

It is impossible to determine one specific reason why the score of Japanese has not improved, however, some possible reasons can be cited as follows.

1. The time allotted for reading has decreased in the classroom due to the increase in time for listening and speaking.
2. The ability to grasp sentence structure has diminished since rapid reading is emphasized instead of intensive reading and because less emphasis is being placed on grammar.
3. Emphasis on guessing the meaning of unknown words from the context and a movement away from memorizing vocabulary has decreased reading ability in the long term.
4. Not all of Japanese examinees seriously aim study abroad. Some of them take TOEFL to evaluate their English ability and motivate themselves. It is natural that their scores are not very high.

### 3 The Next Revision of the National Curriculum Standards.

The criticism of the recent decline in academic standards has been leveled not only against English education but against subjects other than English. The critics claims that today's decline is caused by the current guidelines, which are based on the more relaxed education policy (*yutori kyouiku*) implemented to ease the previous cramming system

of education. The rising criticism obliged MEXT to revise the national curriculum standards again to be implemented in a few years.

Here are major points for change in English education.

1. Elementary schools will introduce English (most of them have already started in reality).
2. Junior high schools will increase English classes from 3 hours to 4 per week.
3. Compulsory words to be learnt at junior high school will increase from 900 to 1200 and the same vocabulary and grammatical points are taught repeatedly.
4. English courses for senior high school will be revised as follows.

#### Present Courses (~2012)

English I · II  
Oral Communication I · II  
Reading  
Writing

#### New Courses (2013~)

Communication English Basic I · II · III  
English Expression I · II  
English Conversation

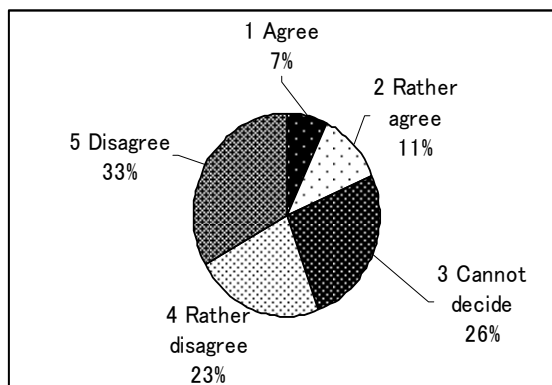
5. The guideline for senior high school suggests that English should be taught in English in principle.

Though MEXT is inclined to continue to reinforce communication in English as a whole, it aims at increasing linguistic knowledge, such as vocabulary and grammar for junior high students, realizing that the current relaxed guidelines have led to a lack of linguistic knowledge, and not the improvement of listening and speaking skills. At senior high school, however, all English courses will be communication-oriented and aimed at fostering a positive attitude toward communication through English. School teachers suspect now that it will be more difficult for students to acquire enough basic linguistic knowledge required for communication in English and college entrance examinations.

The most controversial revision is No. 5 above, which means that 'a direct method' will be adopted so as to give students more chances to use English. Although the idea itself is commendable and a lot of schools have already implemented the method, many teachers suspect that it will be difficult to adopt the policy uniformly across all upper secondary schools in Japan for such reasons as students' abilities and the requirement of detailed grammatical knowledge and translating skills for college entrance exams requiring. A questionnaire conducted with approximately 100 3<sup>rd</sup> graders of Kyoto Prefectural Kizu High School

(17-18 years old) shows that over the half of them oppose the revision according to the following graph.

Table 4: "Do you agree or disagree with the idea that English should be taught in English?"  
Based on the Questionnaire for 3<sup>rd</sup> graders of



Kyoto Prefectural Kizu High School

According to the questionnaire, students who agreed said that they could improve their listening skills and acquire English expressions and vocabulary subconsciously. Others worry that their understanding of English would decrease, classes would not proceed smoothly, and that they might lose their motivation to keep up with the teaching if they couldn't follow it. The 'direct method' is a good way to teach a foreign language to help students get into the habit of thinking in the target language without translation if they are motivated to learn it. Some claim that students will become successfully accustomed to the method quickly even if they are first confused with it. Actually many teachers have adopted this method, while others suspect that it might increase the number of the students who hate English because not all students are sufficiently motivated to keep up with classes taught only in English. We can assume that the adoption of the 'Direct method' has to be left to each school's judgment.

## Conclusion

The revision of National Curriculum Standards approximately every decade has not necessarily improved the English ability of Japanese people so far. It is undeniable that every revision has some defects and whatever revision is made, improving English ability in a very short time is impossible. English teachers have to be aware of the necessity of expanding students' knowledge of grammar and vocabulary to improve their communicative competence in English and too therefore, much emphasis on speaking and listening should be avoided. Although college entrance examinations have been said to be one of the barriers to fostering

communicative English at Japanese high schools, the truth is that fewer and fewer colleges are requiring examinees to have translation skill and trivial grammatical knowledge. In the end, it is the English teachers who actually control the classroom and English education rather than the National Curriculum Standards.

## References

- ETS, TOEFL Internet-based and paper-based Tests January 2008-December 2008 Test Data. ETS, Available from: URL <http://www.ets.org/toefl>
- Kumko Torikai, (2002) TOEFL, TOEIC Nihonnji no eigoryoku (TOEFL, TOEIC& English Ability of Japanese), Kodansha, Tokyo
- The Ministry of Education (1977) Koutougakkou gakushu shidouyoryo shinkyutaisho hyou (Senior High School National Curriculum Standards), the Ministry of Education, Tokyo
- The Ministry of Education, Culture, Sports & Science (2008) Chugakkou Gakushu Shidouyoryo Shinkyutaisho Hyou (Comparison between the Present and New Junior High School National Curriculum Standards) 2009 Available from [http://www.mext.go.jp/a\\_menu/shotou/new-cs/youryou/chu/index.htm](http://www.mext.go.jp/a_menu/shotou/new-cs/youryou/chu/index.htm)
- The Ministry of Education, Culture, Sports & Science, (2008) Koutougakkou gakushu shidouyoryo shinkyutaisho hyou (Comparison between the Present and New Senior High School National Curriculum Standards) 2009 Available from [http://www.mext.go.jp/a\\_menu/shotou/new-cs/youryou/index.htm](http://www.mext.go.jp/a_menu/shotou/new-cs/youryou/index.htm)

# Epistemic Modals by Korean University Students: Comprehension vs. Production

Hera Chu

English Department, Yong In University, Korea

e-mail: hera\_chu@yahoo.co.kr

## Abstract

Epistemic modals (EM) refer to the expressions presenting a writer's or speaker's evaluation of, and the degree of confidence in the knowledge being described. Much research on EM has recently been done in English as foreign language settings. However, it tends to be limited in describing the learners' production of EM, especially their preferences of different epistemic grammatical and lexical forms. Without researching how exactly learners perceive EM, the results of EM production cannot fully explain the process of acquiring EM as well as their preferences on epistemic devices. Thus, this study explores the correlation between production and comprehension of EM by 49 Korean university students. Two kinds of production test were administered. The participants were first asked to describe or respond to a situation open to various interpretations so that EM can be used for deduction. Then they were also asked to rewrite sentences using appropriate EM. Semantic contrast pairs of EM were designed and used for the comprehension test. Correlation results suggest the semantic distinction ability of EM may develop in parallel with syntactic competence of EM. Syntactic

competence combined with semantic competence of EM is essential in the actual uses of EM.

## Key words

epistemic modals, comprehension, syntactic competence, production

## Introduction

Epistemic modals are used to express the degree of speaker or writer commitment to the truth of the proposition expressed in their utterance (Coates, 1983; Palmer, 1990). Semantically, they involve notions of possibility and necessity, and logical probability of a conclusion. Examples of epistemic modal uses are as follows,

- 1) You must be John's wife. (Necessity)
- 2) It may rain tomorrow. (Possibility)
- 3) He can't be from the U.S. He doesn't speak English. (Logical probability)

Marking the above-mentioned meanings requires raising personal opinions and making evaluations on what he/she says or write. And in this process, epistemic modals with other epistemic devices such as adverbials (e.g. certainly, probably), mental verbs

(e.g. think, guess, suppose), and adjectives (e.g. possible, likely) become a linguistic means to show one's thoughts such as personal opinions and evaluations (Biber et al., 1999).

Despite this value of EM as one of the most important communication means, modal verbs, particularly EM, are often considered redundant. There have been a number of studies on clarifying the concept of EM in EFL<sup>1</sup>, but research on EM in terms of how learners understand and produce EM has not been very brisk. Or it tends to be limited in describing the learners' production of EM, especially their preferences of different epistemic grammatical and lexical forms (Chu, 2007; Oh, 2007). However, without researching how exactly learners perceive EM, the results of EM production cannot fully explain the process of acquiring EM as well as their preferences on epistemic devices. Thus, this study explores the correlation between production and comprehension of EM by 49 Korean university students.

## 1 Methods

### 1.1 Participants

Forty-nine students enrolled in Yong In University, Korea participated in this study. Their level of English proficiency varied from high elementary to intermediate.

---

<sup>1</sup> EM is considered one of the most difficult areas for EFL students to grasp due primarily to the multiple meanings attached to them. Modal verbs can be categorized into two by meaning: epistemic and deontic modals. Deontic modals mark the notions of obligation and permission (Palmer, 1986).

### 1.2 Procedures

Two types of tests involving comprehension and production of epistemic modals were given to the participants. The production test was administered prior to the comprehension test. The participants were first asked to write their answers freely in response to a situation about which various interpretations can be made. As the questions in this production test were designed to lead the participants to use a certain degree of epistemic sense in their answers, it was expected their ability to use epistemic modals in their writing would be revealed. However, there was no guarantee that they would bring their epistemic evaluation into their answers relying solely on the modal verbs, if not at all.

In order to examine the participants' ability to use modals only in their sentence, a rewriting task was designed. In this test, the participants had to rewrite sentences using one of the seven modal verbs given. The sentences given for the rewriting task included expressions that could be replaced by modal verbs, namely, modal verb substitutes such as *possible*, *perhaps*, and *I'm sure* etc. The focus of this test was on whether they were capable of making grammatically correct sentences using a right form of modals as well as selecting the most appropriate modal for the sentence they have. Thus, two types of production tests, the free writing task and the rewriting task were administered.

One week after the production task, the same participants were asked to complete the gap from the two modal choices given (a possibility vs. necessity contrast pair) according to the context to

figure out their ability to understand the relative strength of probability of epistemic modals. As in the rewriting task, the comprehension task was designed in a way that the participants can infer the epistemic degree of modals from the modal verb substitutes given in the sentence.

### 1.3 Research Questions

- 1) To what extent does comprehension of EM correlate with its production?
- 2) On what conditions is the comprehension ability of EM transferred to the production ability of EM?

## 2 Results & discussion

### 2.1 Comprehension vs. free-writing task

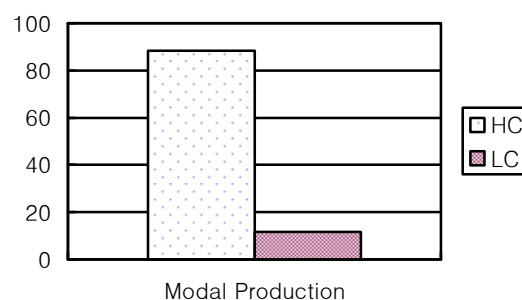
A mean score of the comprehension test was 9.65 out of 15. Of a total of 49 participants, 25 participants above the mean score (here above 10 for the convenience) were classified as a high comprehension (HC) group, and the rest of the participants as a low comprehension (LC) group.

Table 1. Mean score by comprehension group

	N	Mean
High comprehension group	25	11.48
Low comprehension group	24	7.75

The high comprehension group had a mean score of 11.48, while the low comprehension group had a mean score of 7.75.

With relation to modal production in the free writing task, the comprehension score showed a positive relation as in Figure 1.



*F*

figure 1: Percentage of modal production by comprehension score

On the whole, only eight participants used epistemic modals 17 times in their free writing task (production test), and seven out of them belonged to the high comprehension group. In other words, only one student in the low comprehension group expressed his opinion using modals.

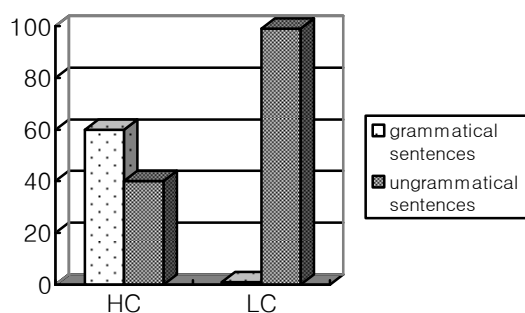
As assumed, however, 20 participants turned out to never employ any of the epistemic devices including modals regardless of their scores acquired in the comprehension test. Considering the mean score of the comprehension test, it seems obvious that the epistemic modal production ability of the participants in this study lags behind their comprehension ability of epistemic modals.

### 2.2 Comprehension vs. Rewriting Task

Comprehension scores seem to have a meaningful relation with the scores obtained in the rewriting task. In the rewriting task, the participants' ability to choose the right form of a modal judging from the modal substitution information was greater in the high comprehension group, although the gap in the mean score between the two groups was only

one point (3.4 > 2.4 out of 6 points. Here, the score indicates the number of right modal choices only). However, 10 participants out of 15 in the high comprehension group scored less than 3, getting only the half right. Similarly, 11 participants (out of 24) from the low comprehension group scored over 3. The results here are not compatible with the scores obtained in the comprehension test. In other words, semantic distinction abilities of modals in the comprehension test do not seem to correspond to those of the rewriting test. In the comprehension test, the key to choosing the right modal lied on the ability to make contrasts between possibility and necessity forms. So to speak, the test required only dichotomy logic, whereas the rewriting task provided participants with a wider range of epistemic modals (e.g. could, might, may, should, must, can't), possibly leading to higher confusions.

More meaningful in the results of the rewriting test was the degree of grammatically correct sentences made by the participants in the two comprehension score groups as shown in Figure 2.



*F*

Figure 2: Percentage of grammatical vs. ungrammatical sentences in the rewriting test by comprehension score

Grammatically correct sentences appeared almost only in the rewriting data of the participants who belonged to the high comprehension group. More than half of the participants (60%) in the high comprehension group were able to make grammatical sentences using a right form of modals in their all sentences. Even when the wrong form of modals was selected and employed in the rewriting test, the participants in the high comprehension group were able to construct right sentences using modals 40% of the time. Although as pointed above, the results obtained between in the comprehension test and in the rewriting test in terms of the ability to distinguish the relative strength of epistemic modals were somewhat inconsistent, the semantic comprehension ability of epistemic modals appeared to relate closely to the syntactic ability to make modal constructions.

### 2.3 Comprehension vs. Production Test

Epistemic modals appeared only 17 times in the data of the free-writing task, 11 times of which were detected in the writings of the participants who belonged to the high comprehension group and the group who scored high in the rewriting task. Four times of modal uses appeared in the free writing data of the high comprehension group and the group who scored low in the rewriting task. The last 2 times were used in the low comprehension group.

Considering the results, it is suggested that epistemic modal verbs are more likely used when the participants have reached a certain level of both

a semantic and syntactic grasp of modals. In this analysis, it is also possible to suggest that the semantic distinction ability of epistemic modals may develop in parallel with its syntactic competence, but which ability was achieved first is uncertain. However, relying only on pure semantic knowledge without the syntactic understanding of how modals are used in sentences does not promise the emergence of EM in the production.

Other forms of epistemic devices used in the free writing task are limited in the use of mental verbs (e.g. I think, I guess) and adverbials (e.g. maybe, certainly). They are used 63 times all together, approximately 3 times higher in the number of use than epistemic modal verbs. Mental verbs and adverbials used for epistemic sense are known to be less complex than modals as they are often used more likely to be a conversational device or an isolated rote member of a complicated system rather than a feature showing a genuine mental state of the writer or speaker (Parafragou, 1998). These other epistemic devices were used almost twice as high in the high comprehension group (42 times) as in the low comprehension group (21 times). Considering that epistemic modals were used almost only in the free writing task by the participants of the high comprehension group, the results can suggest that a high degree of perception and adoption of epistemic modals results in the higher chances of using other forms of epistemic devices.

Nonetheless, as mentioned above, overall 20 participants (45%) didn't use any of the epistemic markers in the free writing task. Looking at the data

of the 20 participants who never used any epistemic devices in their free writing task in more detail, of them, 11 participants were low in their scores of the comprehension test. The rest 9 participants belonged to the high comprehension group and were not able to use modals successfully most of the time in their rewriting sentences. The reason for this could be due to the lack of confidence in their own use of epistemic devices rather than their lacking perception on epistemic notions, considering they are adult learners who are already cognitively mature.

### 3 Conclusion

With relation to modal production, the comprehension score showed a positive relation. On the whole, only eight participants used epistemic modals in their free writing task, and seven out of them belonged to the high comprehension group. However, approximately 40% of the participants never employed any of the epistemic devices including modals regardless of their comprehension scores. Considering the relatively high mean score of the comprehension test, it seems obvious that the production ability of EM by the participants of this study lags behind their comprehension ability of EM. The reason for this could be due to the lack of confidence in their own use of epistemic devices rather than their lacking perception on epistemic notions.

The semantic comprehension ability of epistemic modals appeared to relate closely to the syntactic ability to make modal constructions. Grammatically correct sentences appeared almost only in the

rewriting data of the participants in the high comprehension group.

Epistemic modals appeared 65 % of the time in the production data of the participants who belonged to the high comprehension group and the group who scored high in the rewriting task. The results can show that epistemic modal verbs are more likely used when the participants have reached a certain level of both a semantic and syntactic grasp of modals. In other words, for EM to appear, the semantic distinction ability of epistemic modals ought to develop in parallel with its syntactic competence.

Other forms of epistemic devices were used in the free writing task approximately 3 times higher than epistemic modal verbs. These other epistemic devices were used almost twice as high in the high comprehension group as in the low comprehension group. Considering that epistemic modals were used almost only in the free writing task by the participants of the high comprehension group, the results can suggest that a high degree of perception and adoption of epistemic modals results in the higher chances of using other forms of epistemic devices.

## References

- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *The Longman Grammar of Spoken and Written English*. London: Longman.
- Chu, H. (2007). EFL learners' Critical Thinking in Computer-mediated Discussions. *English Teaching* 62(2), 119-139.

- Coats, J. (1983). *The semantics of modal auxiliaries*. London: Cloom Heim.
- Oh, S.Y. (2007). A Corpus-based Study of Epistemic Modality in Korean Learners' writings in English. *English Teaching*, 62 (2), 147-176.
- Papafragou, A. (1997). Modality in language development: a reconsideration of the evidence. *UCL Working Papers in Linguistics* 9, 1-31.
- Palmer, F. (1986). *Mood and Modality*. Cambridge: CUP
- Palmer, F. (1990). *Modality and English Modals*. (2<sup>nd</sup> ed.). London & New York: Longman.

# Analysis of Listening Test Items in Korean CSAT

Haeng Jung

Honam University

hjung58@daum.net

## Abstract

This study was aimed to analyze the listening test items of the College Scholastic Ability Test (CSAT) in Korea and also to explore in what ways they have been developed to measure students' real listening abilities. To get to the goal, first of all, the test items of each year from 1994 to 2009 were analyzed in terms of 5 categories such as language functions, topics, situations, text type, factual/inferential understanding. Then they were itemized and described under each category in a table to show the changes as the years go by. And the frequency of the items under each category was compared. Finally the changes were examined whether they were made toward the promotion of the Communicative Language Teaching (CLT) in Korean secondary classrooms.

## Keywords

language testing, listening test, washback

## 1 Introduction

The English listening test was introduced as a part of the English test in the College Scholastic Ability Test (CSAT) in Korea for the first time in 1994. It was considered an innovative change of the English test in a highly competitive CSAT test towards the measurement of the oral communicative abilities even though the items were not many. The test was introduced in attempt not only to raise the awareness of the importance of students' English oral communicative abilities in English learning but also to lead to the teachers' integration of teaching of listening into English classes and eventually to the change of the English teaching in secondary schools in Korea from exclusively written English to oral English as well by assessing students' listening abilities in the CSAT (Cheng, 2005; Watanabe, 2000)

The CSAT listening test has been implemented for 15 years from 1994 to 2009. There have been some changes in numbers and the types of the test items. The number of the listening test items comprised only 8 in the tests of 1994 and 1995 and it increased to 10 in the test of 1996. And finally the

portion of the listening test increased almost double to 17 items (34%) from the test of 1997 up to now as of the test for 2009 CSAT. With the increase of the number of the test items, CSAT listening test has also been developed featuring more varied communicative functions and topics.

This study was aimed to analyze the listening test items of the CSAT and also to explore in what ways they have been developed to measure students' real listening abilities. To get to the goal, first of all CSAT English listening test's characteristics were examined. Then the teachers' perceptions about the CSAT listening test and its effects on English teaching and learning were explored. Finally the actual aspects of English teaching and learning practices in secondary English classrooms in association with the test were probed.

## 2. Method

### 2.1. Research questions

This study was aimed to investigate the characteristics of the Korean CSAT English listening test through the item analysis and also to find out the test effects on English teaching and learning in a secondary school context. The following research questions were posed for the purpose of the study.

1. What are the CSAT English listening test's characteristics
2. What are the teachers' perceptions about the listening test and its effects on English teaching?
3. What are the actual aspects of English teaching and learning practices in secondary English classrooms in association with the test?

### 2.2. Data Collection

To answer the research question 1, the test items of CSAT English listening tests conducted over 13 years from 1994 to 2006 were collected and analyzed in terms of 5 categories such as language functions, topic, situation, text type, and factual/inferential understanding. For the research question 2, a questionnaire was employed. A questionnaire consisted of 5 subparts such as basic

information about respondents, overall perceptions of the test, its effects on teaching contents, teaching activities, and teaching materials. For the research question 3, the data were obtained through a total of 8 class observations that were randomly selected and observed for 2 different teachers' classes at 4 different schools.

### 2.3. Data Analysis

To examine the characteristics of the CSAT English listening test, test items from 1994 to 2006 were analyzed in terms of 5 categories such as language functions, topics, situations, text type, factual/inferential understanding. They were itemized and described under each category in a table to show the changes as the years go by. Also the characteristics of the listening test were compared with objectives of listening skills of national curriculum.

Quantitative data from questionnaires were processed with the SPSS (version 12.0) for descriptive statistics and frequency analysis. For teachers' perceptions about the test and its effects on listening contents of the questionnaire, 8 items were given and answered on a 5 point-likert scale, ranging from 1 to 5 (1= strongly disagree, 5=strongly agree) depending on their degree of agreement. Mean score for each item was calculated and compared among items in the same subpart.

Qualitative data from Classroom observation journal written in terms of teaching skills, practice time, materials and activities were also analyzed under each category.

## 3. Results

### 3.1. Characteristics of CSAT English Listening Test

When CSAT English listening test was introduced in 1994 as a way of promoting Communicative Language Teaching (CLT) in Korean secondary classrooms, it seemed to be attempted very cautiously since a big change in the high-stakes test on short notice could impose a big burden on students as well as teachers. The number

of the listening test items comprised only 8 in the tests of 1994 and 1995 and it increased to 10 in the test of 1996 and finally the portion of the listening test increased almost double to 17 items (34%) from the test of 1997 up to now as of the test for 2006 CSAT.

With the increase of the number of the test items, CSAT listening test has been featuring more varied communicative functions and topics since 1997. Table 1, the result of the item analysis of 2000 CSAT English listening test, indicated that the test appeared to have a tendency to be more communicative by including authentic texts such as recorded message, advertisement, announcement, and lecture.

In addition, the test was developed to integrate a speaking skill into a listening though it was tested indirectly through multiple-choice technique within the limit of a paper test. And there appeared 4 or 5 test items at the end that ask students to listen to some conversations or passages and choose the most appropriate dialogue which would come next.

Moreover the number of items that would ask students' inferential understanding increased and students were supposed to employ a high-level thinking skill to infer intention, situation, reason, purpose, main idea, or relationship from the context while they were listening.

### 3.2. Teachers' Perceptions toward the Listening Test

As shown in Table 2, as to the items of 1, 2, 3, and 4 pertaining to the test's contribution to students' learning of listening, teachers reported that the existence of the listening test in CSAT incurred students' motivation in listening (4.25) and also positively affected student's learning of listening (4.18) and listening ability (3.85) and led to the development of students' communicative ability (3.47). However, as to the items such as 5, 6, 7, and 8 about the test's validity and reliability, they showed a mediocre degree of satisfaction. Their levels of agreement were for reliability (3.26), contents (3.15), item type (3.07), and current test (3.02)

Table 1: Item analysis of 2000 CSAT listening test

Item	Function	Topic	Situation	Text Type	Factual/ Inferential (Question)
1	Describing people	Photo	Classroom	Conversation /Picture	Factual (people)
2	Meeting people	Flight schedule	Airport	Conversation	Inferential (place)
3	Expressing feeling	Feeling		Conversation	Inferential (feeling)

4	Understanding message	Message (favor)	Answering machine	Passage (recorded message)	Factual (calling purpose)
5	Shopping	Money	Sports shop	Conversation	Inferential (money)
6	Suggesting	Transportation	Train station	Conversation	Factual (Transport)
7	Talking about family	Family		Conversation	Factual (reason)
8	Understanding topic	Sleeping Exercise	Lecture	Passage (lecture)	Inferential (topic)
9	Understanding situation	Appreciation	Show	Passage (announcement)	Inferential (relation)
10	Understanding purpose	Train schedule	Train station	Conversation	Factual (purpose)
11	Comparing	Job	Advertise	Conversation (advertisement)	Factual (reason)
12	Understanding reason	TV shows		Conversation	Factual (reason)
13	Matching the picture with a dialogue	Situation	Picture description	Conversation (pictures)	Inferential (situation)
14	Checking out	Library use	Library	Conversation	Inferential (next thing)
15	Asking information	Work	Street	Conversation	Inferential (next thing)
16	Finding direction	Direction	Street	Conversation	Inferential (next thing)
17	Speaking after listening			Passage (situation)	Inferential (next thing)

Table 2: Teachers' perceptions or attitudes toward the test

Items	Mean	SD
1. The test helps students develop communicative ability.	3.47	.53
2. The test helps students develop listening ability.	3.85	.65
3. The test promotes students' learning of listening.	4.18	.87
4. The test enhances students' interests or need in listening.	4.25	.72
5. Contents of the test items are valid.	3.15	.74
6. Types of the test items are valid.	3.07	.65
7. The tests are reliable.	3.26	.68
8. The test should be continued as it is.	3.02	.64

(N=385, 1=Strongly disagree, 5=Strongly agree)

### 3.3. Teaching Aspects in the Secondary English Classrooms

To find out the actual aspects of teaching of listening in secondary English classrooms, classroom observation was carried out. Each one of schools in 4 different areas (metropolitan city, city, county, town) was randomly selected. For every school, one class was chosen from regular English classes and the other one was singled out from extra

classes. A total of 8 classroom observations were carried out.

The classroom observation journal was written in terms of language skills, practice time, materials and activities being used in the classes. With regard to language skills being taught in the regular English classes, it was described that reading was the main stream and most of the class time was spent for reading. On the contrary, it was notable

that listening was taught briefly during the introduction stage of each unit sometimes together with speaking, while writing skill was rarely covered.

Nevertheless, observation of English classes outside of regular English classes indicated that listening could not be neglected since it is tested in CSAT. Students were forced to study listening, utilizing morning self-study time before regular classes began or taking away recess time after lunch. Mostly about 20 or 30 minutes were allocated for listening practice almost every day through a school broadcasting system. As for the listening material, EBS listening test preparation books with accompanying CD with them were the most popular, since they were made imitating CSAT listening test format and students could be drilled on the same type of test items as the CSAT listening test.

#### 4. Conclusion and Suggestions

This study was designed to investigate the characteristics of the Korean CSAT English listening test and its effects on English teaching and learning in a secondary school context. The data were collected through the analysis of test items, as well as questionnaire and observation.

The analysis of the development of CSAT listening tests and the test items indicated that the listening skill was likely to play more important part in CSAT English test and the test gradually tended to include more characteristics as a Communicative Language Test. It was clear that the test items covered more variety of communicative functions, or topics and the situations and the texts were adopted from authentic language life. Interestingly, the characteristics of the CSAT listening test seemed to influence the 7th National Curriculum that was implemented from 1998 in a secondary school. It is notable that the test objectives and contents are fairly compatible with the ones about the listening stipulated in the curriculum. This result leads to the conclusion that the test had effects on the curriculum and in part contributed to promotion of Communicative Language Teaching in a secondary classroom.

In another way, some other empirical evidences (Kim & O, 2002; Jung 2008) proved that the listening test had some impacts on teaching and learning practices. Questionnaire analysis indicated that the teachers perceived the listening test had a positive effect on students' motivation, learning of listening, and listening ability, while they showed a mediocre degree of satisfaction about the test reliability, contents, types. Finally, the analysis of the class observation journal manifested that teaching of listening took place not with some

activity-based tasks in a regular class by a teacher but with test preparation books, imitation copies of the CSAT listening test, in a self-study class through school broadcast system.

Based on the above conclusions, several suggestions for CSAT English listening test can be made to bring more beneficial effects to English teaching and learning in secondary school classrooms. First, the types of listening test items need to be varied. Currently the test tends to heavily rely on the type of 'listen and choose'. However, in reality listening is likely to take place accompanying with other responses such as 'listen and act' 'listen and fill in the form' and 'listen and say'. If the test includes different types of test items that are similar to the listening of real life, teaching activities for English listening in a classroom will be varied too.

Secondly, listening test items should be developed incorporating other language skills so that speaking or writing skills can be dealt with together in a classroom. If the test contains integrative test items like 'listen and say' or 'listen and write', teachers will prepare their students to do well on the test items that require overall language skills. It will affect English teaching aspects of secondary schools in which regular class time is spent in mostly reading and other language skills are hardly covered. It is expected that the integrative test will eventually contribute to promoting students' speaking and writing skills in addition to listening.

#### References

- Cheng, L. (2005). Changing language teaching through language testing: a washback study. *Studies in Language Testing*, 21. Cambridge: Cambridge University Press.
- Jung, H. (2008). The washback effects of English listening test in Korean CSAT on secondary school English learning and teaching. *The Linguistic Association of Korea Journal*, 16(3), 143-162.
- Kim, D. J, & O, K. (2002). *Washback on 12th grade English classes of the English tests within Korean university entrance exams. English Teaching*, 57(3), 303-329.
- Watanabe, Y. (2000). Washback effects of the English section of Japanese university entrance examination on instruction in pre-college level EFL, *Language Testing Update* 27, 42-67.

# The Development and Validation of the In-house Can-do Statements for Required Writing Courses

Kahoko Matsumoto

Foreign Language Center, Tokai University,  
mkahoko@tsc.u-tokai.ac.jp

## Abstract

This is an interim report on the process of developing the Can-Do Statements (CDSs) for required writing courses at a Japanese university and the results of their validation. In an EFL situation like that in Japan, some modifications or adjustments of the original CEFR statements were necessary in order to fully realize the purpose and effects of CDS conceptualization. Since the limited exposure to English makes some CEFR statements unrealistic/infeasible and adds more importance to the role of written communication, an attempt was made to develop an EFL version of “localized” CDSs for Japanese college writing courses targeted to different levels of students’ proficiency.

This presentation focuses on, firstly, the challenges faced in the development of localized CDSs that both comprehensively and systematically support a specific curriculum, namely EFL required writing courses at a Japanese university. Secondly, the insights gained from the year-long process of developing and validating in-house CDSs will be discussed. It is hoped that the results, though tentative, may shed light on the future possibilities of using CDSs in an EFL environment, where the proficiency in written communication is a more significant factor than in the ESL counterpart in establishing a successful career in the globalizing world.

## Keywords

Can-do Statements   university writing education  
validation process   EFL situations

## Introduction

In this ICT-driven, globalizing world, English, especially its writing mode has become a dominant medium of communication in all venues; thus increasing importance has been added to English writing education (Warschauer, 2000). However,

English writing education in Japan still focuses on local-level textual aspects (vocabulary, grammar and mechanics) more than global ones such as organization, content and logical development. Concurrently, teachers spend much time correcting such local-level mistakes, so many classrooms are rather teacher-centered.

On the other hand, the concept of Can-do statements (CDSs) as an integral part of the Common European Framework of Reference (CEFR) are based on functional, practical use of language and at the same time, designed to tap “student language awareness” (Council of Europe, 2001).

The application of CDSs to college writing education seems to solve the above two problems; creating the objectives that lead to real-life needs of English writing in the globalizing world, and making the college instruction more student-centered and raising learner awareness on what is required in meaningful written communication, namely more attention to content, organization and flow of discourse.

Yet, we cannot apply CEFR’s writing CDSs which espouse the borderless European situations directly to our EFL situations, where the types and conditions of communication as well as the learning environment are different. Thus, an attempt was made to create a localized/modified CDSs for a university writing course, after a year-long close validation process, both quantitative and qualitative.

## 1. Overall study design

A detailed validation was conducted, both quantitative and qualitative, with various aspects of validity in mind that Weir delineated in his 2005 article for *Language Testing*. First, a series of adjustments were made by reflecting the results of pre-questionnaires given to teachers and students soliciting their responses to the first draft of CDSs (inductive validation). Then, the student self-check lists and teacher assessment scales were developed based on the modified CDSs (as course objectives).

In developing teacher assessment scales, the results of various past studies on the advantages and disadvantages of using holistic and analytic scales were referred to (Weigle, 2002).

Next, the pilot test was conducted to validate the localized CDSs in three different orientations; those of users, curriculum constructors, and assessors. The actual teaching was done based on the syllabi covering the in-house CDSs, also using the student self-check lists and teacher assessment scales. The results of post-questionnaires, class observation and interviews with both students and teachers were used to further modify the CDSs (deductive validation).

## 1.1 Inductive validation

### 1.1.1 Pre-questionnaire to students and teachers

Approx. 300 students in the university's 3 placement levels (Basic, Intermediate and Advanced) answered the pre-questionnaire to check the appropriateness and difficulty level of each Can-do statement. Also 15 teachers who teach them were requested to check the validity of all the CDSs including appropriateness, applicability and difficulty level.

Based on the responses both from students and teachers, a series of modification was made to the original Can-do lists for 3 levels.

### 1.1.2 Preparation of the piloting

First, the curriculum was developed based on the Can-do objectives that include various genres of writing; narrative, expository, argumentative, and work-related functional ones. Then, the teaching material, in the forms of both existing textbooks and customized supplements were collected and/or produced. Secondly, student self-check lists and teacher assessment scales were developed for different genres of writing and for 3 different levels. Basically, the evaluative points in teacher assessment scales correspond to the items in the students' check-lists, but extra attention was paid to make student check-lists easy and simple without the use of technical expressions (see examples: Table 1 and Appendix 1 both are created based on the same CDSs for expository writing).

## 1.2 Deductive validation

A series of pilot test was planned strategically so that the piloting wouldn't disrupt the present teaching as required the syllabus of each course. For each pilot test dealing with a different genre of writing, all the aspects of CDS-based curriculum were tested for its validity, reliability and feasibility, using both quantitative and qualitative analyses.

## 2 The present experiment

In this presentation, the validation results of one genre (unit) of our planned CDS-based curriculum, namely, "expository writing with good supports" was reported as a representative part of our on-going study.

Table 1  
CDS-based teacher assessment sheet for the intermediate level (for expository writing unit)

<b>Writing Assessment Sheet (for Intermediate-level)</b>	
<1=This student cannot do this at all 2=poor 3=fair 4=good 5=excellent>	
<b>1. Organization</b>	This student can write either a long paragraph (more than 200 words) with clear distinction bet. topic sentence, supporting detail and conclusion(ending), <u>or</u> a short essay with clear distinction bet. introduction, body and conclusion.
1-----2-----3-----4-----5	
<b>2. Content</b>	This student can include sufficient knowledge/facts relevant to the thesis statement to keep readers interested.
1-----2-----3-----4-----5	
<b>3. Idea Development</b>	This student can develop ideas in a step-by-step manner from general to specific.
1-----2-----3-----4-----5	
<b>4. Cohesion/Consistency</b>	This student can establish cohesion between sentences and consistency among ideas in a short essay (or a long paragraph).
1-----2-----3-----4-----5	
<b>5. Quality of support and reasoning</b>	This student can use proper supporting statements/examples for the thesis at the paragraph level.
1-----2-----3-----4-----5	
<b>6. Sociolinguistic control</b>	This student can use expressions appropriate to the topic, fulfill situational requirements and address the audience appropriately.
1-----2-----3-----4-----5	
<b>7. Grammatical control</b>	This student can use multiple-clause structures including simple subordinate clauses and those with simple relative pronouns.
1-----2-----3-----4-----5	
<b>8. Conjunctions/transitions/discourse markers</b>	This student can use typical conjunctions/transitions correctly.
1-----2-----3-----4-----5	
<b>9. Vocabulary</b>	This student can use high-frequency vocabulary understandable words and idiomatic language most of the time.
1-----2-----3-----4-----5	
<b>10. Mechanics</b>	This student can use correct punctuation most of the time.
1-----2-----3-----4-----5	

## 2.1 Experimental Design

### 2.1.1 Subjects and procedure

Nine classes of 3 different levels were used for the experiment with approx. 320 students. Teachers were provided with class objectives pertaining "expository writing with good supports" with teaching material. Each class was taught according the given plan for 4 times. The piloting

procedure is as follows.

1. Students responded the self-check lists.
2. Pre-writing test was given in the first class.
3. 4 classes were taught based on the given objectives and using the teaching material provided.
4. Post-writing test was given in the last class (The topics given are of parallel nature for pre- and post-tests)
5. Students responded the same self-check lists again.
6. Teachers assessed the students' pre- and post-writings using the given assessment sheets.
7. Teachers responded open-ended post-questionnaire based on the actual teaching to the Can-do objectives and class observation.

#### 2.1.2 Analysis

The following statistical analyses were conducted and the overall project evaluation was done, reflecting teacher feedback and the interviews with selected students.

1. Comparison between the students' responses to the self-check lists before and after the instruction
2. Comparison between the scores of teacher assessment of pre- and post-writing tests
3. Correlations between students' self-assessment (by self-check lists) and teacher assessment

### 3 Results

The following results were obtained so far.

1. There were significant differences in the average scores of both student self-check list and teacher assessment given before and after the experiment. However, some items (CDSs) representing local-level aspects of writing, notably grammatical and vocabulary control, showed less differences compared to global-level items.
2. Though significant improvements were exhibited by the students at all levels, those in the advanced class rated their own improvements less than the counterparts at intermediate and basic levels. This might be caused by their increased awareness, hence sensitivity to different aspects of writing.
3. As to the students' improvements obtained in terms of the scores of teacher assessment, lower-level students showed more improvements than higher-level ones.
4. Though there was no significant difference among the scores of teacher assessment within the same level, the differences between pre- and post-student check lists (in other words, the improvements perceived by the students themselves) seem to have been affected by the

teacher (who taught the class).

### 4 Implications and future study

We are now in the process of further modifying the "expository writing" CDSs used in this experiment based on different insights gained by both quantitative and qualitative validation.

The same process has been going on in other genres (units) of writing. The difficulty we have faced throughout this process of developing the "localized" CDSs and the curriculum/assessment scales based on them is the judgments of where to draw the line when different kinds of validity are in conflict with each other.

Also, even greater challenge lies in the decision-making that encompasses the limitations deriving from feasibility/practicality and maximization of students' interests, motivation and achievements.

### 5 Appendices

#### Appendix A

Student check-list for expository writing (intermediate level)

Student check list for expository writing (Intermediate Level)		た い い	あ ま り で き ない	ま ま で き る	で き る
1	明確なトピックセンテンス (主題句、それをサポートする詳細な支持文、結論文が区別された長いパラグラフ (200語以上) が書ける	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2	thesis statement (主題) に関連した十分な量の知識や事実を、読み手が興味を持つように組み込んで書ける。	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3	自分のアイデア (内容) を段階的に、一般的なものから具体的なものへと発展 (展開) するように書ける。	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4	文と文をギャップが少いようにつなぎ、文章全体にわたって一貫した考え (主題) で書ける。	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5	パラグラフレベルにおいて、thesis (主題) をきちんとサポートする適切な支持文や例を使うことができる。	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6	トピックやその文章の状況に合う適切な表現を使うことができ、読み手に十分な意図を伝えることができる。	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7	簡単な文法規則や構文を含め文 (= 論議の筋をもつ) を書くことができる。	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8	典型的な接続詞やつなぎ言葉 (= 高校で習う程度) を正確に使うことができる。	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9	一般的によく使われる (= 頻度の高い) 単語やイディオムを正しい場合使うことができる。	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10	句読点 (= パラクチュエーション) だけでなく、大文字の使用、段落の最初を下げることも含む) については、1級レベルに沿って正確に使える。	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### References

- Council of Europe (2001). Common European Framework of Reference for Languages: learning, teaching, assessment. Cambridge University Press.
- Warschauer, M. (2000). The changing global economy and the future of English teaching. *TESOL Quarterly*, 34(3), pp. 511-535.

- Weigle, S. C. (2002). *Assessing Writing*. Cambridge University Press.
- Weir, J. C. (2005). Limitations of the Common European Framework for developing comparable examinations and tests. *Language Testing*, 22(3), pp.281-300.

# Evaluating a Newly Developed Index for Readability Measurement of Japanese EFL Textbooks<sup>1</sup>

Toshiaki Ozasa<sup>1</sup> and George R S Weir<sup>2</sup>

<sup>1</sup>Faculty of Business Administration, Fukuyama Heisei University,

<sup>2</sup>Department of Computer and Information Sciences, University of Strathclyde

tozasa@heisei-u.ac.jp, george.weir@cis.strath.ac.uk

## Abstract

The aims of the present study were (1) to empirically evaluate the validity of a newly developed readability index, *Ozasa-Fukui Year Level, Ver. 2.1*, by correlation analyses, and (2) to compare the readability data of fourteen Japanese EFL textbooks measured by the new index and by *Flesch-Kincaid Grade Level*. The correlation analyses revealed that *Ozasa-Fukui* was the most powerful in differentiating the readability of 84 EFL sentences. It also emerged that the coefficient between Ozasa's criterion and *Flesch-Kincaid* was remarkably high - much higher than that of *Flesch Reading Ease*. The comparative analyses revealed that *Ozasa-Fukui* was more powerful in differentiating Book 1 texts, and that *Flesch-Kincaid* was rather powerful in differentiating the texts of higher years.

## Keywords

Evaluating validity, readability indices, Japanese EFL textbooks, *Ozasa-Fukui Year Level*, *Flesch-Kincaid Grade Level*

## Introduction

In 2008, we developed the *Ozasa-Fukui Year Level, Ver. 2.1* (hereafter, *Ozasa-Fukui*), a new readability index with its application software attuned for Japanese EFL texts. This new readability index is a non-linear function with four variables, sentence length, word length, textbook-based word difficulty and textbook-based idiom-difficulty. It was developed, using the subjective readability evaluation of three sets of representative EFL textbooks of Japan as a dependent variable and the linear function of the four variables as an independent variable. As an independent variable,

the following expression (Diff) was used.

$$\text{Diff} = 0.0993 * \text{words} + 0.4232 * \text{syllables} + 1.0518 * \text{worddiff} + 0.0612 * \text{idiomdiff} + 0.1919$$

In this development, a new readability criterion was employed, which was based on an assessment of 916 textbook sentences by three experienced Japanese EFL teachers well familiar with the EFL textbooks and teaching at middle-grade schools in Japan. The non-linear function thus produced proved to be the highest in explanatory power ( $r^2=.8217$ ) against the solutions of four non-linear methods (Ozasa, Fukui & Hosokawa, 2008; Ozasa, Weir & Fukui, 2008). It was:

$$\text{NewDiff} = 4.6579 * \exp(-17.7116 * 0.3716^{\text{Diff}}) + 1$$
$$[\text{Diff} = 0.0993 * \text{words} + 0.4232 * \text{syllables} + 1.0518 * \text{worddiff} + 0.0612 * \text{idiomdiff} + 0.1919]$$

## 1 Aims of the present study

With these findings as background, a new project was launched to evaluate the validity of the index through correlation analyses and comparative analyses.

More specifically, the present study aims (1) to evaluate the validity of the new index, *Ozasa-Fukui*, by correlation analyses, and (2) to measure the readability of fourteen Japanese EFL textbooks using the new index and *Flesch-Kincaid Grade level* (hereafter, *Flesch-Kincaid*), and quantitatively compare the results.

## 2 Two analyses

### 2.1 Analysis 1

In Analysis 1, coefficients of correlation were computed among the three measurements by *Ozasa-Fukui*, *Flesch-Kincaid*, and *Flesch Reading Ease*, and professional readability judgment of 84 English sentences extracted from five Japanese EFL textbooks, *Columbus 21 English Course*, Vols. 1, 2 and 3, junior high school English textbooks, and

<sup>1</sup>The present study is funded by the Grants-in-Aid for Scientific Research ((c)1, No. 19520535) of the Japan Society for the Promotion of Science. The authors wish to express their heartfelt gratitude for the financial support to the Japan Society for the Promotion of Science.

*Voyager English Course*, vols. 1 and 2, senior high school English textbooks. The readability evaluator was the present first author (Ozasa), a specialist in Japanese TEFL, current author of junior high school English textbooks and former author of senior high school English textbooks.

Table 1: Correlation among Ozasa's Judgment, *Ozasa-Fukui YL*, *Flesch RE* and *Flesch-Kincaid GL*

	Ozasa's	O-F YL	F RE	F-K GL
Ozasa's	1	0.9041	-0.4396	0.7137
O-F YL	0.9041	1	-0.5279	0.7743
F RE	-0.4396	-0.5279	1	-0.916
F-K GL	0.7137	0.7743	-0.916	1

Table 1 shows the correlations with Ozasa's judgment and *Ozasa-Fukui*, *Flesch Reading Ease* and *Flesch-Kincaid*. As is clear in the table, the coefficient of correlation between Ozasa's criterion and *Ozasa-Fukui* was the highest among the three correlations; the coefficients were 0.9041 between Ozasa's criterion and *Ozasa-Fukui*, -0.4391 between Ozasa's criterion and *Flesch Reading Ease*, and 0.7137 between Ozasa's criterion and *Flesch-Kincaid*.

These results clearly indicate that *Ozasa-Fukui* was the most powerful in differentiating the readability of selected 84 EFL sentences. It is noteworthy that the coefficient between Ozasa's criterion and *Flesch-Kincaid* was also remarkably high, much higher than that of *Flesch Reading Ease*, indicating a high validity of the index in measuring the readability of textbook English.

## 2.2 Analysis 2

Using this newly developed index (*Ozasa-Fukui*), and the widely used *Flesch-Kincaid* metric, the fourteen sets of textbooks were measured and analyzed. In this analysis the textbook corpora were 'cleaned' by deleting irrelevant parts of the texts - conversation, poems, etc. - and by tagging irrelevant parts of the text - titles, headlines, etc. - so that the system would not include them in its measurements.

The fourteen sets of English textbooks, each of which consisted of five textbooks, totaling 70 books, were as follows. (The abbreviations in brackets are the textbook identifiers used in the present paper.)

- 1) *Sanders' Union Readers* (C. W. Sanders, 1861 - 63) [*Union*]
- 2) *New National Readers* (C. J. Barnes, 1883-84) [*National*]
- 3) *English Readers: The High School Series* (W. Denning, 1887 [*Denning*])
- 4) *Seisoku Mombusho Eigo Tokuhon* (Education

Ministry English Readers) (Education Ministry, 1889) [*Seisoku*]

- 5) *Standard Choice Readers* (Shobido, 1902) [*Choice*]
- 6) *Kanda's New Series of English Readers (Revised Edition)* (N. Kanda, 1903) [*Kanda NSER*]
- 7) *The Globe Readers* (Y. Okakura, 1907 [*Globe*])
- 8) *New English Drill Books* (K. Kumamoto, 1907 [*NE Drill*])
- 9) *Girls' Taisho New Readers* (U. Tsuda & K. Kumamoto, 1916) [*Taisho*]
- 10) *The Standard English Readers* (H. E. Palmer, 1926-27) [*Standard P*]
- 11) *The Standard English Readers* (T. Takehara, 1932 [*Standard T*])
- 12) *Girls' Pacific Readers* (T. Sawamura, 1939) [*Pacific*]
- 13) *New Jack and Betty: English Step by Step* (1,2,3) (K. Hagiwara, et al, 1952, *New High School English* (1, 2) K. Hagiwara, et al, 1952) [*Jack & Betty*]
- 14) *Sunshine English Course* (1,2,3) (T. Shimaoka, et al, 1996) , *Sunshine English Course* (1,2) (S. Tsuchiya, et al, 1997) [*Sunshine*]

Table 2 shows the results of the readability analysis of the fourteen Japanese EFL textbooks, measured with *Ozasa-Fukui*, while Table 3 lists the results of the readability analysis of the same corpora, measured with *Flesch-Kincaid*.

Table 2: *Ozasa-Fukui Year Level* (Books 1-5)

	Book 1	Book 2	Book 3	Book 4	Book 5
Union	4.32	6.96	6.46	6.87	7.04
National	3.45	4.64	5.38	6.37	6.80
Denning	6.37	6.50	6.45	6.66	6.95
Seisoku	3.13	4.13	6.33	6.93	7.32
Choice	3.37	4.83	5.85	6.27	6.73
Kanda NSER	1.64	3.00	5.02	6.87	6.46
Globe	3.30	5.01	6.31	6.06	6.18
NE Drill	2.10	4.11	4.97	6.13	6.42
Taisho	2.91	3.45	4.68	5.58	6.48
Standard P	3.05	6.40	6.53	5.82	6.70
Standard T	3.08	5.10	5.99	6.32	6.52
Pacific	2.35	4.00	4.64	5.41	6.77
Jack & Betty	1.90	2.90	3.91	5.12	5.53
Sunshine	1.38	3.09	3.66	4.91	5.23

These tables seem to indicate that the measurements of both indices were generally effective in differentiating the readability of these textbook texts. In both of the measurements, the readability values increased progressively from Book 1 to Book 5 in eight textbooks, *National*, *Seisoku*, *Choice*, *NE Drill*, *Taisho*, *Standard T*, *Pacific* and *Sunshine*.

However, it should be noted that these results cannot be accurately compared since these two indices use their own scale, different from each other. In order to overcome this deficiency, the values of the two measurements were statistically transformed into a standardized distribution of  $M=0$  and  $SD=1$ .

Table 3: Flesch-Kincaid Grade Level (Books 1-5)

	Book 1	Book 2	Book 3	Book 4	Book 5
Union	2.75	9.39	7.55	8.25	8.85
National	1.11	2.98	5.31	7.64	8.81
Dening	7.60	8.35	8.35	8.34	9.55
Seisoku	1.46	3.52	6.74	8.79	9.43
Choice	0.99	3.64	5.83	7.68	8.62
Kanda NSER	0.55	1.88	4.27	9.05	7.70
Globe	2.17	3.81	6.92	6.67	7.46
NE Drill	1.30	3.50	3.76	5.62	8.15
Taisho	1.02	1.62	3.79	5.53	7.84
Standard P	2.56	7.74	7.97	7.12	8.95
Standard T	1.58	3.94	6.06	7.46	7.88
Pacific	1.66	3.39	4.09	5.70	8.47
Jack & Betty	1.37	2.37	5.21	6.72	5.06
Sunshine	2.79	3.25	4.10	5.20	5.80

Table 4: Readability Measurement of *Ozasa-Fukui Year Level*, Standardized

	Book 1	Book 2	Book 3	Book 4	Book 5
Union	-0.51	1.16	0.84	1.10	1.21
National	-1.06	-0.31	0.16	0.78	1.06
Dening	0.78	0.87	0.83	0.97	1.15
Seisoku	-1.26	-0.63	0.76	1.14	1.38
Choice	-1.11	-0.19	0.46	0.72	1.01
Kanda NS	-2.21	-1.35	-0.07	1.10	0.84
Globe	-1.16	-0.08	0.75	0.59	0.66
NE Drill	-1.92	-0.65	-0.10	0.63	0.82
Taisho	-1.40	-1.06	-0.28	0.28	0.85
Standard F	-1.32	0.80	0.89	0.44	0.99
Standard T	-1.30	-0.02	0.54	0.75	0.88
Pacific	-1.76	-0.71	-0.31	0.18	1.04
Jack & Bet	-2.04	-1.41	-0.77	-0.01	0.25
Sunshine	-2.37	-1.29	-0.93	-0.14	0.06

Table 5: Readability Measurement of *Flesch-Kincaid Grade Level*, Standardized

	Book 1	Book 2	Book 3	Book 4	Book 5
Union	-0.99	1.46	0.78	1.04	1.26
National	-1.60	-0.91	-0.05	0.81	1.25
Dening	0.80	1.08	1.08	1.07	1.52
Seisoku	-1.47	-0.71	0.48	1.24	1.47
Choice	-1.64	-0.66	0.15	0.83	1.18
Kanda NSI	-1.80	-1.31	-0.43	1.33	0.84
Globe	-1.21	-0.60	0.55	0.46	0.75
NE Drill	-1.53	-0.72	-0.62	0.07	1.00
Taisho	-1.63	-1.41	-0.61	0.03	0.89
Standard P	-1.06	0.85	0.94	0.62	1.30
Standard T	-1.42	-0.55	0.23	0.75	0.90
Pacific	-1.39	-0.76	-0.50	0.10	1.12
Jack & Bet	-1.50	-1.13	-0.08	0.47	-0.14
Sunshine	-0.98	-0.81	-0.49	-0.09	0.13

Table 4, Table 5, Figure 2 and Figure 3 show the standardized readability measurements of the two indices. Since these tables and figures use the same standardized scale, it is possible to compare the measurements of the two indices. Intuitively, it seems that with Books 1 and 2, the *Ozasa-Fukui* measurement is more powerful in differentiation and, in contrast, with higher books, *Flesch-Kincaid* is more powerful than *Ozasa-Fukui*. To get a clearer picture of the features, however, it is necessary to compare these data in detail methodically for each book.

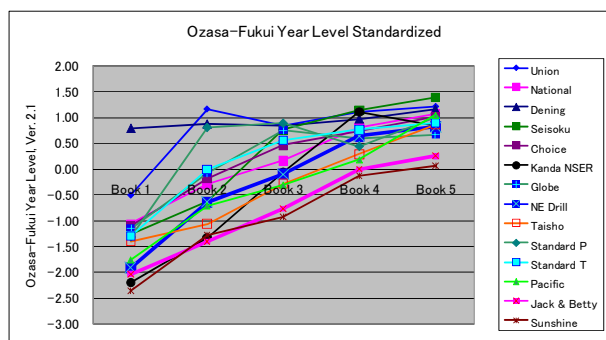


Figure 2: Readability Measurement of *Ozasa-Fukui Year Level*, Standardized

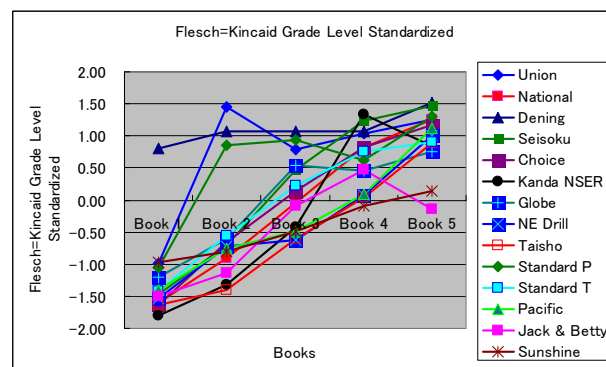


Figure 3: Readability Measurement of *Flesch-Kincaid Grade Level*, Standardized

Table 6 and Table 7 show the standardized readability measurements of the two readability indices of the Book 1 texts of the fourteen sets in question. In order to compare the differentiating power of the indices, the fourteen textbooks were arranged in declining order and differences were computed between the neighboring textbooks. For example, in Table 6, the -0.16 in the cell of the lowest row and the third column from the left was computed by subtracting the value of *Kanda NSER* (-2.21) from the value of *Sunshine* (-2.37). In this analysis it was assumed that the more the number of greater differences there were, the more powerful was the differentiation power of the measurement.

As is clear in these tables, with *Ozasa-Fukui*, eight differences were greater than 0.09, while with *Flesch-Kincaid* seven were greater, which suggests that *Ozasa-Fukui* might be slightly more powerful than *Flesch-Kincaid*, with the Book 1 measurement.

The sum total of the differences for each index was -3.16 and -2.60 respectively, which also implies that *Ozasa-Fukui* is more powerful in differentiation than *Flesch-Kincaid*.

It is also interesting to note that the measurement values of *National* and *Sunshine* for the two indices were remarkably different and this should be examined from a qualitative perspective.

Table 6: Book 1, *Ozasa-Fukui Year level*,

Standardized

O-F	Book 1	
Denning	0.78	
Union	-0.51	-1.30
National	-1.06	-0.55
Choice	-1.11	-0.05
Globe	-1.16	-0.04
Seisoku	-1.26	-0.11
Standard T	-1.30	-0.03
Standard P	-1.32	-0.02
Taisho	-1.40	-0.09
Pacific	-1.76	-0.35
NE Drill	-1.92	-0.16
Jack & Bet	-2.04	-0.13
Kanda NSE	-2.21	-0.16
Sunshine	-2.37	-0.16

Table 7: Book 1, *Flesch-Kincaid Grade Level*, Standardized

F-K	Book 1	
Denning	0.80	
Sunshine	-0.98	-1.78
Union	-0.99	-0.01
Standard P	-1.06	-0.07
Globe	-1.21	-0.14
Pacific	-1.39	-0.19
Standard T	-1.42	-0.03
Seisoku	-1.47	-0.04
Jack & Bet	-1.50	-0.03
NE Drill	-1.53	-0.03
National	-1.60	-0.07
Taisho	-1.63	-0.03
Choice	-1.64	-0.01
Kanda NSE	-1.80	-0.16

Table 8 and Table 9 show the standardized readability data of the Book 2 texts of the fourteen sets measured with the two readability indices. As is clear from the tables, seven differences were greater than 0.09 both with *Ozasa-Fukui* and with *Flesch-Kincaid*, which suggests that the two indices were equal in their differentiation for the Book 2 measurement.

The sum total of the differences for each index was -2.57 and -2.87 respectively, which suggests that *Flesch-Kincaid* is slightly more powerful in differentiation than *Ozasa-Fukui*.

It is also to be noted that the measurement values of *National* were remarkably different between the measurements of the two indices, which should also be carefully examined from a qualitative perspective.

Table 10 and Table 11 show the standardized readability measurements of the Book 3 texts of the fourteen sets measured using the two readability indices in question. As is clear in the tables, with *Ozasa-Fukui*, six differences were greater than 0.09, and with *Flesch-Kincaid* seven differences were also greater than 0.09, which suggests that *Flesch-Kincaid* is slightly more powerful in

differentiation than *Ozasa-Fukui*, for the Book 3 measurement.

Table 8: Book 2, *Ozasa-Fukui Yea Level*, Standardized

O-F	Book 2	
Union	1.16	
Denning	0.87	-0.29
Standard P	0.80	-0.06
Standard T	-0.02	-0.82
Globe	-0.08	-0.06
Choice	-0.19	-0.11
National	-0.31	-0.12
Seisoku	-0.63	-0.32
NE Drill	-0.65	-0.01
Pacific	-0.71	-0.07
Taisho	-1.06	-0.35
Sunshine	-1.29	-0.23
Kanda NSE	-1.35	-0.06
Jack & Bet	-1.41	-0.06

Table 9: Book 2, *Flesch-Kincaid Grade Level*, Standardized

F-K	Book 2	
Union	1.46	
Denning	1.08	-0.38
Standard P	0.85	-0.23
Standard T	-0.55	-1.40
Globe	-0.60	-0.05
Choice	-0.66	-0.06
Seisoku	-0.71	-0.04
NE Drill	-0.72	-0.01
Pacific	-0.76	-0.04
Sunshine	-0.81	-0.05
National	-0.91	-0.10
Jack & Bet	-1.13	-0.23
Kanda NSE	-1.31	-0.18
Taisho	-1.41	-0.10

Table 10: Book 3, *Ozasa-Fukui Year Level*, Standardized

O-F	Book 3	
Standard P	0.89	
Union	0.84	-0.04
Denning	0.83	-0.01
Seisoku	0.76	-0.08
Globe	0.75	-0.01
Standard T	0.54	-0.20
Choice	0.46	-0.09
National	0.16	-0.30
Kanda NSE	-0.07	-0.23
NE Drill	-0.10	-0.03
Taisho	-0.28	-0.18
Pacific	-0.31	-0.03
Jack & Bet	-0.77	-0.46
Sunshine	-0.93	-0.16

The sum total of the differences for each index was -1.81 and -1.69 respectively, which suggests that *Ozasa-Fukui* is slightly more powerful in differentiation than *Flesch-Kincaid*. Here the

conclusions obtained through the two kinds of analyses prove to be mutually contradicting as to differentiating power.

It is also to be noted that the measurement values of *Jack & Betty* was remarkably different between the two indices – a further point that should be examined from a qualitative perspective.

Table 11: Book 3, *Flesch-Kincaid Grade Level*, Standardized

F-K	Book 3	
Denning	1.08	
Standard P	0.94	-0.14
Union	0.78	-0.16
Globe	0.55	-0.23
Seisoku	0.48	-0.07
Standard T	0.23	-0.25
Choice	0.15	-0.08
National	-0.05	-0.19
Jack & Bet	-0.08	-0.04
Kanda NSF	-0.43	-0.35
Sunshine	-0.49	-0.06
Pacific	-0.50	0.00
Taisho	-0.61	-0.11
NE Drill	-0.62	-0.01

Table 12: Book 4, *Ozasa-Fukui Year Level*, Standardized

O-F	Book 4	
Seisoku	1.14	
Union	1.10	-0.04
Kanda NSF	1.10	0.00
Denning	0.97	-0.13
National	0.78	-0.18
Standard T	0.75	-0.03
Choice	0.72	-0.03
NE Drill	0.63	-0.09
Globe	0.59	-0.04
Standard P	0.44	-0.15
Taisho	0.28	-0.15
Pacific	0.18	-0.11
Jack & Bet	-0.01	-0.18
Sunshine	-0.14	-0.13

Table 12 and Table 13 show the standardized readability measurements of the Book 4 texts of the fourteen sets measured with the two readability indices. As is clear from these tables, seven differences were greater than 0.09 both with *Ozasa-Fukui* and with *Flesch-Kincaid*, which suggests that *Ozasa-Fukui* is as powerful in differentiation as *Flesch-Kincaid* with the Book 4 measurement.

However, the sum total of the differences for each index was -1.28 and -1.42 respectively, which suggests that *Flesch-Kincaid* is slightly more powerful in differentiation than *Ozasa-Fukui*.

It is also to be noted that the measurement values of *NE Drill* and *Jack & Betty* were remarkably different between the two indices– a

further point that should be examined from a qualitative perspective.

Table 13: Book 4, *Flesch-Kincaid Grade Level*, Standardized

F-K	Book 4	
Kanda NSF	1.33	
Seisoku	1.24	-0.10
Denning	1.07	-0.17
Union	1.04	-0.03
Choice	0.83	-0.21
National	0.81	-0.01
Standard T	0.75	-0.07
Standard P	0.62	-0.13
Jack & Bet	0.47	-0.15
Globe	0.46	-0.02
Pacific	0.10	-0.36
NE Drill	0.07	-0.03
Taisho	0.03	-0.03
Sunshine	-0.09	-0.12

Table 14: Book 5, *Ozasa-Fukui Year Level*, Standardized

O-F	Book 5	
Seisoku	1.38	
Union	1.21	-0.18
Denning	1.15	-0.06
National	1.06	-0.09
Pacific	1.04	-0.02
Choice	1.01	-0.03
Standard P	0.99	-0.02
Standard T	0.88	-0.11
Taisho	0.85	-0.03
Kanda NSF	0.84	-0.01
NE Drill	0.82	-0.03
Globe	0.66	-0.15
Jack & Bet	0.25	-0.41
Sunshine	0.06	-0.19

Finally, Table 14 and Table 15 show the standardized readability of the Book 5 texts of the fourteen sets measured with the two indices. As is clear from these tables, with *Ozasa-Fukui*, seven differences were greater than 0.09, and with *Flesch-Kincaid* five differences were greater than 0.09. This suggests that *Ozasa-Fukui* is more powerful in differentiation than *Flesch-Kincaid* for the Book 5 measurement.

However, the sum total of the differences for each index was -1.32 and -1.66 respectively, which suggests that *Flesch-Kincaid* is more powerful in differentiation than *Ozasa-Fukui*. This means that the two kinds of obtained data are contradictory to each other in terms of differentiation.

It is also to be noted that the readability values of *Standard P* were remarkably different between the measurements of the two indices, which should be meticulously examined from a qualitative perspective.

Table 15: Book 5, *Flesch-Kincaid Grade Level*, Standardized

F-K	Book 5	
Denning	1.52	
Seisoku	1.47	-0.04
Standard P	1.30	-0.18
Union	1.26	-0.04
National	1.25	-0.01
Choice	1.18	-0.07
Pacific	1.12	-0.06
NE Drill	1.00	-0.12
Standard T	0.90	-0.10
Taisho	0.89	-0.01
Kanda NSE	0.84	-0.05
Globe	0.75	-0.09
Sunshine	0.13	-0.61
Jack & Bet	-0.14	-0.27

### 3 Conclusion and remaining issues

In the present study, the coefficients of correlations were computed between Ozasa's readability judgment (criterion) and the measurements of *Ozasa-Fukui*, *Flesch-Kincaid* and *Flesh Reading Ease*, and the results were compared among the four coefficients. The results revealed that *Ozasa-Fukui* was the most powerful in differentiating the readability of the selected 84 EFL sentences. It also emerged that the coefficient between Ozasa's criterion and *Flesch-Kincaid* was remarkably high, much higher than that between the criterion and *Flesh Reading Ease*, suggesting a high validity of *Flesch-Kincaid* in measuring the readability of Japanese EFL textbook English.

Table 16: Summary of Analysis 2

Book	No of Diff	Sum of Diff
1	OF	OF
2	=	FK
3	FK	OF
4	=	FK
5	OF	FK

Further, as is clear in Table 16, Analysis 2 revealed that *Ozasa-Fukui* was more powerful in differentiating Book 1 texts, and that *Flesch-Kincaid* was generally powerful in differentiating the texts of higher years. In general, however, we could not quantitatively distinguish the two indices.

The analyses carried out in the present study were two-fold, a criterion based analysis and criterion-less comparison of the two types of measurements. The criterion-based analysis succeeded in producing a clear-cut conclusion as to the relative efficiency of the three indices in question. In contrast, the criterion-less comparison could not produce a persuasive or consistent conclusion as to the relative efficiency of the two indices analyzed.

In the next study, only one or two sets of

Japanese EFL textbooks will be used as data, instead of fourteen sets of textbooks as in the present analysis, and will be closely analyzed both from quantitative and qualitative perspectives, with reference to the professional judgment of EFL teachers.

The time has come to provide our readability index for public use so that it may be freely used and evaluated by users around the world. Feedback obtained from these users could greatly enhance our new readability index.

### References

- Anagnostou, N. K. & Weir, G. R. S. (2006). From corpus-based collocation frequencies to readability measure. ICT in the Analysis, Teaching and Learning of Languages (ICTATLL Workshop), 21-22 August, 2006, 33-46.
- Ozasa, T., Fukui, M. & Hosokawa, M. (2008). Development of a readability index attuned for Japanese EFL texts. Proceedings of 4<sup>th</sup> International Workshop in ICT in the Analysis, Teaching and Learning of Languages, September, 2008, University of Sri Lanka, 60-66.
- Ozasa, T., Weir, G. R. S. & Fukui, M. (2008). Toward a readability index for Japanese learners of EFL. Conference Program of the 13<sup>th</sup> Conference of Pan-Pacific Association of Applied Linguistics, 20-22 August, 2008, University of Hawaii, Manoa, U. S. A., 59-62.
- Ozasa, T., Weir, G. R. S. & Fukui, M. (2007). Measuring readability for Japanese learners of English. Proceedings of the 12<sup>th</sup> Conference of Pan-Pacific Association of Applied Linguistics, 122-125.
- Weir, G. R. S. & Ritchie, C., (2006). Estimating readability with the Strathclyde readability measure. ICT in the Analysis, Teaching and Learning of Languages (ICTATLL Workshop), 21-22 August, 2006, 25-32.

# Verb form usage in Japanese EFL Texts

George R S Weir<sup>1</sup> and Toshiaki Ozasa<sup>2</sup>

<sup>1</sup>Department of Computer and Information Sciences, University of Strathclyde

<sup>2</sup>Faculty of Business Administration, Fukuyama Heisei University

george.weir@cis.strath.ac.uk, tozasa@heisei-u.ac.jp

## Abstract

In the present paper, we report a test of our assumption that three sets of historical EFL textbooks exhibit priority of grammatical over vocabulary consideration. This test is performed by profiling the use of verb forms in the three textbook series. Evidence of a managed application of verb forms across the volumes of the textbook series reflects a focus upon grammatical control in textbook design. Coupled to our earlier results on vocabulary profiles, this corroborates our view that the considered textbooks exhibit an emphasis of grammar over vocabulary.

## Keywords

Japanese EFL textbooks, historical linguistics, textual analysis, grammar control

## Introduction

Japan has a long history of EFL teaching and has witnessed a variety of approaches to EFL textbook design (Ozasa & Erikawa, 2004). From an historical perspective, we seek to determine whether there are discernible commonalities in the linguistic characteristics displayed by successive EFL textbooks. This diachronic approach builds upon earlier work at the University of Hiroshima and the development of the Hiroshima Corpus of historical EFL textbooks (e.g., Weir & Ozasa, 2007).

The present study contrasts three EFL textbook series used in Japan at three different historical periods. The first series, Barnes' New National Readers, was published in 1883-84 and is taken to represent the 'early' period of ESL teaching in Japan. The second textbook series, Okakura, Yoshisaburo, The Globe Readers, was published in 1907 and is taken to represent the 'middle' period of ESL teaching in Japan. The third textbook series, Jack and Betty: English Step by Step, was published in 1948 and represents 'recent' ESL teaching in Japan. In the following, we refer to these textbooks as 'National', 'Globe' and 'JandB', respectively.

Previously, we conducted a frequency analysis on the vocabulary in these textbooks (Weir and Ozasa, 2008). This used the Posit textual analysis toolset (Weir, 2007) in order to shed light upon the frequency of use for individual words and multiword sequences across the specified textbook series. Based upon the high degree of hapax legomena (as summarized in Table 1), we noted that only a small proportion of words were considered important enough to merit frequent use.

Table 1: Percentages of hapax legomena

	National	Globe	JandB
1-gram	44.80	50.48	48.38
2-grams	76.68	79.84	76.49
3-grams	92.15	94.23	91.29
4-grams	97.99	98.76	96.62

One might assume that careful textbook design would address vocabulary coverage, as well as grammatical considerations. Furthermore, one might expect more frequent use of those words that were considered more important than others. Yet, the figures from Table 1 show that almost half of the individual words in each of the textbook series are used only once. This led us to propose that the word frequency profiles may result from priority of consideration to grammatical construction rather than vocabulary.

The present paper reports on our efforts to substantiate this view of the considered textbook series.

## 1 Approach

In seeking support for our view that emphasis upon grammar control may underlie the content of the considered textbooks, we sought to analyse verb form usage across each textbook series. Specifically, we look for telling variations in frequency of verb usage as evidence of a managed application of verb forms. Such evidence reflects a focus upon grammatical control in textbook design.

For each book in each textbook series we analyse the frequency of specific verb instances (verb tokens), e.g. how many times ‘ran’ appears, and verb types (i.e., how many different verbs occur). These results are also classified by verb form (part of speech). Seven verb forms are considered: base, past, past participle, gerund, present 3rd, present not 3rd and modal auxiliary.

Table 2: Summary data for Globe Book 1

Input filename	globe-1.txt
Total words (tokens)	1598
Total unique words (types)	452
Type/Token Ratio	3.53
Number of sentences	250
Average sentence length	6.39
Number of characters	8183
Average word length	5.12

For each book in a textbook series, the Posit textual analysis toolset provides extensive data including summary details of the input text as well as totals for part of speech (POS) form tokens and POS form types. Tables 2 and 3 show such summary data for Book 1 of the Globe series.

Table 3: Summary POS data for Globe Book 1

	POS Types	POS Tokens
Nouns	184	429
Verbs	105	298
Adjectives	63	110
Adverbs	37	107
Prepositions	22	119
Determiners	13	142
Personal pronouns	12	210
Possessive pronouns	6	38
Interjections	2	4
Particles	0	0

The Posit tools also provide more detailed information on each part of speech, such as the seven verb forms considered in this analysis. Table 4 shows the totals for types of each verb form in Book 1 of the Globe series.

Table 4: Verb form type totals for Globe Book 1

base_form	34
gerund_form	2
past_form	20
past_participle_form	7
present_3rd_form	15
present_not3rd_form	23
modal_aux	4
total	105

The most detailed level of data analysis provided by the Posit system indicates the number of tokens for each verb form. A sample of this detail for Book 1

of the Globe series is shown in Figure 1 (below).

Base	Present 3 <sup>rd</sup>	Gerund	Past	Present not 3 <sup>rd</sup>	Modal Aux
10 be	54 is	2 going	13 was	19 are	11 will
4 do	4 has	1 turning	5 said	14 have	10 can
3 send	2 loses		4 did	9 am	4 shall
3 let	2 comes		3 were	4 see	2 must
2 write	1 writes		3 got	3 love	
2 thank	1 tells		3 cried	3 do	
2 tell	1 sings		2 let	2 wind	
2 see	1 plays		2 gave	2 want	
2 look	1 kills		2 called	2 play	
2 know	1 keeps		1 went	2 look	
2 hear	1 goes		1 wanted	2 know	
2 come	1 eats		1 tried	2 go	
1 try	1 does		1 thought	1 tell	
...	1 contains		1 stood	1 send	
	1 consists		1 played	...	

Figure 1: Detailed token data for Globe Book 1

Each of the textbook series comprises five books and on the basis of such data for each book in each of the three textbook series, we look for evidence of grammar control. Specifically, is there indication in the usage pattern of verbs across books in a textbook series to suggest verb management?

## 2 Results

### 2.1 Overall verb density

The first step is to consider the overall density of verbs across the books of each textbook. This is measured by determining a proportion of all verb forms against all word forms. For National, this reveals the distribution shown in Figure 2, below.

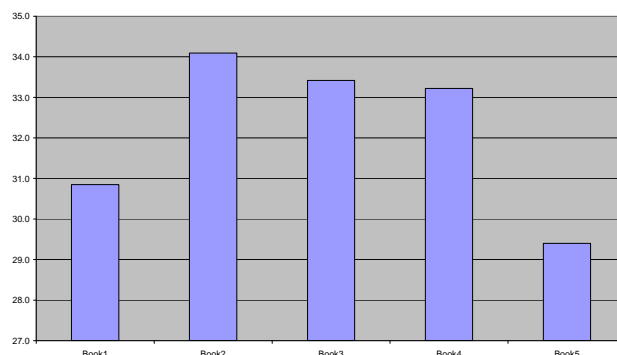


Figure 2: Verb distribution for National

The x-axis considers each book in turn while the y-axis (on a scale of 0% to 35%) measures the percentage of all word forms accounted for by the verb forms in the book. In this case, we can note that the largest proportion of verbs appears in Book 2 (34.1%), while the smallest proportion appears in Book 5 (29.4%). Notably, we can determine that the maximum variation in verb proportion across all five books is only 5%.

Considering similar measures for Globe, we get the results shown in Figure 3 (below). This indicates that the maximum density of verbs is 29.9% (Book 2) and the minimum is 23.2% (Book 1). This exhibits a variation of 6.9%.

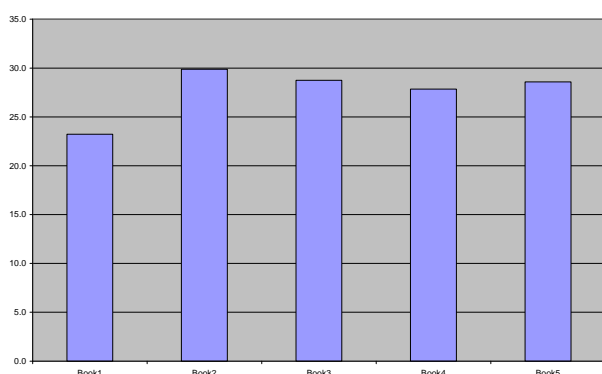


Figure 3: Verb distribution for Globe

The overall verb density for JandB is illustrated in Figure 4 (below).

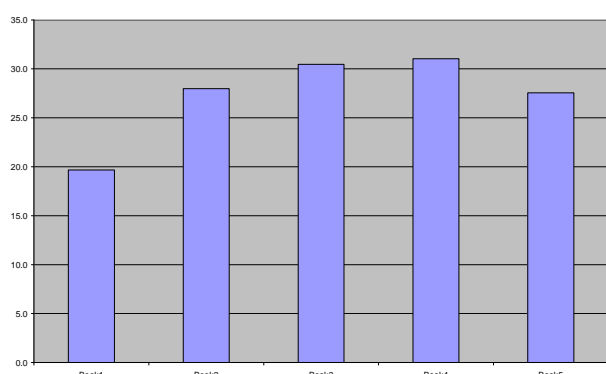


Figure 4: Verb distribution for JandB

For the JandB textbook series, the maximum verb density is 31% (Book 4) and the minimum density is 19.7% (Book 1), giving a variation of 10.3%.

The point of these indications for verb density is to emphasise that the density is fairly consistent across the three textbook series, with an overall value for each series of ~30%. This indicates that the verb density is not influenced by variations in the size of each textbook (measured by number of words).

Figure 5 (below) indicates the progressive increase in word tokens across the books in each series. (The y-axis scale is 10,000 words per segment.) All three textbook series, have more words in each successive book (as the books grow 'more difficult'), yet this has little influence on the verb density.

With this in mind, we move to consider the proportion of all verb types in each book against verb types for the seven verb forms (base, past, past participle, gerund, present 3rd, present not 3rd and modal auxiliary). By noting the changing proportions of different verb forms against the total number of verbs in each book we can map any changes in emphasis (or de-emphasis) toward specific verb forms as we progress through books in any textbook series.

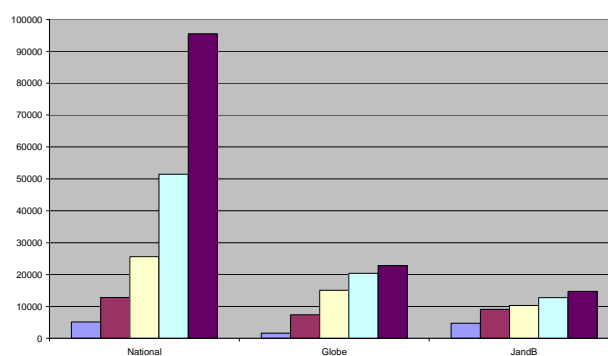


Figure 5: Tokens per book for each textbook series

## 2.2 Verb form usage

In order to gauge any changes in emphasis (or de-emphasis) toward specific verb forms, we graph the frequency of occurrence of our seven verb forms in each of the books (in each series) as a proportion of the total verb forms in each book. (We have used verb types, although the results are similar in nature using tokens instead).

The proportion of verb forms in each book of National is given in Figure 6. The x-axis clusters the results for each of the five books in the series against our seven verb forms. The y-axis indicates the percentage for each verb form against all verb forms, with a scale of 0% to 40%.

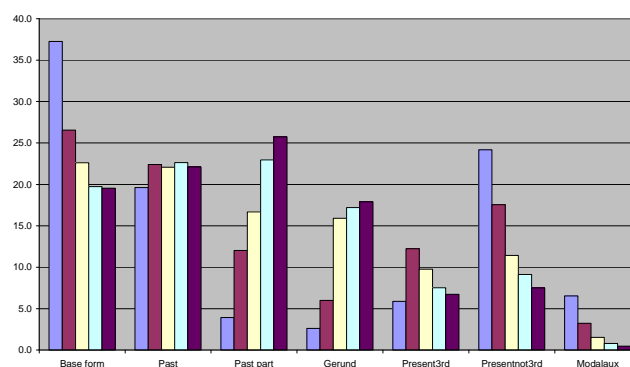


Figure 6: Verb form usage in National

For National, several trends are distinct in this projection. Firstly, some verb forms progressively diminish in their proportion from Book 1 through to Book 5 in the series. This is true for the base form, present 3<sup>rd</sup> (with the exception of Book 1), present not 3<sup>rd</sup>, and modal auxiliary. Secondly, one verb form (past) remains fairly constant across the five books. Thirdly, some verb forms progressively increase in their proportion across the five books. This is true for the past participle and gerund forms.

The proportion of verb forms in each book of Globe is given in Figure 7. As with National, several trends are distinct in the projection for Globe. Firstly, some verb forms progressively diminish in their proportion from Book 1 through to Book 5 in the series. This is true for the base form, present 3<sup>rd</sup> (with the exception of Book 5), present

not 3<sup>rd</sup>, and modal auxiliary. Secondly, some verb forms have no clear overall increase or decrease. This is true for past and gerund. Thirdly, some verb forms progressively increase in their proportion across the five books. This is true for the past participle and gerund forms (although gerund has a spike in Book 2).

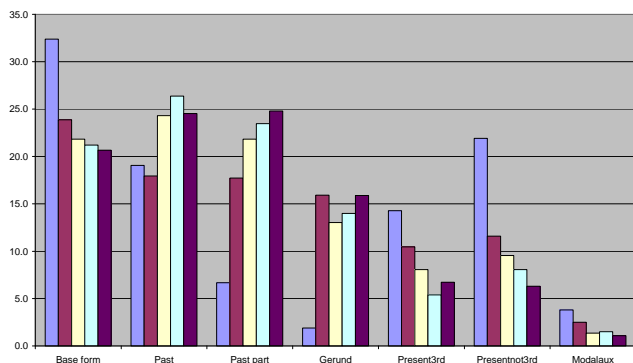


Figure 7: Verb form usage in Globe

The proportion of verb forms in each book of JandB is shown in Figure 8. As with the other two textbook series, several trends are distinct for JandB. Firstly, some verb forms progressively diminish in their proportion from Book 1 through to Book 5 in the series. This is generally true for the base form, present 3<sup>rd</sup>, present not 3<sup>rd</sup>, and modal auxiliary. Secondly, two verb forms (past and gerund) show no distinct trend across the five books. Thirdly, the past participle verb form progressively increases in proportion across the five books.

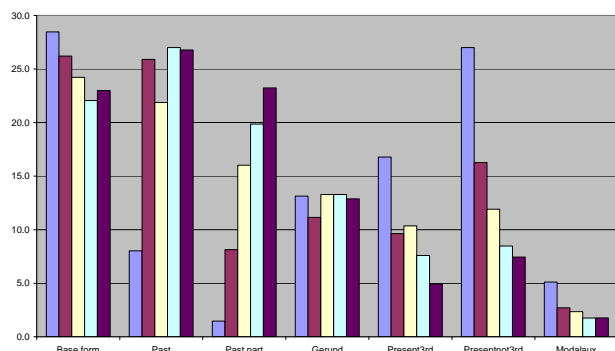


Figure 8: Verb form usage in JandB

### 3 Conclusions

Across each of our three textbook series we can recognise trends in verb usage. For National, there is a progressive reduction in the proportion of base form verbs, present not 3<sup>rd</sup>, and modal auxiliaries from Book 1 through to Book 5. At the same time, there is a progressive increase in the proportions of past participle and gerund forms. The present 3<sup>rd</sup> form also shows a decreasing pattern (with the exception of the change between Book 1 and Book 2) while the past form is fairly static in proportion across the five books in the series.

For Globe, there is progressive reduction in proportion of base form verbs, present 3<sup>rd</sup>, present not 3<sup>rd</sup>, and modal auxiliaries from Book 1 to Book 5. Although the results are less 'smooth' for this series, there are progressive increases in the proportion of past, past participle and gerund forms.

For Jack and Betty, there are clear reductions in the proportion of base form verbs, present 3<sup>rd</sup>, present not 3<sup>rd</sup> and modal auxiliaries from Book 1 through to Book 5. The past form and gerund are fairly static across all books in the series. At the same time, there is a progressive increase in the proportions of past participles.

Worthy of additional note is the fact that the trends are largely similar across the textbooks. This is an interesting result given the different historical periods represented by these EFL materials. Furthermore, the results may be indicative that *lower proportions* of base form, present and modal auxiliaries and *higher proportions* of past participles (and gerunds) are markers of EFL texts in general.

### 4 Summary

We have considered the changing proportion of seven verb forms across the books of three historical Japanese EFL textbook series and found evidence of a managed application of verb forms across each of these textbooks. There are discernible trends in verb usage and these trends are strong evidence of grammar control across books in each textbook series.

In addition, since each of the historical textbook series exhibits similar trends in verb usage, across sixty-five years, we take this as indication of an enduring view on the contribution made by differing verb forms to the complexity of EFL learning materials.

### References

- Ozasa, T. & Erikawa, H. (2004). *Eigo Kyokasho no Rekishiteki Kennkyu* [A Historical Study of English Textbooks]. Tokyo: Jiyusha.
- Weir, G.R.S. (2007). The Posit Text Profiling Toolset. *Proceedings of the 12th Conference of Pan-Pacific Association of Applied Linguistics*, Thailand.
- Weir, G.R.S. & Ozasa, T. (2007). Estimating naturalness in Japanese English textbooks. *Proceedings of the 12th Conference of Pan-Pacific Association of Applied Linguistics*, Thailand.
- Weir, G.R.S. & Ozasa, T. (2008). Multiword vocabulary in Japanese ESL Texts. *Proceedings of the 13th Conference of Pan-Pacific Association of Applied Linguistics*, Hawai'i.

# Conjunctive Adverbials in English Academic Writing by Chinese Speakers: A Corpus Approach

Tung-Yu Kao<sup>1</sup> and Li-Mei Chen<sup>2</sup>

Department of Foreign Languages and Literature, National Cheng Kung University

<sup>1</sup>dodofishletter@hotmail.com, <sup>2</sup>leemay@mail.ncku.edu.tw

## Abstract

The present study aims to investigate the distribution of textual relations performed by CAs (conjunctive adverbials) across different genres based on the corpus compiled in the AWTa project. In this study, forty-five pieces of articles, written by 5 college students, with each student contributing 9 articles of different genres in three years, were selected and analyzed in terms of the seven textual relations and six misuse patterns of CAs. The result shows that there exists a preferred distribution of textual relations signaled by CAs and free from genre time influences. This suggests that the ability to think logically for composing an article is in relation to cognition development, and that once the ability reaches its maturation it would not be subject to other factors. These obtained results can contribute to self-learning resources and teaching material design. For the former, the annotated data can provide authentic data for learners' reference while, for the latter, the findings regarding CAs can offer a new perspective for teachers to teach writing and to arrange their teaching materials other than rhetoric approach.

## Keywords

Conjunctive adverbials, English academic writing, Chinese speaker, corpus approach

## Introduction

The dominance of English makes English writing a crucial platform on which people from different language backgrounds can communicate. From a research point of view, the ability to publish academic articles in English has become an indispensable skill for researchers in order to gain more circulation. Given the trend, it becomes mandatory for English majors to receive academic writing training in college.

Due to the demand to teach English academic writing and cultivate this ability, many textbooks on writing are published, but it is found that most textbooks on writing concentrate on introducing various genres to learners, presupposing and

contending conjunctive adverbials (hereafter, CAs) performing certain textual relations would be more prominent in certain genres (Connelly, 2006; Lannon, 2007; Morenberg & Sommer, 2008; Reid, 2000). For instance, the words or phrases, such as *firstly*, *next*, and *in addition*, are thought to appear more in the process genre, indicating progressive relations in the text. However, while there are many mechanisms other than CAs able to convey different textual relations, few empirical studies were conducted to quantitatively prove whether the difference in the occurrence of CAs in different genres is at a significant level.

In view of the research potential of CAs and their discourse application, the present study aims at investigating the distribution of textual relation performed by CAs and their misuse patterns across genres and time by answering the following three research questions:

1. Does the distribution of textual relation performed by CAs vary according to different genres at a statistically significant level?
2. Does the distribution of textual relation performed by CAs in written register vary according to different genres at a statistically significant level?
3. Does the distribution of textual relation performed by CAs in spoken vary according to different genres at a statistically significant level?

## 1 Classifying framework

The significance of CAs lies in the fact that they direct the interpretation among sentences in text, which leads to the attempt to classify textual relations explicitly indicated by CAs. According to Halliday and Hasan (1976), *additive*, *adversative*, *causal*, and *temporal* were the four types of textual relations regulated, with various subdivisions in each type. Later, Quirk et al. (1985) revised Halliday and Hasan's four-type system as a system of seven types, namely, *listing*, *summative*, *appositional*, *resultive*, *inferential*, *contrastive*, and *transitional*. In addition, Granger and Tyson (1996)

designated one more textual relation, *corroborative*, for the kind of CAs convey writers' attitudes toward and comments on the previous text in the following text.

The present study based its classifying framework for textual relations on Quirk et al's (1985), except combining *resultive* and *inferential* into one category and including the *corroborative* textual relation. The merit of the framework is that most related studies took the same modified framework so that the study is reduplicable. Although the most related studies did not combine *resultive* and *inferential* into one category, the low occurring ration or even the zero occurrence (Altenberg & Tapper, 1998; Tankó, 2004) makes the category collapse have little impact on across-study comparison. Table 1 gives the definition and some examples of each kind of CA sorted by the textual relation it conveys.

Table 1: Classification of Textual Relations of CAs

Textual Relation	Definition	Example
Listing	Mark the next unit of discourse.	First, in addition, then
Transition-al	Shift attention to another topic.	now, in the meantime
Appositive	Provide an example or an equivalent.	in other words, for example
Summative	Conclude the information.	in conclusion, to summarize
Resultive/ Inferential	Mark the second part of the discourse as the result.	accordingly, then, as a result
Contrastive	Show incompatibility between information.	however, on the contrary
Corroborative	Express attitudes toward or comments on the text.	in fact, of course, actually

Given the preliminary classifying framework for textual relations performed by CAs, the present study also take the issue of register into account. Take the following excerpt for example. Both sentences criticize the disadvantages of purchasing some patented software for teaching, with the CA *besides* serving the *Additive* relation to signal one disadvantage plus another. While the textual relation between the two sentences is correctly expressed, *besides*, in fact, is considered informal, and should be avoided in formal writing.

If teachers want the students to buy the software from the publisher, it will cost students extra money. *Besides*, even if the students buy the software, teachers still can't make sure that they make good use of it.

Therefore, in this classifying framework, the use of CAs would be labeled based not only on the textual relation performed but also on the register of the CAs.

## 2 Methodology

The writing samples investigated and the analytic tools employed in the present study were based on the AWTA project (<http://awta.csie.ncku.edu.tw>), which aims at compiling a learner corpus providing inter-sentential information in hopes of developing automated discourse-correcting technology.

Based on the learner corpus compiled in the AWTA project, 45 writing pieces by 5 college students were selected in this study for analysis, with one student contributing 9 pieces which belong to 9 different genres. The nine genres are *Comparison-Contrast*, *Cause-Effect*, *Description*, *Definition*, *Narration*, *Classification*, *Multiple-Strategies*, *Argumentation*, and *Problem-Solving*.

After analyzing the selected data and tallying the counts, all the obtained figures were further analyzed via two-way within-subjects ANOVAs to answer the proposed research questions. A significant level of  $p < .05$  was chosen.

## 3 Results and discussion

All the raw occurring counts of CAs were transformed into the occurring counts per 1000 sentences. The purpose of the transforming the raw counts is to evade the influence stemming from different data sizes of different genres.

By observing these data, it is found that there seems to exist a norm distribution of textual relations performed by CAs without differentiating registers and CAs in written register. In this distribution, the listing and contractive relations are the two most frequently occurring types of CAs while the summative and transitional relations are the two least frequently occurring types. The rest of the textual relations are in the middle. Although few textual relations are performed by spoken CAs, the listing relation appears to be the most frequently occurring kind in the spoken CA form.

### 3.1 Does the distribution of textual relation performed by CAs vary according to different genres at a statistically significant level?

After calculating a two-way within-subjects ANOVA, the results show that there is no interaction between textual relation and genre ( $F(48, 192)=1.070$ ,  $p=0.366$ ) as well as no main effect from genre ( $F(8, 32)=1.697$ ,  $p=0.137$ ). However, there does exist a main effect from textual relation ( $F(6, 24)=10.476$ ,  $p<0.05$ ).

With a closer look at which pair of textual relations have significant differences, the distribution of textual relations is found to conform to the norm of *Listing* ( $M=23.596$ ,  $SD=6.625$ ) and

*Contrastive* (M=14.153, SD=2.199) being the most frequent. *Resultive/Inferential* (M=6.953, SD=0.610), *Appositive* (M=4.849, SD=1.369) and *Corroborative* (M=4.073, SD=0.917) occupying the second most frequent, *Summative* (M=1.200, SD=0.565) and *Transitional* (M=1.147, SD=0.512) are the least frequent.

The reason of presenting the distribution with different compartments rather than in a linear sequence lies in the fact that these compartments differ significantly in terms of frequency but that there is no significant difference between textual relations within the same compartment. For example, the occurring frequencies of the listing and contrastive relations are, except the resultive relation, significantly different from the other relations and have the highest and second highest frequencies, yet the two are not significantly different from each other.

### 3.2 Does the distribution of textual relation performed by CAs in written register vary according to different genres at a statistically significant level?

Through a two-way within-subjects ANOVA, it is found that there is no interaction between textual relation and genre ( $F(48, 144)=0.969$ ,  $p=0.537$ ) as well as no main effect from genre ( $F(8, 24)=2.062$ ,  $p=0.082$ ). However, there does exist the main effect from textual relation ( $F(6, 18)=8.585$ ,  $p<0.05$ ).

The results show that there is no distribution norm of textual relations of written CAs. Most of textual relations do not differ from one another to a significant level. Take the listing and transitional relations as an example. The former does not significantly differ from any textual relations while the latter only differs from the contrastive and resultive/inferential relations, which is very different from what happens in the distribution of textual relations of CAs without differentiating registers.

### 3.3 Does the distribution of textual relation performed by CAs in spoken register vary according to different genres at a statistically significant level?

The scarce occurrence of spoken CAs indicates it is not suitable for running statistical analysis because of many zero occurrences in textual relations. However, it is found through these raw figures that, in most genres, few textual relations are performed by spoken CAs except the listing and contrastive relations. Moreover, the occurrence of the contrastive relation is limited, with only one or two examples. It is the listing relation that appears most

frequently in the spoken CA form.

## 3.4 General discussion

The obtained results examine whether the distribution of textual relations performed by CAs is under the influence of genre. After running a two-way ANOVA, it is found that genre has no role in impacting the distribution of textual relations performed by CAs, and it remains a norm distribution across genres. That is, *Listing* and *Contrastive* are the most frequent. *Resultive/Inferential*, *Appositive* and *Corroborative* are the second most frequent. *Summative* and *Transitional* are the least frequent, which corresponds to Field and Yip's (1992), Liu and Braine's (2005), Chen's (2006) studies as well as Tankó's (2004), Altenberg and Tapper's (1998) and Shen's (2006) studies. However, with a closer look, the distribution norm fails when written CAs are isolated for analysis, and only the listing relation is constantly performed by spoken CAs.

This lack of genre influence might be related to textual-relation-creating mechanism. Textual relations in text can be manifested by various mechanisms, such as discourse-organizing words. Moreover, CAs can be easily paraphrased by discourse-organizing words. (McCarthy, 1991; Winter, 1977; Yu, 2007). As shown in the following example, the CA in (1) can be paraphrased as the noun phrase in (2), and the listing relation is still conveyed. Thus, only focusing on the textual relations performed by CAs may not show dominant textual-discoursal organization in genres.

There are many reasons causing global warming.

- 1) Firstly, the amount of carbon dioxide is three times than it was.
- 2) The first reason is that the amount of carbon dioxide is three times than it was.

On the other hand, textual relations on their own present a constant distribution norm, suggesting there is a preference for employing certain textual relations. For example, the transitional and summative relations occur least frequently at a significant level. This is understandable in that the two relations serve opening and closing functions which only appear at the beginning and at the end no matter how long a textual unit is.

Last but not least, although the distribution norm of textual relations is discovered in CAs, it is not true when CAs are divided into written and spoken registers. This may suggest that the distribution of textual relations is some natural development in cognition while register is learned through education.

#### 4 Conclusions and implications

The present study aims to investigate the distribution of textual relations performed by CAs. There are two major findings. Firstly, there is a distribution norm of textual relations performed by CAs free from influence of genre and time. It is inferred that the preference results from human cognition once the cognitive development is completed. Secondly, time has a role in impacting the use of CAs in different registers. Over time, the use of CAs in written register is increasing while that of CAs in spoken register is decreasing.

Given these obtained results, the contribution of the present study is three-fold. Firstly, this study follows the specifications formatted in the AWTa project, which makes the analyzed data with annotations for CAs serve as authentic samples for learners' reference and training data for developing automated discourse-detecting technology. Secondly, the results provide a new perspective on introducing CAs and the textual relations they perform. For instance, the distribution norm of textual relations performed by CAs across genres and time suggests that students are cognitively prepared to acquire and apply the type of language forms. The traditional textbook arrangement in which CAs are introduced by one textual relation after another in different genres is worthy of second thought.

For further research on textual relations, it is suggested to enlarge the sample size and research scope for investigation. Although it is found there exists a distribution norm of textual relations performed by CAs, which is free from the influence of genres and time, some conditions may have a role in the obtained results. Firstly, to avoid instruction influence, only five participants qualified to submit their writing samples. Nevertheless, with such a small sample size, the inferential power may not be enough and thus undermines the results. Therefore, getting the results from a different statistic design with a large sample size to compare with the results in the present study may consolidate the findings. Secondly, the mechanisms to realize textual relations are by no means limited to the CAs investigated in the present study. Whether the distribution norm of textual relations found in the present study still occurs after including the derived and paraphrased forms of CAs (McCarthy, 1991; Winter, 1977; Yu 2007), such as preposition expressions with references and discourse-organizing words, is another direction for further research.

#### References

Altenberg, B. & Tapper, M. (1998). The use of

- adverbial connectors in advanced Swedish learners' written English. In S. Granger (Ed.), *Learner English on Computer* (pp. 80-93). London: Longman.
- Chen, Cheryl Wei-yu. (2006). The use of CAs in the academic papers of advanced Taiwanese EFL learners. *International Journal of Corpus Linguistics*, 11(1), 113-130.
- Connelly, M. (2006). *Get writing: Sentences and paragraphs*. Australia: Thomson Higher Education.
- Field, Y. & Yip, L. (1992). A comparison of Internal conjunctive cohesion in the English essay writing of Cantonese speakers and native speakers of English. *RELc Journal*, 23, 15-28.
- Granger, S. & Tyson, S. (1996). Connector usage in the English essay writing of native and non-native EFL speakers of English. *World Englishes*, 15, 17-27.
- Halliday, R. & Hasan, D. (1976). *Cohesion in English*. New York: Longman.
- Lannon, J. M. (2007). *The writing process: A concise rhetoric, reader, and handbook*. New York: Longman.
- Liu, M. & Braine, G. (2005) Cohesive features in argumentative writing produced by Chinese undergraduates. *System*, 33, 623-636.
- McCarthy, M. (1991) *Discourse Analysis for Language Teachers*. UK: Cambridge University Press.
- Morenberg, M. & Sommers, J. (2008). *The writer's options: Lessons in style and arrangement*. New York: Pearson Longman.
- Quirk, R., Greenbaum, S., Leech, G. & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.
- Reid, J. M. (2000). *The process of composition*. White Plains, NY: Longman.
- Shen, T.-C. (2006). *Advanced EFL Learners' Use of CAs in Academic Writing*. MA Thesis, Taiwan: National Taiwan Normal University.
- Tankó, G. (2004). The use of adverbial connectors in Hungarian university students' argumentative essays. In Sinclair, J. M. (Ed.), *How to Use Corpora in Language Teaching* (pp. 157-181). Amsterdam: John Benjamins B. V.
- Winter, E. O. (1977). A clause-relational approach to English texts: A study of some predictive lexical items in written discourse. *Instructional Science*, 6, 1-92.
- Yu, Hsueh-ying. (2007). Discourse grammar for academic reading: Textual relationships. *English Teaching & Learning*, 31(2), 159-197.

# **An Acoustic Study of English Language Learning Transfer from Japanese: a case study in CNN English Conversation**

Yi-jun Liu

Department of Foreign Languages and Literature, National Cheng Kung University

fakemak118@yahoo.com.tw

## **Abstract**

Many studies investigate the effects of native language on second language learning, which is restricted to segmental level. Japanese and English are typologically different in terms of prosodic properties. The salient common feature between English and Japanese lies in the accent, but it operates in different ways in these two languages. The aim of this study is to examine the language transfer of Japanese speakers in English conversation in terms of suprasegmental area and discourse functional characteristics. One experiment involves two groups in this study –native American English speakers and Japanese speakers. The collected data is derived from CNN conversation. The experiment will investigate the prosodic differences between English and Japanese groups in post-nuclear position as well as temporal organization of function words in English conversation. As to discourse language, the experiment will make a comparison of reactive tokens as well as accent marking on topic. In addition, the extent of language transfer in different proficiency of Japanese speakers in English conversation will be probed further. The results of this study reveal the language transfer is relative to first language. The comparison between two languages provides several pedagogical implications.

## **Keywords**

language transfer, accent, post-nuclear position, temporal organization reactive-tokens

## **Introduction**

When learning Japanese and English, I find several differences, especially acoustic features, between these two languages. In the process of the second language acquisition or third language learning, it is not surprising that the native language would have the influence on the target language. This raises the motivation to conduct the research regarding language transfer to achieve better understanding for these two languages.

Several acoustic features occurring in Japanese and English can be illustrated as follows. Firstly, Japanese is regarded as pitch accent language and the use of pitch to mark certain syllables in the speech, which is similar to English. However, English is stress accent language, characterized by higher pitch, longer duration and greater intensity, which distinguish from Japanese that is limited to the change of pitch from a H tone to a L tone around the accented syllable(Beckman and Pierrehumbert, 1986a: 256). Secondly, phonemic length contrast determines the lexical meanings in Japanese language, which does not exist in English since the duration only operates at the phonetic level by exhibiting the longer in stressed syllables. Thirdly, English is classified as a stress-timed language, in which the fundamental isochronous unit of timing is the stress foot, while Japanese is classified as a mora-timed language (Ueyama, 2000: 9).

This study explores the language transfer on acoustic features when Japanese speakers are interviewed in English conversation. Two groups involve native English speakers and Japanese speakers in order to examine their differences in terms of intonational structures and discourse language in conversation. There are three objectives intended to be achieved in this research. Firstly, make a comparison of these two groups on intonational structures aspect. I will focus on the point that Japanese speakers how to manipulate the vowel duration in post-nuclear position as well as the temporal organization of function words. Secondly, exhibit their different communication styles in two groups by probing into discourse functional characteristics, such as reactive-tokens, or the accent marking topic. Thirdly, discover the extent of language transfer on different proficiency of English, which ground on the notion that communication styles relative to cultural knowledge is difficulty to remove (Chang, 2007).

## 1 Literature review

### 1.1 Intonational structure

Accented considered the most salient similarity between English Japanese only differs in how word stress is realized in creating characteristic stress patterns of each language (Ohata, 2004: 10). English is stressed accent language so that stressed syllables exhibit the longer duration, change of pitch and greater intensity (Ueyama, 2000: 9). Whereas, Japanese accented words is merely characterized by the change of pitch. Accordingly, this difference distinguishes stress accent from pitch accent language (Japanese). The notion of stress accent in English contributes to the unstressed vowel “shwa /ə/”, which is not correlated to Japanese. I think vowel duration may be difficulty items for Japanese to learn. According to the experiment conducted by Ueyama (2000), it indicates that advanced Japanese learners of English have trouble making the vowel duration different in post-nuclear position since advanced learners employ the fundamental frequency, which is correlated to Japanese lexical trait, as the tool beneficial for lengthening vowels. However, in post-nuclear position, fundamental frequency totally disappears in English so that advanced Japanese learner may encounter the problem of producing a durational contrast between stressed and unstressed vowels (Ueyama, 2000). I predict that fundamental frequency will have a positive transfer, while the vowel duration may be difficulty for all levels of Japanese people. Thus, I will compare different proficiency levels of Japanese learners of English after a contrastive focus in English conversation to confirm if this langue feature is the most challenge for every level of Japanese speakers in learning English.

In case of English stress-timing, complex syllable structures tend to be regarded as stressed syllables, while simple structures, such as function words (CV) will be found in unstressed syllables. Therefore, English speakers tend to cliticize function words so that the duration of function words are reduced as unstressed syllables of content words. On the other hand, Japanese function morphemes should be attached to content words, which rightly received no accent.

Japanese is mora-timed language so that the function morphemes are not be subject to shortening (Ueyama, 2000:110). I suggest this difference concerning with the temporal organization of function words may be neglected in classes, encouraging me to explore this subtle feature.

### 1.2 Discourse functional characteristics

As to discourse language in conversation, English speakers often feel that Japanese speakers are so irruptive since Japanese people like to provide more emotional supports when the speakers are in progress. The emotional support presents one type of back-channel feedbacks with the feature of a non-lexical vocalic form. Emotional support serves as a ‘continuer’, for purpose of revealing interest or claiming understanding, which frequently occurs in Japanese’s non-primary speakers. On the other hand, English non-primary speakers are more inclined to utter a short lexical phrase or word, such as ‘good’, ‘really’ ‘exactly’, in conversation, called as Reactive Expression (Clancy, Thompson, Suzuki and Tao 1996). I suppose that the discrepancy of discourse functional characteristics in conversation reveals the contrast communication style, which is relative to cultural background knowledge. In my view, this aspect concerned with cultural perspective is difficulty to get rid of regardless of their English proficiency. This domain is exactly I want to discover.

Beckman and Venditti (2000) find that Japanese have the tendency to utilize the NP-wa to re-introduce the topic again together with prominence by probing into the spontaneous speech of Japanese language corpus. On the contrary, English speakers view the re-introduced topic as old information without putting accent on it. Moreover, as the experiment about language transfer conducted by Wennerstrom (1996) indicated that Japanese group is prone to put equal prominence on each item notwithstanding the importance in discourse context. In response to the above finding, I think Japanese may have their communicative strategy to signal the relationship between old and new information, which entails more attention to examine.

## 2 Method

### 2.1 Subjects

There are two groups participating in this study. One consists of two native English speakers who both grow from New York city. The other embraces two native Japanese speakers who exhibit different English proficiency in CNN interview conversation. Paris Hilton (female), a famous American socialite and Howard Schultz (male), a chairman of Starbucks form American English group. Hamasaki Ayumi (female), a famous Japanese singer and Yoshizo Shimano (male), a CEO of bicycle corporation construct Japanese English group. Hamasaki Ayumi often engages code-switching in Japanese so that Ayumi is classified as beginning level, as compared to Yoshizo Shimano who has

better command of English in CNN interview conversation.

## 2.2 Data collection

Data collection is mainly derived from the book “Master Listening with CNN News-- Interviews with Celebrities”, including American group subjects and Yoshizo Shimano. The CNN conversation of Hamasaki Ayumi is obtained from the internet:

<http://www.veoh.com/videos/e1749709ab7Nnnr>.

Therefore, the data in this study are authentic.

## 2.3 Measurement

Software Parat was utilized to analyze the fundamental frequency and vowel duration. Fundamental frequency determines tonal organization, such as intonational structure. Vowel duration is manipulated in different ways in two languages.

## 3 Data analysis

### 3.1 Intonation structure

The findings can be displayed in the figure 1, 2, 3, and 4.

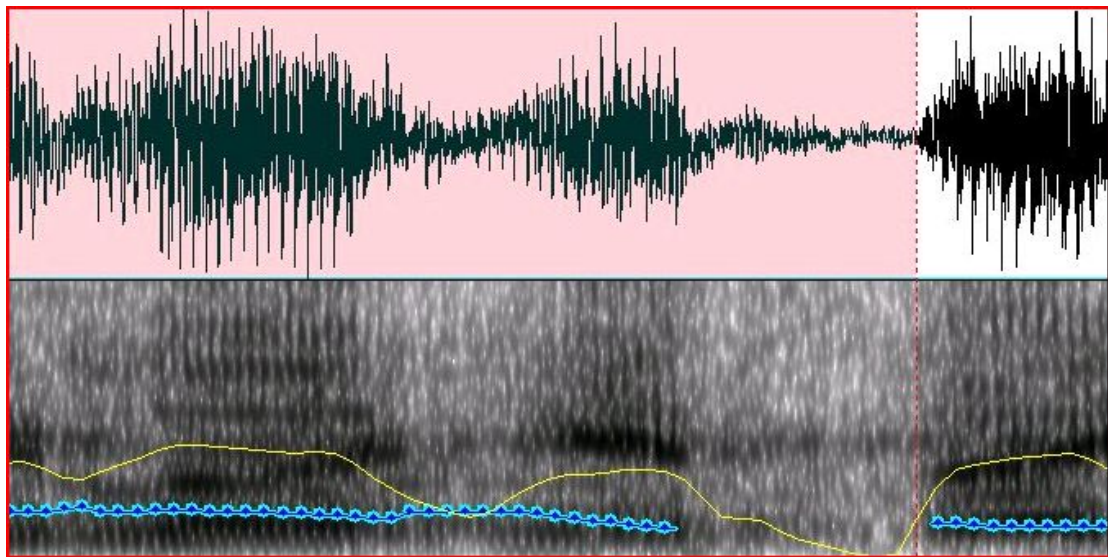


Figure 1. Male American English Speaker

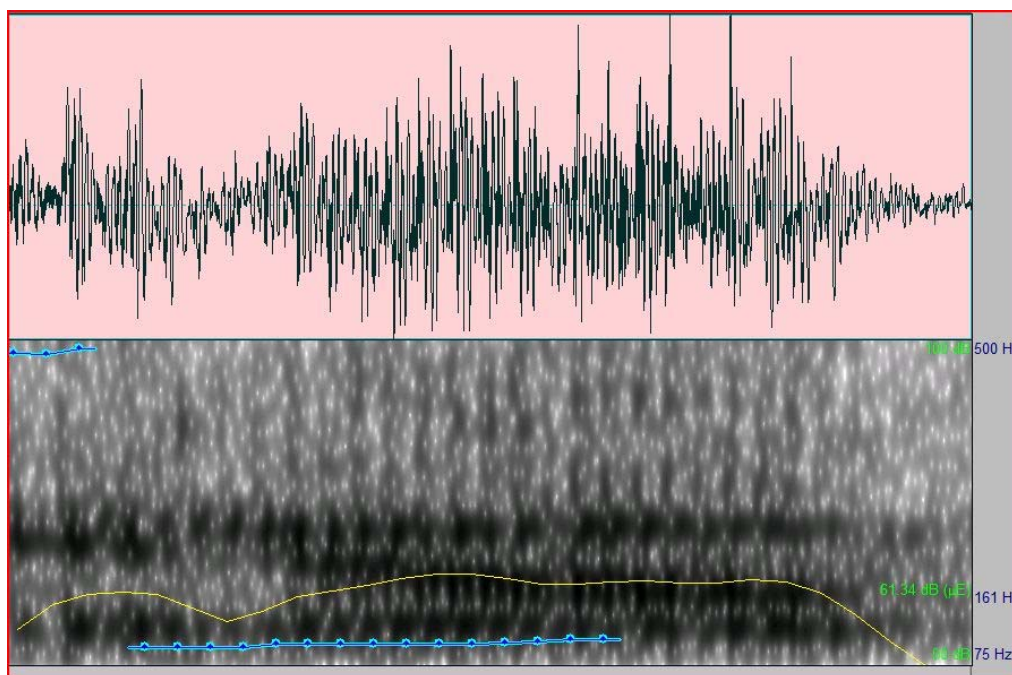


Figure 2. Male Japanese Speaker

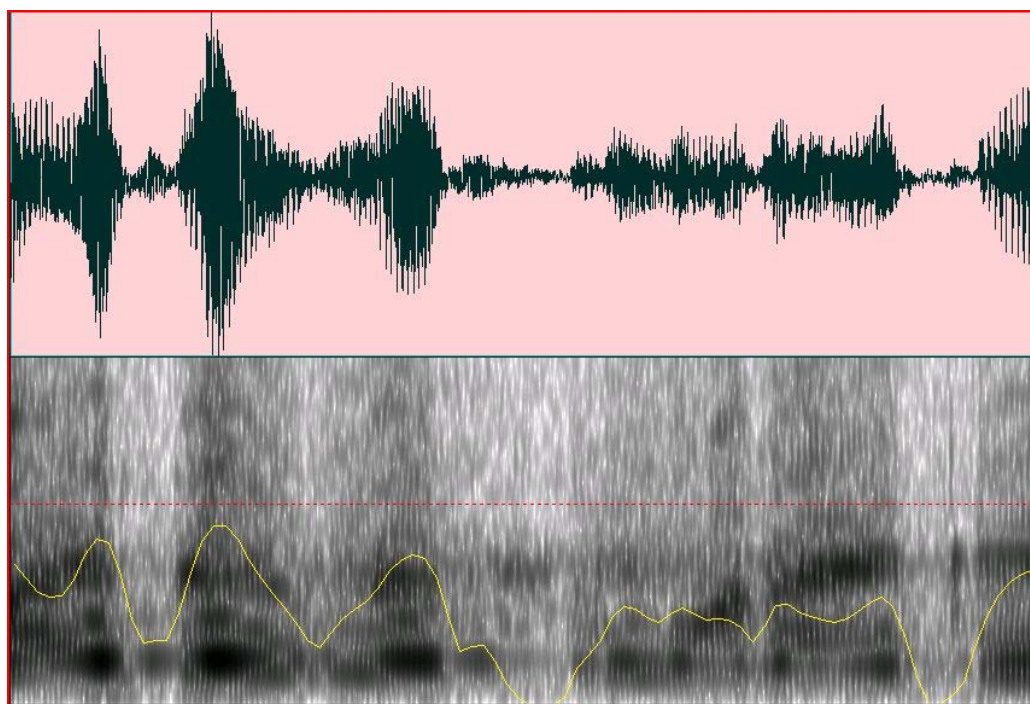


Figure 3. Female American Speaker

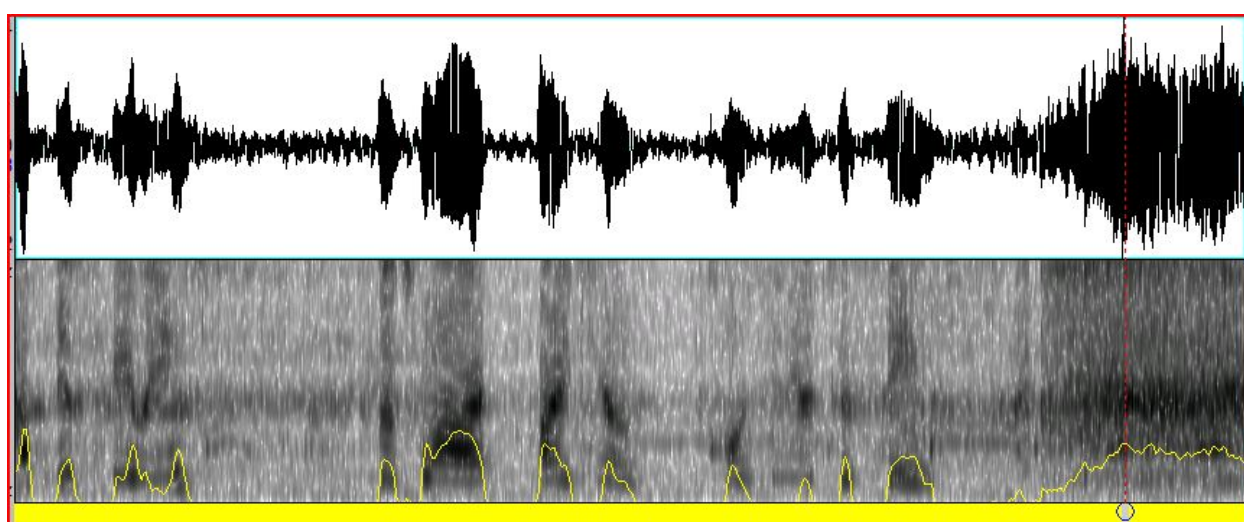


Figure 4. Female Japanese Speaker

The results indicate Japanese speakers have difficulty producing a durational contrast in post-nuclear position. However, the change of pitch can be positively transferred in Japanese English. Advanced subject (Figure 2) is inclined to perform post-nuclear deaccentuation more reliably. This confirms the findings Ueyama (2000). As for the function words, Japanese speakers, especially, low level subject, tend to extend the duration of function words as compared to American English speakers. Therefore, we discern that the language transfer is highly relative to the influence of native language.

### 3.2 Discourse functional characteristics

Hamasaki Ayumi really provided more emotional supports, such as non-vocalic back-channels in conversation. By contrast, Yoshizo Shimano did not

have an irruptive style as I predict, which may result from their different occupation. Hamasaki Ayumi is a famous singer so that she should exhibit her lively personality, whereas Yoshizo Shimano is a CEO so that he would be careful in personal conducts. Besides, I think this phenomenon can probably be explained by their educational backgrounds. As to the accent marking, I find Japanese speakers influenced by native language, they are prone to put the topic in the initial position of sentences with prominence. The topic is already mentioned in the previous. Consequently, I find that the language transfer is not restricted to the prosodic properties, which also contains the transfer from Japanese syntactic structure. Specifically, Japanese language involves the pattern “NP-wa” to re-introduce the topic, which is actually exhibited in

Japanese English.

### 3.3 The extent of language transfer

Advanced subject, Yoshizo Shimano, is more native-like since he could produce more contrast duration and do not perform the irruptive communication style. However, two Japanese speakers cannot have the ability to cliticize function words in English, especially low-level subject.

## 4 Discussion and conclusion

The study is aimed to examine the langue transfer of Japanese performed by Japanese speakers in English CNN conversation. The findings can be summarized as follows. Firstly, in the aspect of intonational structures, Japanese speakers regardless their English proficiency, have difficulty producing durational differences after contrastive focus. As to the change of pitch, advanced subject do well in suppressing F0 contrast after a contrastive focus than the beginning subject. Accordingly, the results can be supported from Ueyama's experiment (2000). The reduced function words could be challenge for Japanese speakers, which reflect on the fact Japanese speakers, especially low-level subject (Hamasaki), tend to extend the duration of function words. This langue transfer phenomenon can be explained by the nature of Japanese language that is a mora-timed language. In addition, Japanese grammatical morphemes are suffixes certainly attached to words. Consequently, it poses the difficulty for native Japanese speakers to cliticize English function words (Ueyama 2000: 121).

Secondly, in the area of discourse functional characteristics, it is surprising that the advanced speaker do not reveal the irruptive style by offering emotional support in progress when interviewing. However, the Japanese communication style can be found in Ayumi's conversation. I think this result is not inconsistent my hypothesis may come from the limitation of the sample sizes, background knowledge of subjects, such as their occupation, educational backgrounds and so on. The accent marking on the old information really occurs in Japanese English. As compared to advanced speaker, low level subject like to re-introduce the old information in the beginning of sentences with prominence. This sentence pattern is equal to the emphasize sentence –NP- wa, which exhibits the language transfer not only comes from the prosodic properties but also result from the syntactic structures. Compared to English speakers, Japanese subjects fail to display the wider pitch range in English intonation patterns. As Ohata suggested (2004), the narrow pitch movement of Japanese

English may contribute to the misinterpretation of English native speakers. For instance, when Japanese speaker fails to lower the pitch at the end of a sentence, the utterance might be perceived as a continuation of the speech. This example of misinterpretation could be the most problems for Japanese learners in English communication. I consider this problem could occur in Taiwanese learners, which entails more attention to investigate. Thirdly, the degree of langue transfer would be more in low-level subject in terms of intonational structures and discourse language. The most common as well as difficulty learning items should be the reduced the function words and duration contrast in post-nucleus position. I assume that the suprasegmental features may be the ignored domains in Japanese English teaching. Moreover, I find that English speech performed by Japanese speakers may be staccato-like, which may be difficulty for American English speakers to comprehend. The fact of staccato characteristic influenced by their native language mora-timed also corresponds to Taiwanese English due to a syllable-timed feature in Mandarin. I suggest English instruction for Japanese as well as Taiwanese learners should emphasize the development of the awareness for stressed syllables, namely, loudness, vowel length and clarity, and pitch, and cultivate the ability of recognizing the stress-timed rhythm in English.

Furthermore, even if the communication style is not so significantly influenced by native language when observing the advanced subject, I recommend the cultural knowledge should be incorporated into English instruction, which is really ignored in Taiwanese English teaching. Cultural knowledge introduced into English teaching would cultivate the students' international perspective so that they can learn how to be tolerant different culturally communication styles.

## 5 Reflections

According to behaviorism, langue learning is a process of habit formation. When learning a new language, we are tried to get rid of the influences of native language or other familiar languages. This phenomenon reflects on my third language learning of Japanese. Sometimes, I am really afraid of my English pronunciation with Japanese accent. The notion of behaviorism reflects on the impression of the public that the master of Japanese will decrease the English proficiency, after all, there are many differences between these two languages, which raises the motivation to do this study regarding the langue transfer from Japanese on English learning. This study involving the comparison between English and Japanese improves better

understanding of these two languages. I am more aware of some similar features, such as accent, stress, how to manipulate in these two languages. I discern the language transfer of Japanese people on English learning from this research so that I can pay more attention to these language transfers from Japanese. Moreover, I learn the factors inducing these language transfers. This is to say I really achieve the comprehensive understanding of English and Japanese. In fact, learning English, I think it is difficulty to control the vowel duration in stressed or unstressed syllables, which also occurs in Japanese people since Japanese and Mandarin Chinese are not classified as stress-timed language. Through this research, I get several insights about how to control the vowel duration of English. On the other hand, when learning Japanese, pitch accent, a trait of accent word poses difficulty on me. I do not know how to mark the pitch accent on the word. Due to this research, I finally realize the pitch accent on Japanese word only involves the pitch change from a H tone to a L tone. I am really glad to be a graduate student in linguistics field so that I can take the course of English intonation analysis, which not only solve my English learning problems but also improve the understanding of Japanese.

This study provides the differences of these languages, which is not confined to intonation structure aspect. I also learn the discourse area in these two languages operates differently, which cause a contrast communication style. I think the discourse language aspect is relative to cultural background knowledge, which is hard to remove and often neglected in English teaching. Therefore, I suggest the discourse language should be incorporated into English teaching, enable students to discern the different communication styles and then learn to be tolerant. In fact, many English speakers feel Chinese people tend to be an inactive communication style as compared to Japanese's irruptive style. In the future, I want to engage the international business secretary and thus the knowledge about different cultural communication style are beneficial for me to interact or communicate with people from all over the world smoothly.

## References

- Beckman, M. E. & Pierrehumbert, J. B. (1986a). Intonational structural in Japanese and English. *Phonology Yearbook*, 3, 255-309.
- Beckman, M. E. & Pierrehumbert, J. B. (1986b). Japanese prosodic phrasing and intonation synthesis. *Proceedings of the 24th annual meeting on Association for Computational Linguistics* (pp. 173-180). New York: Association for Computational Linguistics.
- Beckman, M. E. & Venditti, J. J. (2000). Tagging prosody and discourse structure in elicited spontaneous speech. *Proc. of Science and Technology Agency Priority Program Symposium on Spontaneous Speech* (pp. 87-98). Tokyo.
- Clancy, P. M., Thompson, S. A., Suzuki, R., & Tao, H. (1996). The conversational use of reactive tokens in English, Japanese, and Mandarin. *Journal of Pragmatics*, 26, 355-387.
- Gundel, J & Torsten, F. (2001). "Topic and Focus." *Handbook of Pragmatic Theory*. Oxford: Blackwell.
- Katagiri, Y. (2007). Dialogue functions of Japanese sentence-final particles 'Yo' and 'Ne'. *Journal of Pragmatics*, 39, 1313-1323.
- Ohata, K. (2004). Phonological differences between Japanese and English: Several potentially problematic areas of pronunciation for Japanese ESL/EFL learners. *Asian EFL Journal*, 6(4), Article 5. Retrieved December 9, 2008, from [http://www.asian-efl-journal.com/december\\_2004\\_KO.php](http://www.asian-efl-journal.com/december_2004_KO.php).
- Chang, S. J. (2007) A Graphic Approach to Teaching Communication-Concerned, Culture-General Constructs: Its Rationale, Design and Application. *English Teaching and Learning*, 31(1): 1-42.
- Ueyama, M. & Jun, S-A. (1998). Focus realization in Japanese English and Korean English intonation. In Hajime Hoji (Ed.), *Japanese and Korean Linguistics 7*. CSLI: Stanford University Press.
- Ueyama, M (2000). Prosodic Transfer: An Acoustic Study of L2 English vs. L2 Japanese. Doctoral dissertation, University of California at Los Angeles.
- Venditti, J. J. (2005). The J\_ToBI model of Japanese intonation. In Sun-Ah Jun (Ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing* (pp.172-200). Oxford: Oxford University Press.
- Ward, N. & Tsukahara, W. (2000). Prosodic features which cue back-channel responses in English and Japanese. *Journal of Pragmatics*, 32, 1177-1207.
- Wang, L.S. (2008). Master listening with CNN news—interview with celebrities. Taipei: LiveABC Interactive corporation.
- Wennerstrom, A. (1994). Intonational meaning in English Discourse: A study of non-native speakers. *Applied Linguistics*, 15(4), 400-420.

# The effects of conference participation on communication apprehension

Rieko Matsuoka

School of Nursing, National College of Nursing, Japan

matsuokar@adm.ncn.ac.jp

## Abstract

This study examines the process by which the Japanese college students reduced their level of communication apprehension through their experiences of working as interns at an international conference. Data were obtained from the focus group meeting of nine Japanese students who volunteered to work as interns at the international conference in 2008. A 120-minute meeting was audio-recorded and transcribed. Using a content analysis approach (Krippendorff, 2004), the transcribed data were coded into meaningful units, which were categorized into concepts representing the participants' psychological features. Elicited concepts such as perfectionism, competitiveness, and other-directedness were discussed as the possible causal factors of communication apprehension. Socio-psychological and anthropological studies on Japanese communication behaviors (e.g., Kuwayama, 1992; Lebra, 2004; Maynard, 1997) were examined for interpreting the data. As previous studies (e.g., Matsuoka, 2006) suggest, communication apprehension is regarded as the strongest antecedent to willingness to communicate, and some studies revealed that the Japanese communication apprehension is the highest amongst many nationalities including other East Asian countries; therefore, the teaching professionals should find the ways of reducing the Japanese learners' levels of communication apprehension in order to raise their willingness to communicate. Therefore, this study has important pedagogical implications in reducing students' communication apprehension in English.

## Keywords

communication apprehension, content analysis, international conference. Japanese sense of self.

## 1. Introduction

In this global society, English has become an international 'lingua franca' and has been regarded as a necessary language for being a member of the

international community. Even in such a milieu, the high degree of communication apprehension has impeded the communicative competence among Japanese learners of English.

Communication apprehension, the mental construct defined as "an individual's level of fear or anxiety associated with either real or anticipated communication with another person or persons" (McCroskey, 1977, p. 78), has been identified as the strongest factor in reducing the level of willingness to communicate among Japanese college students (Matsuoka, 2006). Communication apprehension is also identified as a psychological phenomenon called social anxiety. People experience this social anxiety whilst speaking in front of others, as Young (1991) has also observed. Accordingly, communication apprehension is likely to be a primary reason for the avoidance or disruption of communication.

Using content analysis approach, meaningful codes or units were elicited and categorized into concepts which seem able to explain the Japanese sense of self. These categorized concepts consisting of codes or units may explain the core mentality of Japanese learners or Japanese sense of self.

The focal point of this study is to find out the way in which the experiences of being interns at an international conference were successful in reducing the levels of communication apprehension of serious learners of English were able to reduce the level of communication apprehension.

## 2. Methods

### 2.1 Participants

Nine students majoring nursing at the national college participated in this study. They helped the international conference as interns for three days in November in 2008. Five of them were females and four are males. Five of them (three: female; two: male) were the third year students, whose age ranged from 20 to 21; four of them (two: female; two: male) were the second year students whose

age ranged from 18 to 19. The average age was 19.67. Four participants out of five third year students volunteered to work as interns in 2007, and they knew what they were expected to do. One first year student was a returnee and lived in the UK for three years; the other first year student was from the international high school where foreign students existed and English education was emphasized. All three female second year students had some experiences of traveling overseas. Two students out of the remaining four students helped at the conference in the previous year. The rest of two students did not have any experiences of contacting native speakers of English outside of the classroom settings, and one of them (first year male student) confessed he had high anxiety in speaking English.

## 2.2 Procedure

Nine participants joined a 120-minute meeting, and the conversation there was audio-recorded and transcribed. The questions asked at this meeting included a) whether they had any changes in their communication behaviors in English, b) what sorts of difficulties or problems they had during the conference sessions, and c) how they felt in speaking English. In a relaxed condition, they were expected to express their feelings, reflecting their experiences.

## 2.3 Content analysis approach

Content analysis is the research method making valid inferences from the text (Weber, 1990). In order to make valid inferences, Krippendorff (2004) postulates that the following six points need to be addressed; a) data, b) definition of the data, c) the population of the data d) context, e) the boundaries, and f) the target of the inferences.

The data for this present study were the transcribed text of the conversation at the focus-group meeting, and were defined as the experiences as interns at the international conference. The population of the data was the Japanese college students who are eager to improve their communication competence in English. The context was the international conference where the participants helped as interns. The boundary of analysis exists in the contents affecting the participants' communication behaviors in English. Lastly the target of the inferences was to explore the mechanism of reducing the Japanese learners' communication apprehension.

The inferences are made through coding in order that the latent meanings should be elicited from manifest contents. Then the elicited codes are categorized into viable factors.

A hypothesis is also required for content

analysis. The hypothesis of this study was "the participation as an intern at the international conference can reduce the level of communication apprehension.

## 3. Findings and Discussion

Several meaningful codes elicited from the data seem to develop the viable factors of reducing the communication apprehension levels. The frequent codes were 'to gain many more opportunities to use English than in their usual lives', 'to understand the importance of communication even when the communication skills are not good enough', 'to gain confidence of speaking English after making themselves understand themselves', and 'to feel happy and confident in speaking English'. In clear contrast to prior studies (e.g., Matsuoka, 2006), which revealed the Japanese overly conscious mentality toward outside world or other people (Kuwayama, 1992; Lebra, 2004, Maynard, 1997), these elicited codes may build up the promising factors reducing their communication apprehension levels.

The viable factors from these codes should be 'positive attitudes towards using English' and 'self efficacy of speaking English'. Among others, these two factors should function effectively for Japanese learners to reduce their communication apprehension levels.

## 4. Concluding remarks

The results of the present study support the hypothesis that the participation as an intern at the international conference can reduce the level of communication apprehension', and suggest that the Japanese sense of self manifested as the mentality of 'other-directedness' could be defeated should the optimal occasions of using English and experiencing the self-efficacy be given.

## Selected References

- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*. 2<sup>nd</sup> edition, CA: Sage.
- Kuwayama, T. (1992). The reference other orientation. In N. R. Rosenberger (Ed.), *Japanese sense of self* (pp. 121-151). New York: CUP.
- Lebra, T. S. (2004). *The Japanese self in cultural logic*. Honolulu: UHP.
- Matsuoka, R. (2006). *Japanese college students' Willingness to Communicate in English*. Doctoral Dissertation. Philadelphia: Temple University
- Young, D. J. (1991). Creating a low-anxiety classroom environment: What does language anxiety research suggest? *The Modern Language Journal*, 75, 426-439.

# Taiwanese College Students' Perceptions of Plagiarism: Is it a Cultural Issue?

Shih-Chieh Chien

Department of Foreign Languages and Literature, National Sun Yat-sen University

chien.paul@gmail.com

## Abstract

The present study aims to investigate Taiwanese college students' perceptions of plagiarism. Data were collected through semi-structured in-depth interviews with students. In this study, it identified a variety of reasons why students may plagiarize. Most of the students participating in this study do not think plagiarism is a cultural issue. Rather, they regard it as a worldwide phenomenon. There are a number of factors that may lead to plagiarism, including the rapid evolution of information and communications technology making it much easier for students to copy and paste information via Internet, lack of adequate study skills and improper preparation particularly in poor time management, poor academic writing and referencing skills, lack of confidence to write in their own words due to difficulties and unfamiliarity with a foreign language, being under lots of pressure because of many too hard assignments, high expectations of gaining the best grade, utter laziness, individual differences etc. This study suggests that education plays a very important role in training and making students aware of the ethos and values of academic honesty and integrity, and the acknowledgement of sources of information as they relate to citing other people's scholarly works.

## Keywords

plagiarism; cultural; perceptions; Taiwanese

## Introduction

Plagiarism, as defined in the Collins Dictionary and Thesaurus (1991) is "to appropriate, borrow, crib, infringe, lift, pirate, steal, thief – from another work or author" (p. 753). The plagiarism in higher education is becoming a nerve-racking problem as such practices not only undermine the quality and standards of education but also contradicts the rule of intellectual property rights. Some studies have pointed out cultural factors as a cause why foreign students plagiarize and

indicated that they have a dissimilar cultural understanding of plagiarism (Briggs 2003; Juwah, 2003). The concept that plagiarism is ethically wrong is a common view in academia. Kollich (1983) portrays how he pursued offending plagiarists like an avenging god and explains what they have done as deception. He regards the act of plagiarism as an offense and makes out the need for punishment against what he sees as a crime. In academia, plagiarism is viewed as academic dishonesty or academic fraud. It is related to the unwarranted claim to the plagiarizing author's standing that is accomplished through false declaration of authorship.

The emergence in concern with regard to plagiarism in higher education is underscored in a number of recent articles (e.g., Ashworth, Bannister, & Thorne 1997, Carroll & Appleton, 2001, Errey, 2002). It is clear that there are different levels of understanding of plagiarism in different regions. These concerns are supported by some recent literature. Plagiarism is, however, not considered a problem in some other cultures. In some regions it may be acceptable, even flattering, to copy the work of experts. In some cases it is deemed more humble than boldly promoting their own viewpoints about something (Carroll & Appleton, 2001) and students who want to write something particularly clearly see paraphrasing the source as a strange thing to do when the source itself is written better than they could ever reword or rewrite it in their own language (Biggs & Watkins, 1996). The ethics of plagiarism is associated with the ideological standpoint of the ownership of texts. Scollon (1995) indicates that in the treatment of academic plagiarism there is an assumption of a common ideology of thinking which presupposes that the author is the only creator and originator of his/her texts. As a result, plagiarism is a violation against the author and accordingly viewed as ethically wrong.

## 1 Research design

The present study sought to explore whether there are notable cultural differences in the student's interpretations of plagiarism and to reflect on some of the implications for teaching practice. Twenty one college students from a national university in Taiwan participated in the study. An explanatory case study approach as described by Yin (1993; 1994) was the methodology used in implementing the study. The aims of the study were to (1) explore students' understanding of the term plagiarism; (2) identify the factors that may cause some students to plagiarize; (3) scrutinize and identify effective strategies and methods for conquering the factors that potentially cause students to plagiarize; and (4) build a knowledge base which would be helpful in informing the pedagogical guidelines and practices of coping with factors that cause students to plagiarize. The present study addressed the following questions:

1. Do students understand the term of plagiarism and its influence on academic writing?
2. What factors influence on how students understand plagiarism?
3. For what reasons do students plagiarize?

The investigation consisted of in-depth semi-structured interviews with the students in order to establish a broad understanding of the issues concerning plagiarism as perceived by them. The questions covered the students' level of understanding of plagiarism, academic writing, referencing conventions and academic integrity, the consequences linked to plagiarism, and students' prior educational background to help ascertain conceptions or misconception of plagiarism. Student participation in the interviews was voluntary and anonymity was ensured. All the data were transcribed verbatim, analyzed and reported. They were analyzed qualitatively. The results were presented in the form of qualitative comments. The students' perceptions of plagiarism were the evidence to help illustrate points.

## 2 Results

The present study comprised of interviews with the students in order to ascertain a wide-ranging understanding of the issues concerning plagiarism.

### 2.1 Students' perception of plagiarism

In the present study, most students mentioned that there is a need for greater understanding of plagiarism. Based on the interviews with students, a majority of them do think it is necessary to rewrite

or use quotation marks to cite an author's words even though the author is well known and respected. Hence it is an important step to avoid incidences of plagiarism. In addition, in academia, plagiarism is a serious offense that can lead to punishments such as a failing score on a particular assignment or for a course. Most of the students participating in this study do not think plagiarism is a cultural issue. Rather, they regard it as a worldwide phenomenon and plagiarism does not seem to be unique to students in Asia.

### 2.2 Plagiarism in coursework

In the present study, several students stressed how plagiarism is often unintentional owing to the way they make notes while researching and writing their essays. They elucidated that some students may gather plentiful electronic and non-electronic references whilst writing the notes they prepare when researching a particular essay question. In this way they indicated that it was easy to lose track of what is their own work and what is the work of others as they piece it together in their ultimate coursework submission. In this way students regarded some form of plagiarism as being unintentional. For instance, a student suggested that not completely referencing the patchwork could come about because of time constraints and / or the poor time management of research projects. He suggested that time is an issue saying that "most of the cases arise when students are lack of time, and merely do not have time to think about it."

### 2.3 Ability and Plagiarism

The question of academic ability or competency in English and/or the topic was seen to be associated with understanding why some students plagiarize. On the basis of their insights gained in learning English, several students said that, by reason of English not being the first language, that "taking a bit here and there helps with gathering meaning across. It is difficult to paraphrase if you are not a native speaker." Others also mentioned that there are only so many ways that issues could be written, and usually if an author had written something clearly and you agreed with it, then there was nothing wrong with copying this. As one student commented, "All the ways for saying something have already been said, and thus we have to use the same words. But this is about words and not concepts." Academic writing may be particularly challenging for students who do not have a solid grasp of

English.

## 2.4 Importance of Examinations

All the students in the present study held the view that cheating in an exam is not acceptable. One reason for this is that as the exam is the final evaluation then students should have, “an equal chance to prove themselves, cheating is something that is immoral.” Another student agreed with this saying that people work hard for exams, and they are a very nerve-racking period, which results students in resenting others that take notes into exams or cheat to some extent. The other student also said that the teacher’s requirements are not as high with exams as they are with coursework, and thus students feel less stress to cheat. In addition, as has already been elucidated, exams often bear a much higher percentage than coursework, and as such the risk of being caught cheating in exams was seen as being substantial by students. Although another student echoed with the above views, she also added “teachers are well conscious of degrees of plagiarism in coursework – while an exam they feel that you should do it alone.”

## 2.5 Technology

Technology is an interesting theme to pursue when exploring plagiarism in coursework. One theme emerged, the role of the Internet in coursework.

There is a perception that over the past decade there has been a rise in the incidences of plagiarism at schools. This is usually due to the increase in availability of material via the Internet. While plagiarism in scholarship has a long history, the ease of access to materials on the Internet, where articles come out as electronic text, makes it easier for students to copy and paste information directly into their assignment, thereby exacerbating the problem of plagiarism. Students mentioned they made use of the Internet mainly for finding information related to a particular academic topic. The issue of patching should be noted at this point as several students mentioned that some keep some windows open at one moment in time and copy and paste words and sentences into the essay they are working on. They acknowledged that this was unethical in terms of not clearly giving the credit of others’ work, or on the other hand, very tempting with regard to passing it off as their own work. This is said to be most likely among EFL students because of their levels of language proficiency, or among students having problems understanding material, or merely because of time pressures.

## 2.6 Learning and plagiarism

One interesting view that emerged amongst

students was that plagiarism is inevitably intertwined with student learning and development. For example, a student commented that when students plagiarize work well, it often still necessitates an understanding of the topic, and therefore reveals a degree of learning, saying that:

If you take all the words/sentences from other writers – then you need to do the work to assemble it together – you have learned and had a certain understanding of the subject matter, it is not merely copying.

Some students supported this point of view, stating that being able to make an argument in a coursework, even though some of it was plagiarized from a variety of sources, showed some degree of learning.

## 2.7 Motivations for Cheating

One of the other reasons that may induce students to cheat in exams or to plagiarize during their coursework is owing to pressures arising from accomplishing a certain level of grades. Grades were seen by the students to be one of the major incentives behind cheating. The students mentioned that competition was somehow fierce at the school. For example, one student explained that it was particularly fierce and tough in certain subjects such as English and American Literature. He went on to explain that gaining a high grade is important if students want to gain entry onto studying abroad. A good grade is also important in terms of finding a decent job. Another student also indicated, “Grades almost mean everything when students have no prior work experience. Grades are also one of the very important things that companies judge them on.”

## 3 Conclusion and Implications

In conclusion, this study reveals that the practices we deem plagiaristic are the result of several diverse and complex influences. It is clear from the interviews that most of the students think there are no significant local and cultural differences in terms of the practices of plagiarism. On the one hand, the ideological basis of the concept of plagiarism (due to learning skills, language, time pressures, grades, coursework evaluation etc.) could potentially offer reasons why students may plagiarize. It seems that the students agree with the generally accepted definition of plagiarism; that of stealing someone else’s words and sentences and passing it off as their own, is ethically wrong. Pedagogical models, assessment practices,

writing practices and institutional arrangements may place more weight in terms of teaching students how to avoid plagiarism. Teachers need to help students improve academic writing practices, referencing conventions, language, and study skills, thereby enabling them to increase their confidence, handle academic work better, and thus avoid incidences of plagiarism.

Finally, this study, owing to its limited scope, is tentative indicative of the issues at stake that may be generalizable. There is still more work to do. For instance, more detailed qualitative studies of writing practices might help us to gain insights into how 'cut and paste' writing and patch writing are employed to make arguments. We need to realize what kind of learning is involved and how to develop teaching techniques to help the students go on to independent writing practices. We need to have a better understanding of the skill gap that EFL students get there. It seems that if we were serious about plagiarism, it would be indispensable for us to throw away the ideological stances and do more comprehensive research to understand the process of these practices.

## References

- Ashworth, P., Bannister, P. & Thorne, P. (1997). Guilty in whose eyes: University students perception of cheating and plagiarism, *Studies in Higher Education*, 22, 2, pp.187-203
- Biggs, J. & Watkins, D. (1996). *The Chinese Learner*, Comparative Education Research Centre, Hong Kong
- Briggs, R. (2003). Shameless! Reconceiving the problem of plagiarism. *Australian Universities Review*, 46, 1, 19-23.
- Carroll, J. & Appleton, J. (2001). Plagiarism: A Good Practice Guide, JISC, Oxford Brookes University
- Errey, L. (2002). Plagiarism: Something Fishy? ...Or Just a Fish out of Water?, OCSLD
- Juwah, C. (2003). The Robert Gordon University Seminar Series on Plagiarism.
- Kolich, A. M. (1983). Plagiarism: The worm of reason, *College English*, 4, 141-148
- Scollon, R. (1995). Plagiarism and Ideology: identity in intercultural discourse, *Language in Society*, 24, 1, 1-28.
- The Collins Dictionary and Thesaurus (1991). Glasgow: Harper Collins Publishers.
- Yin, R. K., (1993). *Applications of case study research*. Newbury Park, CA: Sage.
- Yin, R. K., (1994). *Case study research: Design and methods*. Thousand Oaks, CA: Sage.

## Importance of Teaching Gender-neutral Language in EFL Classrooms: Students' Survey Results

Chiyo Myojin

Core Studies Department, Kochi University of Technology

myojin.chiyo@kochi-tech.ac.jp

### Abstract.

The use of gender-neutral language instead of gender-biased one is rapidly becoming important in the business and academic world. In fact, gender-neutral words such as “human being”, “person” and “chairperson” are commonly used as generic meanings in English-speaking countries, whereas the use of gender-biased words such as “man”, “he”, and “chairman” has been avoided recently.

Therefore, it seems very important for us EFL teachers to equip our students with ways to avoid sexist language by teaching it in classrooms. Otherwise, our students may not be aware of such language and may even use it in English-speaking settings in the future. Therefore, in order to enable our students to skillfully interact in authentic situations, as EFL teachers, we should have a responsibility to equip our students with an understanding of gender-biased language and appropriate substitutions for such language.

First, this paper explained why gender-neutral language should be used. Second, this paper reviewed exactly what gender-biased terms have been replaced by what gender-neutral terms in English-speaking countries recently (Myojin, 2004). Finally, this examines how much EFL learners are currently aware of gender-biased language by administering a questionnaire survey toward about 66 Japanese university students.

### Key words

Gender-neutral language, Gender-biased terms, genetic pronoun, EFL classrooms

### Introduction

Since our language and society reflect one another, it is quite important for us as communicators to recognize and respect change in the meaning and acceptability of words. Concern about the use of sexist language is part of our increased awareness that the perceived meanings of some words have changed as men and women's roles in our society have changed.

The previous paper (myojin, 2004) investigated how gender-related English terminology has shifted lately as men and women's roles in the society have changed. Overall, the results of the investigation indicate that nowadays, there seems to be quite remarkable concern about the meaning and acceptability of sexist words among both English speaking people. That is to say, as for the use of “he” and “man” as gender-neutral meanings, most of the English speaking subjects have been trying to avoid it by using several different alternatives. In addition, some compounds ending in -man such as *cameraman*, *salesman* and *stewardess* seem to be becoming dead words due to people's deliberate effort to avoid those sexist words. The findings of this study conclude that people are trying to change their language to accommodate concerns about

fairness to both sexes. As a consequence, the movement to reduce sexism in English has been becoming quite remarkable recently as men and women's roles have changed in both societies. However, no research has investigated how much EFL learners are currently aware of gender-biased language. Therefore, this study examined how much EFL learners are currently aware of gender-biased language by administering a questionnaire survey toward first-year students at a university in Japan.

## 1 Experiment

The subjects consist of 66 1st-year students (58 males and 8 females) at a technology university in Japan.

The following are the questions in the questionnaire administered toward the students.

### Question 1

"Choose an appropriate pronoun for ( ) in the following sentence from the choices below."

Each student has ( ) own computer.

1. his      2. her      3. his or her      4. their
5. I don't know

*more and more used as a generic pronoun lately?*

Only one male student answered *Yes* for this question, whereas 10 students answered *No*. In addition, as for the student who chose *Yes*; he was requested to choose one from the following:

1. *I have heard about it, but still his seems better as a generic pronoun.*
2. *I have heard about it, but I had forgot it.*
3. *Other comments:*

He chose 1. from the above.

Next, the following question was asked for those who chose *her* (0%) or *their* (48%)

Q: *Why did you choose it?*

The numbers of the students who chose the following answers are;

- *Their is the gender neutral* ..... 8/32

### Question 2

"The - man compound words such as *policeman*, *salesman* *cameraman*, and *stewardess* have been dead words these days.

Instead, *police officer*, *salesperson*, *photographer* and *flight attendants* have been replaced. Have you heard this fact before?"

## 2 Results

As for the answers for Question 1, "Which pronoun will you choose for ( ) in Sentence Each student has ( ) own computer?", the numbers and the percentage of pronouns that were chosen are as follows:

- *his* ..... 11 students (2 females) out of 66 (17%)
- *her* ..... 0%
- *his or her* ..... 21 (2 females) (32%)
- *their* ..... 32 (3 females) (48%)

Furthermore, the following question was asked for the 11 students (17%) who chose *his*:

Q: *Have you heard that his or her has been*

students

- *Their has the meaning of "general person."* ..... 2
- *Their has the meaning of "all the students."* ..... 5
- *Their has the meaning of "plural."* ..... 4
- *His excludes the meaning of "woman"* ..... 3  
(2 chose *Yes* for Q II)
- *I don't know exactly why, but I felt* ..... 4
- *Other reasons* ..... 3
- *Blank* ..... 3

The question below was asked for those who chose *his or her* (32%) :Q: *Have you heard that his or her has been more and more used as a generic pronoun lately?*

7 students (33%) (1 female) answered *Yes*, whereas 14 (67%) (1 female) selected *No*.

Furthermore, the following two questions were asked for those 7 students who chose *Yes* for the question above. And the numbers of the students who selected answers for the questions are as follows;

Q: *When did you know?*

A: •junior or senior high school ..... 6  
 •university ..... 0  
 •don't remember ..... 1

Q: *How did you know?*

A: •*His or her* was in a textbook and the teacher explained ..... 3  
 •*His or her* was in a textbook, but the teacher didn't explain, and I found it myself ..... 1  
 •*His or her* was not in a textbook, but the teacher explained.... 2  
 •*His or her* was not in a textbook, nor did the teacher explain, but I found it myself by reading a newspaper .....1

As for Question 2, that is:

The - man compound words such as *policeman, salesman, cameraman, and stewardess* have been dead words these days. Instead, *police officer, salesperson, photographer and flight attendants* have been replaced (Myojin, 2004) Q: *Have you heard this fact before?*

46 students (6 females) (70%) answered *Yes*, while 20 (2 females) (30%) chose *No*.

Moreover, the following two questions were asked for those who chose *Yes*;

Q: *When did you know?*

A:• junior or senior high school ..... 40 (87%)  
 • university ..... 0 (0%)  
 • don't remember ..... 5 (11%)  
 • no answer ..... 1 (2%)

Q: *How did you know this fact?*

• It was explained in a textbook and the teacher explained .....9/46

• It was explained in a textbook, but the teacher didn't explain, and I found it myself .....3  
 • It was not explained in a textbook, but the teacher explained.... 11  
 • It was not explained in a textbook, nor did the teacher explain, but I found it myself .....11  
 • by reading a newspaper ..... 3  
 •by TV ..... 1  
 • I don't remember ..... 2  
 • by cartoon .....1  
 • at a prep school .....1  
 • through daily life .....1  
 • common sense ..... 1  
 •Others ..... 2  
 •No answer .....10

### 3 Findings

Regarding the generic pronoun for “each”, 48% of the students chose “their”, 32% “his or her”, and 17% “his”. Nobody selected “her”. Overall, 88% of the students have never heard that “his or her” has been more and more used as a generic pronoun lately, whereas 12% of the students have heard before and most of them learned in high school. Although 67% of the students who chose “his or her” have never heard that it has been more and more used as a generic pronoun lately, they can imagine “his or her” seems preferable somehow. As mentioned earlier, almost half of all the 66 students chose “their”, but only 3 students clearly knew that “his” is a gender-biased pronoun and 10 students kind of noticed that “his” was not a good pronoun and that “their” seemed better because it includes the meaning of “gender neutral” or “general person”. However, the important thing is that almost half of the students who chose “their” didn’t seem to know that “their” is not a grammatically proper pronoun for “each”.

With respect to the –man words, as much as 70% (46) of the students have heard; the - man compound words such as *policeman*, *salesman*, *cameraman*, and *stewardess* have been dead words these days. Instead, *police officer*, *salesperson*, *photographer* and *flight attendants* have been replaced. Regarding the students who have heard the fact, 87% of them knew the fact in junior or senior high school. 43% of them said that this was because their teachers explained it. 26% mentioned that it was explained in the textbook. 24% of the students found it by themselves in various ways such as newspaper, TV and cartoon.

#### 4 Conclusion

Overall, 88% of all the students have never heard "his or her" has been more and more used as a generic pronoun lately. Although 67% of them have never heard the fact above, they were quite sensitive about the gender language. That is, 67% of the students who chose "his or her" have never heard, but they are likely to imagine "his or her" seems preferable. In addition, 13 out of 32 who chose "their" felt in some ways that "his" did not seem to be an appropriate pronoun for "each".

As for the –man words, unlike the results of the survey about the generic pronoun, as many as 70% of the students have heard that the -man words have been replaced by other words recently. 87% of them learned it in high school; in details, about half of them were explained by their teachers. In conclusion, as for the - man words, our EFL students were much more aware that they were gender-biased terms than "his or her". This seems to be because they have more chances to get the knowledge through their daily lives; for example, it is likely nowadays that we tend to avoid words like "stewardess" even in Japanese.

It is true that recently many English-speaking countries have passed strict law prohibiting

discrimination based on gender. As a consequence, quite a few organizations, especially, academic organizations prohibit gender-biased language in their writing and public speaking (Ferguson, 2004). Therefore, it seems very important for us EFL teachers to equip our students with ways to avoid sexist language by teaching it in EFL classrooms. If teachers don't teach gender-biased language in class at all, their students may not be aware of it and may even use it in their careers of English-speaking settings. Then, it is possible that their offensive manners like that may not be acceptable, and may even end up reducing their chances of succeeding in the settings. Thus, in order to enable our students to skillfully interact in authentic situations, as EFL teachers, we should have a responsibility to equip our students with an understanding of gender-biased language and appropriate substitutions for such language.

#### References

- Erlich, S., and King, R. (1998). 'Abortion :Beyond the Personhood Mrgument'. In J.Saul, *Feminism*,. New York: Oxford University Press.
- Ferguson, C. (2004). Sexist Language Persists in the EFL Classroom. English Teaching Forum Online.
- Moulton, J. (1981). 'The Myth of the Neutral "Man"'. In Vetterling-Braggin (1981::100-115).
- Myojin, C. (2004). Language and Gender: Shift in English and Japanese Terminology. The 34th Annual Convention of the Communication Association of Japan
- Saul, J. (2003). *Feminism*,. New York: Oxford University Press.

# Metaphorical use in lay and expert cancer discourse and conception of cancer in Thai culture

Worawanna Petchkij

Chulalongkorn University and the Royal Golden Jubilee, Thailand Research Fund

worawanna.p@gmail.com

## Abstract

Metaphors in everyday language are linguistic evidence showing us how native speakers in a culture conceptualize abstract concepts or experiences. Metaphors are also rich in Thai cancer discourse. They are used as communicative tools when patients with cancer and doctors talk about having cancer which is a complex experience that is difficult to understand and share verbally with others. This research investigates how Thai cancer patients and physicians conceptualize the disease through their metaphorical expressions. Conceptual metaphor theory (Lakoff and Johnson, 1980) and the relations between culture and metaphor suggested by Kovecses (2002, 2005) were employed. The data were cancer metaphors in both spoken and written cancer discourses produced by Thai patients with cancer and physicians. The results illustrate some conceptual metaphors that are different from what previous works found in western culture. Some show how Buddhist and Thai agricultural culture influence their thought and metaphorical use. Moreover, some conceptual metaphors also show different perceptions between lays who have direct experience of cancer and experts who experience it scientifically.

## Keywords

conceptual metaphor, cancer discourse, Thai, lay and expert

## Introduction

According to Lakoff and Johnson (1980), metaphors are abundant in our everyday language. They are used systematically, conventionally and unconsciously by native speakers and their system reflects the way we understand more abstract concepts in term of other more concrete ones. In other words, some words from a more concrete domain e.g. verbs and nouns occur systematically without our awareness when we talk about another more abstract concept. For example, words from the domain of war always occur in natural speech when native speakers of English talk about argument.

These linguistic realizations are pervasive in English and their system reflects conceptual metaphor “Argument is war” in cognition of English speakers.

Metaphors are rich in medical and illness discourse as well. They are useful when physicians communicate about diseases and medicine (Fleischman, 1989; Ross, 1989) and when patients try to communicate to others about their subjective understanding of illness (Radley, 1993; Gibbs and Frank, 2002). The most common metaphor in medical discourse is war or military metaphor (Fleischman, 1989) while multiple metaphors are language tools when patients talk about their illness (Gibbs and Franks, 2002).

As metaphorical use is related to native speakers’ experience and culture (Lakoff and Johnson, 1980; Kovecses, 2002), this research paper found that some cancer metaphors in Thai cancer discourse reflect conceptual metaphors about cancer that are different from western culture. In addition, there are also some differences between patients with cancer and physicians who have different experiences and levels of knowledge about cancer.

## 1 Methodology

This research finds out conceptual metaphors from metaphorical use by patients and physicians in Thai cancer discourses. The variations of cancer conceptions between these two groups and between Thai and western cancer metaphors are also examined.

### 1.1 Data collection

Cancer metaphors used by Thai patients with cancer and physicians were drawn from both of their spoken and written cancer discourses. Spoken cancer discourses were collected at three hospitals in Bangkok. This research was approved by the ethic committees of each hospital. Some ethical documents and consent forms were also shown to both patients and physicians to ask for their

permission.

### 1.1.1 Spoken cancer metaphors

The researcher interviewed 70 patients with cancer. They are patients who were waiting for their doctors at five different out-patient departments (OPD), namely oncology, radiology, gynecology, orthopedic and surgical. 51 of them were females and 19 are males. Their ages range between 18-71 years old and they suffer from different kinds of cancer. The researcher conducted a tape-recorded interview. The open-ended questions adapted from Gibbs and Frank (2002) were used to encourage patients to talk freely about their cancer experiences. Each interview lasted between 1.46 - 46 minutes; the total number of interview time is 628 minutes.

Some spoken cancer metaphors were also drawn from Doctor-patient conversation. 9 doctors agreed to be involved in this research. They are oncologist (1), surgeon (3), orthopedist (2), radiologist (2), and gynecologist (1). They allowed the researcher to record their conversation with cancer patients with MP3-recorder in their examining rooms. To get natural language data, the researcher did not stay in the room while recording.

The sound recorded were transcribed into Thai written language. Spoken cancer metaphors were then drawn from transcribed patients' interview discourses and doctor-patient conversation.

### 1.1.2 Written cancer metaphors

Patients' written cancer metaphors were drawn from 26 pocket books concerning cancer experiences written by Thai cancer patients. Some were also drawn from 27 articles about cancer experience of different Thai cancer patients in *Chivajit*, a famous alternative medicine magazine in Thailand. Physicians' metaphors were collected from 34 pocket books concerning cancer and its treatments and 36 cancer articles from a health magazine, *Klai-mor*, written by Thai doctor themselves.

To avoid perspectives influenced from other cultures, translations from other languages are excluded from this study. In addition, these physicians are both modern and alternative medicine doctors. Patients included those undergo modern treatment and alternative medicine. Therefore, their metaphorical uses represent understanding of cancer from both modern and alternative standpoints.

## 1.2 Data analysis

According to Lakoff and Johnson (1980) and Gibbs (1994), the researcher analyzes the collected data by using these two main criteria. First, conceptual metaphors are reflected from metaphorical

expressions about cancer that are systematic and conventional. Second, metaphorical expressions include metaphors that are creative and less conventional but are influenced by the underlying conceptual metaphors in the first criterion.

When problems in identifying certain metaphorical expressions or indicating their conventionality occur, the researcher consult meanings in a dictionary and the use of these meanings in a corpus as suggested by Semino, Heywood, and Short (2004). The dictionary consulted in this research is the Thai National Dictionary (1999) and the corpus is Thai Concordance provided by Department of Linguistics, Chulalongkorn University.

Conceptual metaphors found in patients' and physicians' cancer discourses are compared between themselves and between Thai and English discourses. Influences of culture on cancer metaphors are also examined by using the idea of variations of metaphors in a culture and among cultures suggested by Kovecses (2002, 2005)

## 2 Lays' and Experts' conceptions of cancer in Thai culture

Cancer metaphors in Thai cancer discourses produced by Thai patients with cancer and physicians are words from six main domains which are germ, weed, animal, disaster, friend, and general object. They are used as metaphors systematically and conventionally in Thai cancer discourses and reflect six cancer conceptual metaphors shared in the cognition of patients and physician as shown in Table 1.

Table1: Thai Lays' and experts' conceptions of cancer

Conceptual metaphors	Patients with cancer	Physician
Cancer is a germ	•	•
Cancer is a weed	•	•
Cancer is a bad animal	•	•
Cancer is a disaster	•	•
Cancer is a friend	•	×
Cancer is an alien object	•	•

According to table 1, these six conceptual metaphors demonstrate similar ways of cancer conceptualization among patients and physicians. Both of them talk about cancer in terms of a germ, a weed, a bad animal, a disaster, and an alien object. Only friend metaphors are mostly found in patients' language not physicians'

Germ metaphors are found in the context about emergence of cancer in patients' bodies.

Some verbs from germ domain like to infect, to culture, to incubate, to resist to a medicine and to spread are pervasive. Cancer is compared to infection and cancer cells are like germs that cause infection without patients' notice. Weed metaphors are mostly used in the context concerning tumors. Cancer is unwanted weeds that grow up on their organ. Cancer cells are seeds that can be brought from its tree to other places or to other organs.

Animal metaphors are frequently found in the context regarding metastasis. In this context, cancer is compared with a fierce animal that can move from a place to another by itself. Metastasis cancer cells are bad animals that bite human tissue and organs as their food. Disaster metaphors are found in the context about symptoms. As cancer always shows its symptoms when patients are in last stages of cancer, it is compared with disaster that comes quickly and destroys people's lives. Symptoms are warning signals before a disaster and patients are victims who cannot escape or survive.

Object metaphors are mostly found in the context of screening and treatment where cancer is compared with an alien object that can be put in and taken out of human bodies. They are alien object that need to be removed or taken out. In addition, we need to seek, search or find them because they are hidden somewhere in our bodies.

Friend metaphors are mostly found in patients' language not physicians. They occur in the context about chronicity of the disease. As a chronic disease, cancer is a friend who patients have to live with until the end of their lives. However, cancer can be both a good friend and unsolicited one who live in patients' bodies. Patients' bodies are houses where they live together. Patients are the house owners who want to live peacefully with a good guest who will not destroy their houses and disturb their daily lives.

These metaphors are abundant in physicians' language as well. As experts, physicians employ the same group of metaphors which reflect the same conceptions of cancer, except friend metaphor. Physicians who know well that cancer is not an infectious disease use germ metaphors in the context of emergence and metastasis of cancer. They also communicate about metastasis by using animal metaphors.

Weed, disaster and object metaphors are employed in the same contexts as in patients' language. Friend metaphors are rarely used. However, it is interesting that friend metaphors are found in the language of physicians with cancer. These physicians employ friend metaphors in the context of chronicity as other patients do.

### 3. Thai and Western conceptions of cancer

Some metaphors may be universal and reflect the same conceptual metaphors among cultures. However, many metaphors are based on cultural considerations and cognitive process of various kinds. In addition, many are non universal and they show variation among languages and cultures. Kovecses (2005:4),

In English cancer discourses, war metaphors are common among physicians and patients in western culture. (Gibbs and Franks, 2002; Gwyn, 1999; Penson et al, 2004; Reisfield and Wilson, 2004; Teucher, 2003) In their cognition, cancer is compared with an enemy that both of them need to fight or destroy. In Thai culture, it is found that cancer is compared with a friend whom patients have to live with until the end of their life. It is a good example of how thinking of cancer in a positive way is possible in Buddhist culture like Thailand where Buddhism is the professed religion of more than ninety percent of all Thais.

Semino, Heywood, and Short (2004) studied cancer metaphors in doctor-patients conversation from a hospital in the northern part of England. They found some interesting animal metaphors that are related to western culture. These metaphors reflect two conceptual metaphors, Cancer is a horse and Cancer is a hibernating animal. As animal is a common source domain that can be found in many languages (Kovecses, 2002), animal metaphors are also found in Thai cancer discourse. However, Thai patients and physicians talk about cancer in term of a fierce or bad animal in general, not any specific kind of animal. There are no words from the domain of hibernating animal which is a typical animal in cold area as western countries.

Weed metaphors is another cancer metaphor that seem to be unique to Thai culture which is basically an agrarian society. Plant is another common source domain mentioned by Kovecses (2002) too but it is not mentioned in previous research on cancer metaphor. In Thai cancer discourse, words from plant domain are pervasive as Thailand is abundantly blessed with natural resource and trees. Thais are familiar with the nature of plant around them as Thailand is located in a tropical zone where trees, and weeds can shoot or grow almost everywhere around them.

### 4. Discussion

Conceptual metaphors found in this study demonstrate shared understanding of cancer among Thai physicians and patients. They communicate about cancer by using the same groups or systems of metaphors. These metaphors reflect the same ways they conceptualize cancer concept, so that is why metaphorical communication among them is understandable or possible.

For patients, these conceptual metaphors may illustrate how they understand an unfamiliar complex concept, cancer, in terms other more concrete ones. However, for physicians, these conceptual metaphors might be evidences of how metaphorical understanding occurs with familiar or known concept as cancer is not a new or unfamiliar concept for these experts. In other words, they give us an idea that metaphorical understanding in our cognition might be able to occur with recognized but difficult to communicate concepts. Physicians may find it difficult to communicate this complex disease to patients and these metaphors are very handy for them to communicate with lays. Besides understanding cancer in scientific ways, physicians therefore also conceptualize cancer through other concepts that are sharable or communicable.

Friend metaphors show how direct physical experiences of cancer influence patients' thoughts as they are found only in the language of patients and physicians with cancer. The way this metaphor is rarely found in general physicians who do not have direct experience of having cancer may be evidence supporting how direct experience influence our metaphors and thought.

Thai culture seems to have much influence on friend metaphor as well as it is not mentioned in previous research. As most Thai people are Buddhists, thinking about cancer in terms of a friend which is a positive way of thinking might have some influences from Buddhist belief such as beliefs about the truth of life. For example, most Buddhists know about the process of one's life, birth, aging, illness, and death. Everyone will surely go through this process. No one could escape death. Cancer is in the fourth step of this process, illness. Therefore they can do nothing but accept the disease as any other kinds of life threatening may happen anywhere and anytime.

## 5. Conclusion

Cancer metaphors in Thai cancer discourses reflect six conceptual metaphors about cancer in Thai culture which are cancer is a germ, cancer is a weed, cancer is a bad animal, cancer is disaster, cancer is a friend and cancer is an alien object. All of them are found in the language of patients and physicians with cancer. Physicians have the same conceptions of cancer except friend concept. In addition, conceptions of cancer in terms of weed and friend seem to be unique to Thai culture as they are not found in previous research.

## References

- Fleischman, Suzanne. 2003. *Language and Medicine*. In *The handbook of Discourse Analysis*. Edited by Deborah Schiffrin, Deborah Tannen, and Heidi E. Hamilton. P. 470 -502 USA : Blackwell
- Gibbs R. (1994) *The Poetics of Minds: Figurative Thought, Language, and Understanding*. New York: Cambridge University Press.
- Gibbs, R. and Franks, H. (2002) "Embodied Metaphor in Women's narratives about Their Experiences with Cancer" *Health Communication* 14 (2).139-165.
- Gwyn, R. (1999) "Captain of My own ship: Metaphor and the Discourse of Chronic Illness". In Lynne Cameron and Grayham Low (ed.). *Researching and Applying Metaphor*. United Kingdom: Cambridge university press. 203-220
- Kovecses, Z. (2002). *Metaphor: A Practical Introduction*. Oxford: Oxford University Press.
- Kovecses, Zoltan. 2005. *Metaphor in Culture: Universal and Variation*. Cambridge: Cambridge University Press.
- Lakoff, G. and Johnson, M. (1980). *Metaphors We Live by*. Chicago: University of Chicago Press.
- Penson, R.T et al. (2004). "Cancer as Metaphor". *The Oncologist* 9. 708-716.
- Radley, A. (1993). "The role of metaphor in adjustment to chronic illness". In Alan Radley (ed.). *Worlds of Illness*, London: Routledge. 109 - 123
- Reisfield, G.M. and George R.W. (2004). "Use of Metaphor in the Discourse on Cancer". *Journal of Clinical Oncology* 22:19. 4024 – 4027
- Ross, J.W. (1989). "The Militarization of Disease: Do We Really Want a War on AIDS?". *Sounding* 72:1. 39-58
- Semino, E, Heywood, J and Mick, Sh. (2004). "Methodological Problems in the Analysis of Metaphors in A Corpus of Conversations about Cancer". *Journal of Pragmatics* 36.1271-1294.
- Teucher, U. (2003). "The Therapeutic Psychopoetics of Cancer Metaphors: Challenges in Interdisciplinary". *History of Intellectual Culture* 3:1.1-15.
- Weiss, M. (1997). "Signifying the Pandemics: Metaphors of AIDS, Cancer, and Heart Disease." *Medical Anthropology Quarterly* 11:4.456-476.

## Acknowledgement

The researcher would like to express deepest gratitude to her advisor, Assoc. Prof. Dr. Krisadawan Hongladarom, for her support and guidance throughout the research. She acknowledges with thanks the financial support from the Thailand Research Fund through the Royal Golden Jubilee Ph.D. Program (Grant No. PHD/ 01 12/2548) to Ms. Worawanna and Assoc. Prof. Dr. Krisadawan. She would also like to extend her thanks to all Thai patients with cancer and physicians who participated in this research, including the authors of the written sources.

# Emotions, Visual Rhetoric, and Pragmatic Inferencing in Campaigning Discourse

Vincent Tao-Hsun Chang

Graduate Institute of Linguistics, National Chengchi University

Vince.T.Chang@gmail.com, 94555001@nccu.edu.tw

## Abstract

The audience's comprehension and interpretation in multimodal communication are investigated by looking into four ads for the Beijing Olympics 2008. The ads creatively crafted the Games' logo with brilliant pictures and colour symbolism to attract the audience's attention, to initiate cognitive contextual effects, viz. pun, irony, metaphor and humour, and to perform various communicative functions thereafter. They convey too the prominent Chinese civilisation, exhibiting the aesthetic and cultural consumption, artistic commodification and identity politics reified within the symbolic domain of campaigning discourse.

## Keywords

Advertising discourse, Chinese civilisation, emotions, implicature, multimodality, pragmatic inference, social cognition, visual rhetoric

## 1 Introduction

The studies on advertising language (in Taiwan) are mainly from marketing and advertising, and the researches from linguistic perspectives are minority, especially as compared to those in the United States, UK and Europe. This paper thus presents a qualitative study on four ads for the Olympics 2008, intending to explore the research aims listed below. Furthermore, as mentioned by van Leeuwen (2002) regarding visual communication, e.g. advertising, no matter in commercials or political campaigns, in print form or through other visual/sound effects, is no longer limited to directly selling products and has provided a site where emergent various interesting and appealing forms. One of the most significant features is that longer texts and/or storytelling in a literary style have largely been applied by many institutions and corporations intending in particular to construct certain cultural image, such as the advertiser/Mainland China discussed in this study. It weakly communicates a lot of messages, inviting the audience members as potential consumers to join the stories and

interpretation process, and helps on the other hand to shape itself as a cultural landmark.

## 2 Research questions

This study attempts to investigate popular culture expressed in advertising discourse, and explores the ideology and power relations shaped in the ads by society and culture, and their possible effects on society and culture, trying to explain and render plausible interpretations to the following research questions:

- (1) *What does the role of audience play in the interpretation process in media communication? How can she process and interpret the communicated information from advertiser?*
- (2) *How are diverse linguistic/communicative strategies and advertising appeals used to affect and persuade the audience, as the communicator/advertiser/enterprise attempts to build certain corporate image?*
- (3) *What cognitive effects with emotions could be perceived and inferred by different audience through the visual rhetorical strategies and 'communicative acts' (van Leeuwen 2002) employed in institutional discourse/specialised communication?*
- (4) *What is the social meaning accompanying or behind language use? What competing ideologies and changing cultural values can be seen from advertising discourse, and further to shape social cognition?*
- (5) *How will the audience construct their cultural understanding through popular culture, e.g. advertising?*

## 3 Theoretical framework

As mentioned above, the studies on language in ads in Taiwan from linguistic viewpoint are still few, while from multimodal perspective, they are far less. So this current study proposes a qualitative study on four captions for the Olympics 2008, in attempts to examine the research issues outlined above.

Analytical framework is based upon Sperber and Wilson's *Relevance Theory* (1986/1995). Since proposed in 1986 by them, the implications of Relevance on communication and cognition have been widely applied in various fields, including literature, linguistics, psychology, political language, language education, advertising and film studies, etc. (Sperber & Wilson, 1986/1995: 255-6, 259-60). In *Relevance*, Sperber & Wilson define 'optimal relevance' from the hearer's processing ability in terms of her<sup>1</sup> processing effort and those possible effects:

An utterance, on a given interpretation, is optimally relevant if and only if:

- (a) it achieves enough effects to be worth the hearer's attention;
- (b) it puts the hearer to no gratuitous effort in achieving those effects. (Smith & Wilson, 1992: 5)

Also they define *Communicative Principle of Relevance* and *criterion of consistency with the principle of relevance* as follows:

#### *Communicative Principle of Relevance*

Every act of communication communicates a presumption of its own optimal relevance. (Sperber & Wilson, 1995: 158)

#### *Criterion of consistency with the principle of relevance*

An utterance, on a given interpretation, is consistent with the principle of relevance if and only if the speaker might rationally have expected it to be optimally relevant to the hearer on that interpretation. (Smith & Wilson, 1992: 6)

In *Relevance*, Sperber & Wilson mainly discussed the examples from word, phrase and sentence levels. While as Blakemore (1992: 165-6) analysed the example of irony, 'irony is not always restricted to a couple of lines or a single utterance. In many cases it extends over a whole poem or story...' in which ironical effects are achieved through processing a whole text or, a level of a more global organisation, i.e. *macrostructure* (van Dijk, 1977: 130). Thus, since the data for analysis in this study are in single integrated textual structures that I apply *Relevance Theory* to analyse larger units, from macrostructure level to see the relevance reached by the audience and diverse pragmatic functions.

Aside from pragmatic analysis, on the other hand, and also suggested by Sperber & Wilson

(1995: 279), this paper further examines the sociocultural contexts to see the inseparable relationship between language use and social function, the major concern of the critical linguists (Fairclough, 1989, 1995a, 1995b):

It is an approach which is, I believe, suitable for use in the sort of research into social and cultural change [...] What in particular makes it suitable for such work is that it foregrounds links between social practice and language, and the systematic investigation of connections between the nature of social processes and properties of language texts.....It is moreover a 'critical' approach to discourse analysis in the sense that it sets out to make visible through analysis, and to criticize, connections between properties of texts and social processes and relations (ideologies, power relations) which are generally not obvious to people who produce and interpret those texts, and whose effectiveness depends upon this opacity. (Fairclough, 1995a: 96-7)

## 4 Data analysis

In this section, I am dealing with the four ads for the Olympics 2008 respectively in the light of their picture by picture to deliver cognitive effects. Let us see caption (A) first:

[A]



From caption (A), the slogan reads as *Aoyun re, rebian jingcheng!* 'Olympic Fever Heats All of Beijing,' as the other campaigns do, simply and plainly expressing the atmosphere of hot and the mood of expectation. The picture, aside from proposing also echoes this feeling of heat. The sweet round dumplings are in a bowl with boiling water and vapour [sweetness is happiness!]. Besides, the dumplings signifies reunion and harmony for

<sup>1</sup> In this paper I also consider the audience to be female, and the addresser to be male, in accord with Tanaka (1994: xv).

Chinese people/culture, the round bowl also support this point (round, pronounced as *yuan*, is a homophone of *tuanyuan* “reunion”). The dumplings, consumed during Chinese New Year, further enhance the meaning of reunion. Moreover, the symbol *bat* (*bianfu* in Mandarin Chinese) represents good luck and is symbolic of good fortune for its pronunciation *fuqi*.

Being examined from another level, this caption strongly implicates that ‘all the athletes and audiences from everywhere will reunion in year 2008,’ a happy reunion with expectation and somehow though competition. Also, this caption weakly conveys and implicates the following series of messages:

(1)

We expect that all of you could participate in this world event.

Come to join us to experience the Chinese culture.

Chinese culture is far-reaching and longstanding.

Come and try our passion and friendship. (We are friendly!)<sup>2</sup>

Come to the Game, visit our beautiful scenery, try the Chinese gastronomy, taste our local specialty (, feel a different China).

Look, the hotness of Olympics is spreading all the Beijing City.

The Olympic fever represents our passion and expectation.

[...]

By using this multimodal communicative style involving visual rhetoric, the speaker could leave a wide space of interpretation to the audience (accompanying larger responsibility at the same time). Moreover, he would not communicate only the strong implicatures in (1), for in that case he might directly express in another way to save her processing efforts, as shown in (2) and (3):

(2) We are waiting for you to participate in the Olympic Games. See you in 2008.

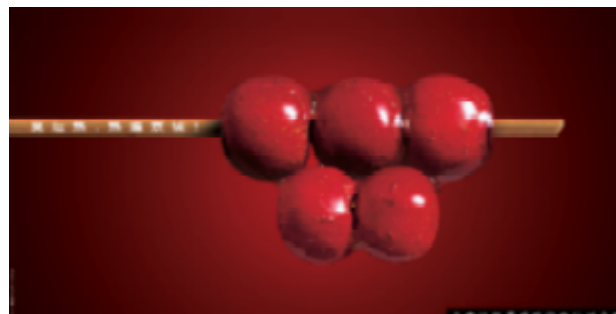
(3) Come to join us, play the Game, and experience the Chinese traditional culture.

As compared to (A), except the propositional content, the audience could not infer any weak implicatures from (2) or (3).

Sweet dumplings in (A) are one of the most prominent characteristics of Chinese culture and social norms/routines; in (B), another typical sweet

snack — *tanghulu* (plum with hot syrup coating) comes into play, which is not so formal nor being positioned higher class as sweet dumplings’ inherent meanings:

[B]



Likewise, caption (B) conveys that mood of heat through the hot syrup and red colour. The hot syrup may also convey a state of liquid to implicate that the Olympics 2008 are fluent, moving and active, with the help of the chopstick, straightforward and speedy. Also, this ad (weakly) communicates the following possible implicatures:

(4)

It is (sociostereotypically) regarded that *tanghulu* is not so high taste (as sweet dumplings).

(However,) we still wanna introduce it to you.

Snacks (usually) signify sort of close relations with short distance.

Beijing City and the host authority of Olympic Games 2008 welcome various groups of people to join the Game, and play the Game.

[...]

As mentioned above, not merely the formal icons are employed to promote the image for the host institution, Beijing City, but the foodstand on the street with much lower social meaning/connotation is applied in the campaign to introduce and promote Chinese culture and tradition as well. This can well shorten social distance amongst the different social groups/strata, from lower level, much lower social strata to higher level, the social elite [As compared to (A), this one] further (fairly weakly) implicates sociometaphorically that:

(5)

Chinese culture embraces different kinds of elements and flavours.

Mainland China is wide-ranging and very generous.

We welcome various kinds of people, including, say, street vendor, to the Game.

By doing so, we are considerate and generous.

A considerate and generous government is worthy of respect and support.

[...]

<sup>2</sup> The words in parentheses represent further optional explanation for implicatures and are relative to ‘degree,’ whereas a square bracket with dots inside [...] means the indeterminacy of implicatures.

Interestingly ironical here is that, if the audience could infer and derive these weak implicatures, it's really incorporates politics into exercises. A salient humorous and ironic effect is best depicted in (C):

[C]



The slogan reads “*Zhongguoren Jiu Ai Gao Yundong Ma! 2008 Jian*” which is used as a pun and can be paraphrasably as follows:

(6a) The Chinese people just love to make political movements. See you in 2008.

(6b) The Chinese people love sports so much. See you in 2008.

Along with the photo of Chairman *Mao Zedong*, the audience is not hard to associate (6a) with those drastic and great political movements happened decades ago. But after further processing against this ad for the Olympics, the audience will recover the contextual effects and abandon the previous inference. As for one meaning or two meanings being activated simultaneously by the speaker's intention (cf. Tanaka, 1994), [especially considering the economic reform led/directed by *Teng Xiaoping*] it still needs further examination.

The problem of human rights in Mainland China has long been questioned internationally, notoriously lying in the event of *Tiananmen Square* on June 4<sup>th</sup> 1989. For Beijing authority and Chinese people, to get the host right is not very easy and being viewed as an opportunity to refresh and renew the negative impression by this world event. This can be evidenced from Chairman *Mao*'s friendly smile on the same *Tiananmen Square* with red wall, warm-hearted, informing the audience that that time has passed. Thus, in this way, such a

communicative strategy provides multidimensional thoughts for audience. Moreover, the irony and humour soften the presumably political tension and unpleasant atmosphere. According to Sperber & Wilson, in recognition of irony (Sperber & Wilson, 1995; Wilson & Sperber, 1992; Blakemore, 1992), which is a major source to trigger humour, it demands, first, an echoic element or an attributed thought; second, the speaker's attitude of mockery or rejection. From the text and context here, we can retrieve an echoic element and attributed thought:

(7)

The speaker believes that the Chinese is supposed to be sport-loving, like those movements happened decades ago.

The speaker believes that the Chinese people had made many political movements long time ago.

And the speaker's attitude of mockery can be expressed below:

(8) Funny how it is for the people simply compare (serious and formal) political movements with physical exercises.

With the trait of short phrases and hence ease of memory, the audience could easily memorise the language elements, the slogans, which are still widely used in advertising (in view of economic consideration, saving money for advertisers). Such advertising style of baldly demonstrating political claims and commercial appeals, however, may not necessarily be longer stayed in, or deeply influence, the audience's mind. It possibly makes one getting tired and causes 'memory fatigue' during, say, a period of election campaign, though 'language attacks' among candidates and competitors are quite common, legitimate, reasonable and plausible within media, especially in the election arena. Multimodality, on the other hand, embracing ample semiotic codes, seems to be not so drastic but could be more humanistic and wide-angled. Here, I am not eliminating or underestimating the possible effects of advertising slogans, as illustrated in the related studies, but rather more emphasising on the richer and diverse contextual effects created by multimodal communication. Different triggers will bring the audience different effects in different contexts at different times and places. As for what communicative strategies are more 'effective' than others, it could be approached through two avenues: the first falls into a quantitative study to reach empirical findings; the other resorts to time factor, leaving the challenge of effectiveness to the dynamic relations between the speaker and hearer. The more the audience could accept, the longer and

more effectively the ads would work. Effectiveness results in social continuity, ineffectiveness leads to social change. This is a vital contribution to, and underlies, the dialectical relationship (Fairclough, 1989, 1995a, 1995b) between sociocultural value structures and social practice, and is one of the core elements in the process of socialisation.

The last caption (D) adopts an actress of Peking (Beijing) opera to echo the fever of the Beijing Olympic Games:

[D]



Based on our common sense, the opera dress itself is hot and heavy enough, though in this picture there is not showing the lower part of the actress. Colour symbolism via the ruby red on the head and the make-up around the eyes is manipulated to strengthen the fever and heat. With the eyes of concentration and self-confidence, again, it tells the audience that ‘We are ready (for the champion),’ ‘Be careful, we are not weak’...and so forth.

Both the male and female athletes, actors and actresses, can well be famous and outstanding. However, as illustrated in this series of captions for the Olympics, the advertiser only employed female roles (little girl and opera actress) to express the ambition and confidence of China towards the Game, giving rise to somewhat femininity to soften the severe competition. With only the lion to show a certain degree of masculinity. As for Chairman Mao, designed by NIKE, could be viewed as inter-/national figure to gain much more attention of the audience due to his fame and powerful influence.

Multimodal communication is much more interesting than baldly promote and sell products, services and images. By exploiting multimodality to shelter stronger measures and claims in political campaign and in commercial advertising to shift voter and buyer’s concern, leading directly the

audience to a certain domain and line of thinking, is another linguistic strategy, or termed ‘macro-level hedging.’ The audience will be led to another domain of interpretation supplied by the visual dimension or the story lines, and is at least encouraged not to focus the seller’s motive. This is evidenced by the data analysed above. Vaguely communicating the main gist of Chinese culture, friends’ rapport and other ideological loadings, these ads place special emphasis on higher involvement (*renqingwei* [人情味] in Mandarin Chinese) and harmony, as compared to Western cultures, that are among the most prominent values within Chinese community and generally outweigh many others.

## 5 Concluding remarks

From the analyses in this study, we can know that the audience is active, rather than passive, in the interpretation process, where ‘relevance’ is crucial to the text/discourse interpretability. The audience would achieve her optimal relevance through the interaction of linguistic form, the shared cognitive environment of communicator and audience, and the ‘criterion of consistency with the principle of relevance’ (Wilson & Sperber, 1992). As demonstrated in this study, those strong and weak implicatures are resulted from the author’s (my) higher involvement and active processing. They are indeterminate and unlimited for inferencing, and might be processed further, as long as one is willing to process and thinks it is worthwhile, i.e. she will be rewarded with additional contextual effects, which would outweigh her processing efforts. Different audiences will receive different readings and interpretations, the ‘absolute levels’ are diverse; also, even the same audience will receive different readings and interpretations under different degrees of involvement and circumstances. However, they would reach at any rate the ‘optimal relevance’ for themselves.

Communicative strategies enrich interpersonal connection, which themselves enrich life. Mainland China released the four ads incorporating various images into the pictures, namely sweet dumpling in a round bowl, *tanghulu* ‘plum with syrup coating,’ smiling Chairman Mao at Tiananmen Square, and an actress of Peking opera. Instead of claiming ‘come n’ join us to expect and welcome the Olympics 2008,’ they impress the audience through these four vivid pictures. If you couldn’t spare time to watch the Games, it doesn’t matter, but do remember we host the Beijing Olympics 2008 and enjoy the captions we designed for you. Sweet dumplings (*Yuanxiao* or *Tangyuan*) is inviting you that Beijing Olympics is like Chinese family reunion. An informal local snack is telling you the

Olympic fever is bloody running. A friendly smile of Chairman Mao is surprising you our politeness and friendship with the sense of humour yet without political tension. An opera actress with profession is informing you that 'we are ready for those competitions,' and is promoting another traditional cultural asset. Besides, colour symbolism is another important strategy to convey the sense of fever and ambition — red, along with yellow and blue, are the principal ones used in these captions. The advertiser (Mainland China) intends not only to reshape those long-established stereotypes, thus shorten the social distance between the East and the West as well as between consumer and advertiser in an amiable way, but to build specific cultural image through globalisation and might thus gain itself competitive edges.

Promotion of products, services and images is the ultimate goal of advertising. The selling motive, however, as illustrated in this study, has largely been hidden and melted by persuasion through multimodal communication in an ever increasingly prominent trend. Macro-level hedging via social metaphor is also initiated intertextually to shift consumers' buying concern. This distance leaves a wider space to the audience's imagining, as many other rhetorical strategies do, leaving an indeterminate domain of interpretation to the audience. While nowadays there are many speech contexts much freer, more open and direct than before, the competition between direct and indirect persuasion (e.g. multimodality, storytelling) is then a challenge to language users and time. This (competition) reflects one important property in mass media, the tools of ideological representation, being able to embrace competing forces. Advertising, lending itself as a symbolic domain of dominant ideologies (Lull 1995), incorporates fashion-driven discourse by reflecting social cognition and cultural patterns, it also invites the audience members as potential consumers to recognize these prominent values. It is socioculturally shaped but it also constitute sociocultural cognition, in ways that may be transformative as well as reproductive (Fairclough 1995b: 34), thus maintaining the dialectical relationship between social structures and social practice. This functional and critical linguistic study of advertising texts heightens sensitivity to language and its (mis)uses, poses a preliminary stage to cross-cultural communication and interdisciplinary studies, which themselves constitute promising research issues to overcome the simplistic account of Eastern vs. Western cultures, and provides useful insights into a multitude of issues, including symbolic power, social stereotypes, changes in social trends and

attitudes, group and ethnic identity, cultural identity and iconicity, verbal art, and gender. Comparative studies in terms of diachronic perspective and different genres are also research areas worthy of further investigation.

## References

- Blakemore, D. (1992). *Understanding utterances: An introduction to pragmatics*. Oxford: Blackwell.
- Bourdieu, P. (1991). *Language and symbolic power*. Cambridge: Polity.
- Davidson, M. (1992). *The consumerist manifesto: Advertising in postmodern times*. London: Routledge.
- Fairclough, N. (1989). *Language and power*. London: Longman.
- Fairclough, N. (1995a). *Critical discourse analysis: The critical study of language*. London: Longman.
- Fairclough, N. (1995b). *Media discourse*. London: Edward Arnold.
- Gibbs, R. W. Jr. (1993). Process and products in making sense of tropes. In Andrew Ortony (Ed.), *Metaphor and thought*, 2<sup>nd</sup> ed. (pp. 252-276). New York: Cambridge.
- Lull, J. (1995). *Media, communication, culture: A global approach*. Cambridge: Polity.
- Pilkington, A. (1992). Poetic effects. *Lingua* 87: 29-51.
- Simpson, P. (1993). *Language, ideology and point of view*. London: Routledge.
- Smith, N. and Wilson, D. (1992). Introduction. *Lingua* 87: 1-10.
- Sperber, D. & Wilson, D. (1995). *Relevance: Communication and cognition*. 2<sup>nd</sup> ed. Oxford: Blackwell.
- Tanaka, K. (1994). *Advertising language: A pragmatic approach to advertisements in Britain and Japan*. London: Routledge.
- van Dijk, T. A. (1977). *Text and context: Explorations in the semantics and pragmatics of discourse*. London: Longman.
- van Dijk, T. A. (1994). Discourse and cognition in society. In David Crowley and David Mitchell (Eds.), *Communication Theory Today* (pp.107-126). Cambridge: Polity.
- van Leeuwen, T. (2002). Ten reasons why linguists should not ignore visual communication. GURT 2002 Plenary speech. Georgetown University, U.S.A.
- Wilson, D. & Sperber, D. (1992). On verbal irony. *Lingua* 87: 53-76.
- Wilson, D. & Sperber, D. (1993). Linguistic form and relevance. *Lingua* 90: 1-25.

# Developing ESP Course Material for International Economics Students

Tony Cripps

Faculty of Economics, Ritsumeikan University

cripps@ec.ritsumei.ac.jp

## Abstract

This paper provides a chronology of the development of English for Specific Purposes (ESP) course material for international economics students at Ritsumeikan University. At first, a brief outline will be provided of the events that led to the establishment of a small material development team for the Communication and Writing 3 (CW3) course for second-year students at the university. The paper then explicates the rationale which underpins the creation of the material and focuses on the assessment of the material by one group of CW3 teachers (n=7) and the subsequent revision of the material that culminated in the creation of a new CW3 text.

The latter section of the paper examines the assessment of the revised CW3 material by two classes of international economics students (n=32). Data was collected through an online questionnaire and semi-structured interviews. The paper concludes with a brief introduction to the on-going development of an online Autonomy Resource Center (ARC), which is being created to support the CW3 course and other courses at the university. It is hoped that this paper will provide insight into the practical aspects of ESP material development.

## Keywords

Autonomy, English for Specific Purposes, Material Development

## Introduction

Selecting appropriate course material for students can be frustrating. For teachers who have English for Specific Purposes (ESP) classes this often seems an impossible task due to the dearth of suitable material which matches their students' needs. This paper outlines how ESP material for international economics students at Ritsumeikan University was developed. The author focuses on the material creation process, the reaction of teachers and students to the material, and how the ESP material has been refined.

## 1 Background

In 2006 an international economics course was established at Ritsumeikan University's Biwako Campus in Shiga prefecture, Japan. Apart from fundamental economics subjects taught in Japanese, students also take English classes. In the first two semesters first-year students take the following skills-based English courses; Reading 1 & 2; CALL (Computer Assisted Language Learning) 1 & 2; Listening 1 & 2; Communication and Writing 1 & 2 (see Figure 1). For these courses teachers usually use standard skills-based texts, or skills-based software (CALL 1 & 2). However, the focus shifts to English for Specific Purposes in the second year.

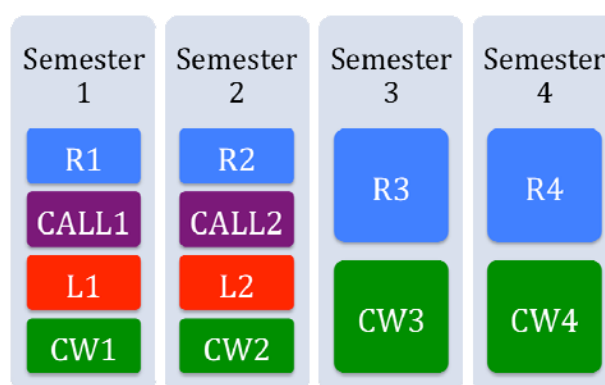


Figure 1: Four-semester overview

After conducting a needs analysis with faculty members, a small team of English teachers within the economics faculty developed the ESP material used for the Reading 3 & 4 (R3 & R4) and Communication and Writing 3 & 4 courses (CW3 & CW4). This paper focuses on the material created for the CW3 course.

### 1.1 Fostering Autonomy

An important consideration when creating the ESP material was the need to foster learner autonomy. Benson's assertion that "those who lack autonomy can develop it given 'appropriate support'"

(2001:2) is fully supported by the material development team.

Kohonen (1992) and Little (1996) stress that autonomy implies 'interdependence' and thus, learner autonomy is fostered not in isolation, but through social interaction. Therefore, a major consideration when designing our ESP material was the need to create innovative material which requires students to work closely together.

## 2 Material Development

It is very difficult to find existing material which exactly fits the needs of our students (Johns, 1990). After researching the existing ESP material for students of economics it was felt that none of the material addressed our students' specific needs and material development teams for the R3/R4 and CW3/CW4 courses were formed.

The use of material which is authentic and promotes self-direction are two of the chief characteristics of an ESP course (Carter, 1983; Gatehouse, 2001). The need to develop our students' English ability through exposure to authentic material and the promotion of learner autonomy are two important elements which underpin the CW3 course. Our aim is to help students develop the English skills which they will need in the future, such as the ability to write academic essays (or reports) on topics related to international economics, and to present their research findings through formal academic presentations.

Since no suitable ESP material could be found, the CW3 team designed and created a CW3 text which includes 20 academic workshops. These workshops focus on academic writing and academic presentations. The academic writing workshops provide guidance on topics such as essay structure, in-text citation and referencing, and proofreading. In the presentation workshops students are given advice on topics such as body language, eye contact, presenting research data, and answering questions.

In addition to the CW3 text, supplementary material was created for the course. Three economics professors agreed to be videoed while they answered questions on fundamental aspects of economics such as supply and demand, opportunity cost, and the causes of economic inequality. A DVD lecture series on economics was also added to the course to give students the chance to explore authentic material. Additional support material was created and piloted during the first semester. After

the pilot stage was completed the CW3 material was distributed to all CW3 teachers. This included the CW3 text, course notes, professors' DVD comments, DVD scripts, comprehension and discussion exercises, and listening tests. After using the CW3 material for one semester, teachers were asked for their feedback. Their responses allowed the material development team to refine the CW3 material that had been created.

## 3 Teacher Feedback

Teachers who used the first CW3 text (n=7) were asked for their comments on the text and the support material at the end of the course. Their comments were collected using an online survey site (SurveyMonkey.com) and were followed up by informal interviews. Their responses were analyzed using a grounded theory approach (see Radnor, 2002).

An overwhelming majority of teachers (n=6) provided positive comments about the CW3 text and other material. However, they felt that the DVD lectures were more useful than the videos of the economics professors. The positive aspects of the supplementary material are summarized below:

- Useful DVD material
- Video material provided salient examples of good body language
- Discussion questions facilitated class debate
- Students appreciated seeing their professors speak English
- Interesting to teach

However, there was one teacher who said that the material was not useful at all and he stated that the material was far too difficult. His comments serve to highlight the fact that material creation cannot please everyone.

## 4 Initial Refinements

After analyzing the teachers' feedback, the CW3 text was rewritten and the support material was revised. The main refinements focused on making material for the CW3 text which was easier to use in class. Several practice listening tests were also created to give students listening support for both the formal DVD lectures and the TOEIC test.

## 5 Student Feedback

The refinements above culminated in the creation of a new CW3 text and support material. At the end of the 2008 spring semester, two classes of

international economics students (n=32) were asked for their feedback on the CW3 course and the material they used. The data was collected through an online questionnaire and semi-structured interviews. Analysis was conducted using a grounded theory approach.

The students' comments about the course were generally very positive. In contrast to the English teachers, many students stated that they liked to see their Japanese professors speaking about economic themes in English as they found this extremely motivating. Their feedback on the CW3 textbook was also positive. Typical comments were as follows:

- *"The text is very useful for me to use English in future."*
- *"I can understand more than textbook of CW1 and CW2."*

There was, however, mixed reaction to the formal DVD lecture material:

- *"The DVD lecture for note-taking is a little difficult."*
- *"The DVD lecture is too difficult!"*
- *"We have to listen to the DVD carefully. It is nice for us."*
- *"The DVD lecture was very difficult, but I realize that I should study English more through that DVD, so that was good, I think."*

It should be noted that the students with comparatively higher TOEIC scores were far more positive about the CW3 material (especially the DVD lectures) than the students with comparatively lower TOEIC scores. The students' comments resulted in further refinement of the CW3 text and support material.

## 6 Building the ARC

Following the feedback from the teachers and the students it was clear that further material development was necessary. Many students commented that they want to review the DVD material in their own time after class. They also commented that they want to have online access to additional material which would help them improve their listening ability and their test-taking techniques.

The feedback from the teachers and students acted as the catalyst for the creation of an online autonomy supportive resource center (ARC). The

conceptual framework for the ARC became a reality when the ARC research team was awarded a research grant by the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT). With the help of MEXT funding, course material was created and designed using the Moodle platform. Moodle was chosen because of its underlying social-constructivist philosophy (Moodle, 2009: para. 1). The Moodle platform allows course administrators and designers to control and develop the online environment. Through the ARC, CW3 students now have access to video content, listening tests, essay guidelines, TOEIC practice tests, and other online self-study material. Figure 2 provides an overview of the available ESP material for the CW3 course.

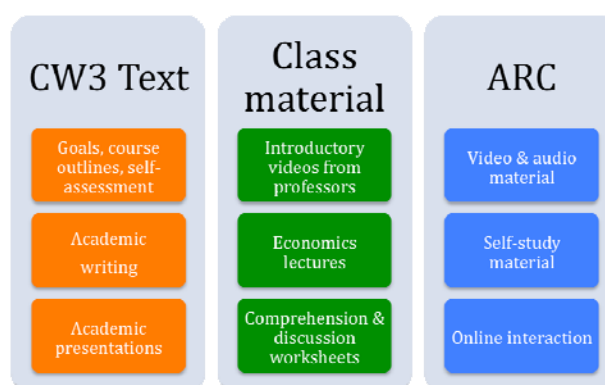


Figure 2: ESP Material Overview

The ARC is also being woven into two new CALL courses for first-year students. Through the integration of the ARC into these new courses students are being encouraged to take part in simultaneous online discussions and asynchronous online forums. It is hoped that these developments will lead to the creation of new ESP material aimed at providing further online support and, perhaps, even for a totally independent online course.

## 7 Conclusion

The development of ESP material requires time, funding, and dedicated colleagues. In order to create effective and engaging resources, educators need to be aware of the limitations of the material they create. Therefore, reassessment of the material through consultation with teachers and students is essential. By involving the stakeholders in the feedback process, material developers can provide innovative ESP material that teachers and students will use. It must be borne in mind that effective material development is an ongoing process. Considering the changing technological and pedagogical landscapes, we should constantly be

challenging ourselves to create even better material.

## References

- Benson, P. (2001). *Teaching and researching autonomy in language learning*. London: Longman.
- Carter, D. (1983). Some propositions about ESP. *The ESP Journal*, 2, 131-137.
- Gatehouse, K. (2001). Key Issues in English for Specific Purposes (ESP) Curriculum Development. *The Internet TESL Journal*, Vol. VII, No. 10. Retrieved May 2, 2009 <http://www.iteslj.org/>
- Johns, G. (1990). ESP textbooks: Do they really exist? *English for Specific Purposes*, 9, 89-93.
- Kohonen, V. (1992). Experiential language learning: second language learning as cooperative learner education. In Nunan, D. (ed.), *Collaborative language learning and teaching* (pp.14-39). Cambridge: CUP.
- Little, D. (1996). Freedom to learn and compulsion to interact: promoting learner autonomy through the use of information systems and information technologies. In Pemberton, R. et al (eds.), *Taking Control: Autonomy in Language Learning* (pp. 203-218). Hong Kong: HKUP.
- Moodle. (n.d.). *Philosophy*. Retrieved May 2, 2009. <http://docs.moodle.org/en/Philosophy>
- Radnor, H. (2002). *Researching your professional practice*. Buckingham: OUP.

## Effects of Review Activities on EFL Learning

Hsiao-fang Cheng<sup>1</sup> & Nina Chiulan Lin<sup>2</sup>

<sup>1</sup>Language Center, National United University, <sup>2</sup>Department of Applied Foreign Languages, Takming University of Science and Technology

<sup>1</sup>hfcheng@nuu.edu.tw; <sup>2</sup>ninachiulan@yahoo.com.tw

### Abstract

The utmost goal of foreign language instruction is aimed at helping the learner master the language. At the same time the learner shall become equipped with linguistic, pragmatic and social-linguistic competence. This study was done to explore if review activities in EFL classes should be mandatory for learners to learn the new knowledge. One hundred and fifty non-English majors participated in this study. They were divided into three groups, two as experimental groups and one as the control group. Group A received written review treatments whereas Group B received oral review treatments once every three weeks. Group C, the control group, did not receive any review activities after the lessons were taught. At the end of the semester, all the participants received a post-test on what they had learnt. The results showed that the participants in Group A (written review) and Group B (oral review), performed substantially better than the control group. A significant difference in gained scores was found between the control group and the experimental groups, indicating that providing review opportunities is both necessary and helpful for achieving the learning outcomes, i.e., the retention of new knowledge. Drawing on the finding, some pedagogical suggestions were made.

**Key words:** review activities, linguistic

competence, learning outcome

### 1. Introduction

Foreign language instruction is directed at helping students to gain linguistic competence in terms of phonology, vocabulary, syntax, and functions of the target language, and above all, pragmatic and socio-linguistic competence which is the use of the new knowledge in real-world communication (Hughes, 2003). In many foreign language learning settings, the target language is unfortunately not widely used in the community where the learning takes place. Foreign language learners receive instruction and are given opportunities to practice only in the classroom. Usually classroom instruction focuses on the forms and structures of the language within the context of communicative interaction. Indisputably, instruction becomes a series of cooperative activities between teachers and students (Yu, 2002). Teacher-student interaction promotes the acquisition and internalization of new language forms that are characteristic of EFL learning activities.

In many EFL teaching situations, teachers tend to make professional decisions to ensure that learning takes place effectively and are expected to take control of the teaching processes in their classrooms (Nunan & Lamb, 1996). Learners seem to learn what has been taught; however, it is not

certain that they learn everything that they are taught. Even though something is taught or made available for them to absorb, it does not mean that the learners will digest it right away. This is especially true for under-motivated non-English majors who study English as a subject instead of a language. They have not had a chance to use English for real-life communication; how can they progress to a level deeper than what has been presented to them? Therefore, what in-class activities should teachers provide to help students master the target language? Drawing on Output Hypothesis (Swain & Lapkin, 1995), Interaction Hypothesis (Long, 1996), and Noticing Hypothesis (Schmidt, 2001), the researchers attempt to investigate if it is mandatory to provide in-class review activities, either in written or spoken form, to promote learner's learning outcome, and most important of all, what kind of review opportunities contribute the most to student's learning performance.

## 2. Literature Review

Output Hypothesis (Swain & Lapkin, 1995) provides a detailed description of each component stage and depicts the interrelated and dynamic processes of language acquisition. The model proposes five stages during the learners' conversion of input to output. Noticing Hypothesis (Schmidt, 2001) stressed language awareness and language restructuring. In the course of interaction, learners learn to identify the differences between the input and their own output. Through practice they convert the noticed items into acquired items and restructure them in oral or written tasks -- they use them productively. In the early stages of L2 development, ideas are separate units and the only link between them is their common reference to a given topic (Ellis, 1997). The result of restructuring is often reflected in what is known as U-shaped

behavior (Gas & Selinker, 2001). U-shaped behavior refers to three stages of linguistic use. Error frequency follows a low-high-low pattern as development takes place. In the earliest stage, a learner produces some linguistic forms that conform to target-like norms (i.e., is error-free). At Stage 2, a learner appears to lose what was known at Stage 1. The linguistic behavior at Stage 2 deviates from his native language form. Stage 3 looks just like Stage 1 in that there is again correct target language usage (Ellis, 1990; Gass & Selinder, 2001).

Clark and Clark (1977) stated that exact wordings are only stored for very short periods of time and unless actively rehearsed, are lost very quickly. Schematization of knowledge occurs when learners are required to think about the material at a deeper level in terms of how concepts inter-relate (Herbert & Burt, 2003). Drawing on the U-shaped learning behavior, learners need to review the newly absorbed knowledge so as to convert it into generalized conceptual knowledge. Studies showed that good EFL learners pay close attention to the formal properties of the target language (Ellis, 1994), attend to form and monitor one's own and others' speech (Reiss, 1985), as well as pay attention to how the target language is used (Stevick, 1989). It is obvious that the more the new knowledge is schematized, the better the output results. However, what is the best method to enhance the schematization of new knowledge? Ding (2007) conducted an interview with three university English majors and documented text memorization and imitation as the most effective methods of learning English. In Ding's study the participants are highly motivated in learning English, so they are willing to memorize the texts. For those under-motivated learners, whether rote learning is an appropriate and interesting method to motivate learning is worth pondering. Therefore,

the researchers, speaking from their personal experience, propose an alternative method -- in-class review activities after lecturing—with the prospect of promoting learner's motivation and learning outcome. It is hoped that such activities can be good ways to improve learners' comprehension and memory.

The following research questions will be addressed.

- (1) Is it mandatory to provide review activities to enhance students' learning outcome?
- (2) Which review activities, written or oral, contribute to the excessive acquisition of new knowledge?
- (3) What were students' attitudes toward and their perceptions on the provision of review opportunities after oral teaching?

### 3. Method

#### 3.1 Participants

One hundred and fifty non-English majors participated in this study. Their average age was 19.2 years. All of them had 8.1 years of learning English as a foreign language at school in Taiwan. None of them had been to English speaking countries.

#### 3.2 Instrument and Design

Three contextualized dialogues on interviews and three short stories were utilized as teaching materials. A TOEIC practice test serving as pre-test and a criterion-referenced test based on the teaching materials were utilized to gather data for the investigation.

The participants were divided into three groups based on their academic majors. The pretest results showed that the three groups in the study were indeed homogeneous in terms of English language ability ( $df=2$ ,  $F=0.000$ ,  $p>0.05$ ).

#### 3.3 Procedures

The three groups received a pre-test at the beginning of the treatment and a post-test. According to Wu and Lin (2000), the interval between each review cannot be too long or too short. It was decided that Group A and B be given a review opportunity once every three weeks. However, the control Group C did not receive any review activities. Group A received a written review; Group B received an oral review. In the thirteenth week all the subjects took the post-test without prior notice.

Paired *t*-tests were used to analyze the results of a 'before' and 'after' treatment. After the post-test all the subjects completed a post-exercise survey concerning their attitudes toward and perceptions of the treatment.

## 4. Results and Discussion

### 4.1 Statistical Results and Discussion

A one-way ANOVA analysis was applied to see whether a significant statistical difference was found in the post-test scores among the three groups. A significant statistical difference was found in the post-test scores among the three groups ( $F=12.933$ ,  $df=2$ ,  $p=0.000$ ), indicating that the treatment did influence subjects' learning outcome.

The Scheffe test revealed that after the treatment there was a significant statistical difference between Group A and Group C as well as between Group B and Group C. These findings indicated that review opportunities, whether oral or written, did facilitate learner's memory and output. Group A and Group B performed similarly, and that each of these groups performed better than Group C. This difference could be a result of practice effects for Group A and B as the material taught was repeated once every three weeks for a total of three times before the post-test. On the contrary, Group C did not have any review opportunities.

The mean score of Group A was somewhat higher than that of Group B, suggesting that written review activities appeared to facilitate the internalization of language or memory slightly better than oral review activities.

Results of the paired *t*-tests showed that the gained scores in Group A and B between pre-test and post-test was statistically significant, indicating that learners in Group A and B truly progressed after the treatment. However, for Group C, no significant difference was found between the learners' pre-test and post-test mean score ( $p=0.162$ ), indicating that learners in Group C did not improve their English ability to a significant level after oral teaching. It can be inferred that giving instruction with subsequent review activities is necessary in EFL classes.

To sum up, subjects receiving three review activities did perform better in the post achievement test, and it is concluded that review opportunities did facilitate the subjects' memory and output. Research Question One and Two can be considered answered with the following statement: Review opportunities promoted subjects' memory and output and written review activities appeared to facilitate better achievement. A possible explanation is that learners in the experimental group had chances at regular intervals to recall what had been taught, so new knowledge was less likely to be forgotten. It is likely that knowledge that had been reviewed several times could have been processed at a more profound level and then converted into intake and integrated into the learners' developing foreign/second language learning system. Therefore, learners could successfully retrieve the needed knowledge in their output, leading to a higher performance outcome.

It was also found that the performance result of Group A showed evidence of U-shaped behavior. The average score of the first review test was 83;

the second review test 63; the third test 62 and the post achievement test 71. A backsliding phenomenon, i.e. deviated output, was found in the second and third review tests. But in the post-achievement test, the members of this group tended to produce more target-like responses as they had in the first test.

#### **4.2 Results of the Post-exercise Survey**

Question 1 was "Do you think providing review activities in class after oral teaching help you remember what had been taught? Why?" All the participants in the experimental groups stated that reviewing the learned material regularly and in different formats did help them a lot. Ninety-six percent of the participants admitted that they would not have reviewed the lesson on their own after instruction; therefore, the review opportunities helped them recall the newly acquired knowledge when they were about to forget. They further mentioned that they could remember the content better than the syntactic structures and vocabulary. To master the linguistic forms, they needed to practice several times.

All the participants in the control group admitted that they would not have reviewed the lesson on their own if they did not have a quiz on what they had learned. They agreed that they could learn well under a little pressure and believed that review opportunities if given would have promoted their learning outcome.

Question 2 was "Why didn't you perform substantially better in the post-test?" All participants in the control group stated that they could not remember very well what had been taught in class because the interval between instruction and post achievement test had been too long. Ninety-six percent of them mentioned that they had only a vague idea about the content and they could not remember the detailed information, the syntactic structure or vocabulary, because they

did not review what was taught immediately after class.

## 5. Conclusion

This study has provided some relative answers to improving learning outcomes. The results revealed that reviewing what had been taught at regular intervals after instruction promoted learners' performance and written review facilitated better recall of information taught than oral review. No review activities available after the instruction led to poor retention of information taught. It was found that U-shaped learning behavior and practice effects did occur in the process of learning. Generally speaking, learners had positive attitudes toward review activities after oral teaching and preferred reviewing in different formats.

## 6. Pedagogical Implications

The practical goal of this study is to investigate whether providing review opportunities in class is likely to enhance learning outcomes. Drawing on the results of this study, some pedagogical implications are proposed.

First, it is necessary to provide review opportunities after lecturing. Second, providing review opportunities in different formats avoids boredom and promotes some measure of success and motivation. Third, the most appropriate number of reviews ranges from two to three times, time permitting. If time is not available, as U-shaped behavior suggests, reviewing the new information at least once is mandatory.

There was only one group of Chinese technical college students participating in this study. The results presented in this paper cannot be taken as conclusive and definitely do not allow for generalization.

## 7. References

- Brindley, G. & Slatyer, H. (2002). Exploring task difficulty in ESL listening assessment. *Language Testing*, 19(4), 369-394.
- Clark, H., & Clark, E. (1977). *Psychology and language: An introduction to psycholinguistics*. New York: Harcourt Brace Jovanovich.
- Ding, Y. (2007). Text memorization and imitation: The practices of successful Chinese learners of English. *System*, 35, 271-280.
- Ellis, R. (1990). *Understanding second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (1997). *SLA research and language teaching*. Oxford: Oxford University Press.
- Gass, S. M., & Selinger, L. (2001). *Second language acquisition: An introductory course*. Mahwah, N. J.: L. Erlbaum Associates.
- Herbert, D. M. B., & Burt, J. S. (2003). The effects of different review opportunities on schematization of knowledge. *Learning and Instruction*, 13, 73-92.
- Hughes, A. (2003). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Long, M.H. (1996). The role of the linguistic environment in second language acquisition. In Ritchie, W.C., Bhatia, T.K. (Eds.), *Handbook of Second Language Acquisition*. San Diego, CA: Academic Press. Pp.413-468.
- Nunan, D., & Lamb, C. (1996). *The self-directed teacher: Managing the learning process*. Cambridge: Cambridge University Press.
- Reiss, M. (1985). The good language learner:

Another look. *Canadian Modern Language Review*, 41, 511-523.

Schmidt, R. (2001). Attention. In Robinson, P. (Ed.), *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press. pp.3-32.

Stevick, E. (1989). *Success with foreign languages: Seven who achieved it and what worked for them*. Hamel Hempstead: Prentice Hall.

Swain, M., Lapkin, S. (1995). Problems in output

and the cognitive processes they generate :

A step towards second language learning.

*Applied Linguistics* (16), 371-391.

Yu, M. N. (2002). *Educational test and evaluation: Achievement tests and instructional evaluation*. Taipei: Psychology Publishing Co. [In Chinese]

Wu, W. Y., & Lin, C. H. (2000). *Business research methods*. Taipei: Hua-tai Cultural Center. [In Chinese]

# A study on the stress experienced by native English teachers in Korea

Jongbok Lee

School of Education, Mokwon University

jblee@mokwon.ac.kr

## Abstract

The purpose of this study is to find out the factors of teaching stresses from native English teachers who are working at a Korean university. These days, college English education managed by inviting native English teachers becomes popular. In this situation, it is important to figure out how to increase efficiency of English education conducted by native English teachers. Therefore, as one of the methods to manage native English teachers, this survey focused on the teaching stresses by taking a survey toward 26 native English teachers who engaged in N university.

## Introduction

This study set its focus on stress of native English teachers when they could feel in their teaching. To investigate the stress, this study took a survey toward 26 native English teachers working in N university in Korea. Main purposes of this study can be suggested as following;

1. To find out amount of teaching stress of native English teachers and divide the stress into factors by factor analysis.
2. To divide the participants into groups by their personal background and figure out trend of stress

## 1. Literature Review

### 1.1. Current Situation of the inviting native English teacher program.

Native English teachers have been working in Korea since 1976. At that time, about 200 young English speakers from USA, who were affiliated

with America Peace Corps, were dispatched to middle and high schools all around Korea. After that, through the Fulbright program of Korean-American Educational Commission, about 30 English natives with certain qualifications (under 30 with B.A. degree) had dispatched to schools in Korea every year since 1992. In 1997, with the Plan of Inviting & Utilizing Native English Assistant Teachers was established by Ministry of Education, the program has conducted actively since 1998.

Looking into the number of invited Native teachers in Korea per each year, the number of teachers has rapidly soared until 1997. As Korea began to be under management of International Monetary Fund(IMF) from 1997, the number of the natives decreased. After recovering economical crisis, as demands for improving English communication abilities began to elevate, the number of native English teachers was also increased and it reached to 2,294, officially registered in EPIK(Ministry Of Education & Human Resources Development, 2007).

However, beyond officially registered Native English teachers, there were more English natives who individually came to Korea with the purpose of teaching English conversation. In 2006 the number of unregistered native English teachers was 12,560, which accounted for 91.5% of all the natives with E-2 Visa and about 10,000 more than that listed in EPIK at the same year, (Ministry Of Education & Human Resources Development, 2006a; 2006b). In addition, there were 29,263 natives without E-2 Visa, who staying in Korea with the purpose of teaching English conversation and that was about 16,000 more than the number of aforementioned data. Thus, much more English natives were expected to work in Korea as English teachers. Recently, there were 35,457 native English teachers who dwelled in Korea in 2008(Korea Immigration Service, 2008).

## 1.2 Current Situation of managing native English Teachers

Looking into the management guideline for native English teachers, which provided by EPIK(Ministry Of Education & Human Resources Development, 2005), while it mainly mentioned external conditions such as invitees' qualification, residence, salary and working hours for the program manager, it didn't contain internal information such as difficulties which native English teachers would face with in their daily life and stress when they would feel in teaching. Referring preceding studies dealing with teachers' stress (Blasé, 1986; Koh & Kim, 1994; Mark & Richard, 1991a, 1991b; Shin, 2004), they presented that teachers' stress affected teaching and relationships between students and teachers. Based on the result, it can be elicited that native English teachers' stress is not only one of the main factors which influence their teaching, but also related to the efficiency of the inviting native English teacher program.

## 1.3 Teaching Stress

Teaching stress, change of teacher's psychological stability with physiological change, happens as a result of teacher's activities. Also, it means the stress when teachers recognize outside demands as threats to their psychological stability or their self-esteem.

Kyriacou and Sutcliffe(1978) suggested that teachers feel stress due to students' motivation to study, school, work, students' insincere attitude to teachers, students' disorder, working conditions, troubles with colleagues, quality of classes. Also, Litt and Turk(1985) pointed out that when teachers encountered a problem threatening their present situations, they feel stress and negative feelings if it requires much more ability than what they originally have.

Considering aforementioned results, teachers' stress influences negative effects not only on themselves, but also on the relationships between students and teachers(Kelly & Berthelson, 1995; Kim & Shin, 2000).

Since studies on native English teachers' stress

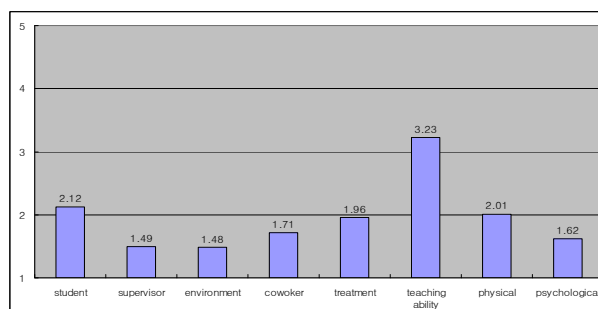
have not conducted widely, it is hard to apply the results of studies mentioned above, which mainly set their focuses on kindergarten teachers, to this study. However, main elements, which influence teachers' stress, suggested in the studies, Students' Attitude(Kyriacou & Sutcliffe, 1978) and Great Pressure for Teaching and Extra Work(Korean Federation of Teachers' Association), are supposed to be commonly applied to all the people working in education field regardless of the age of their students.

Therefore, based on the results supporting that teachers' stress affects huge influence on teachers' teaching and the relationships with students, it can be possibly predictable that native English teachers' stress also can influence their teaching.

## 2. Methodology

A survey was carried out with 26 native English teachers engaged in N university in Korea to find out stress factors in their teaching. The questionnaire was consisted of questions selected from preceding studies(Caplan, Cobb & French, 1975; Davison & Cooper, 1983; Kim, 1991; Kim, 1997) dealing with teachers' stresses and working stresses. Questions about teaching stress were collected from Kim(1991)'s and Kim(1997)'s studies. Questions related to teaching were selected among questions asking working stresses gathered from studies of Caplan et al. (1975) and Davison and Copper(1983). After deleting similar questions from all the collected questions, the questionnaire was consisted of 56 questions.

## 3. Results and Discussion



<Figure 1> teaching stresses of native English teachers

<Figure 1>, which presents results of the survey, shows eight categories of teaching stresses(student, supervisor, environment, coworker, treatment, teaching ability, physical and psychological) sorted by factor analysis. Average of all the questions was 1.49 which means that participants hardly felt stress in their teaching. Reason behind this could be found in participants' unique background. All of them worked in same university managed by Christian foundation. So, they are all Christians and most of them had strong passion in their teaching. Consequently, the reasons why the participants felt little stress in their teaching can be found in their religious background and working motivation.

Participants felt the highest stress in the teaching ability factor(3.23). It means that most of participants recognized their insufficiency and were not satisfied with their teaching ability. Quick and Quick(1984) insisted that the acceptable stress can facilitate tension and efficiency in working and take the positive role. Based on that, stress of teaching ability can be sorted as a positive stress. Therefore participant didn't feel much stress in their teaching except a positive stress, teaching ability' which encourage participants' tension and motivation in their teaching.

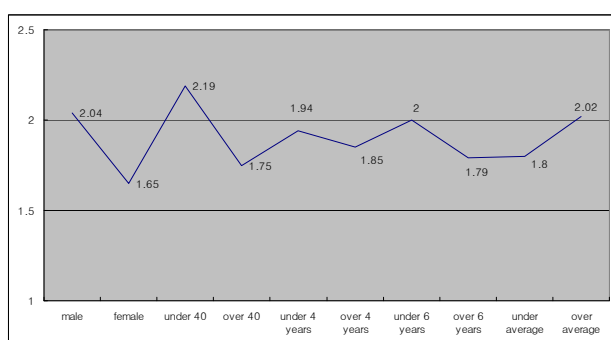


Figure 2 differences of teaching stress tendency according to personal backgrounds

To find out differences of the teaching stress tendency according to personal backgrounds, the participants were divided into 2 groups in accordance with 5 categories(sex, age, period of residence in Korea, teaching experience, salary).

As each Category was consisted of two groups, the number of groups was 10. <Figure 2> presents the results of the teaching stress tendency. Criteria

for age, period of residence, and salary was set based on average of the participants(age=40; residence period=4 years; salary=2,100thousand). In case of teaching experience, since there were huge differences among all of the participants, accumulative number of the participants was applied as a criterion.

Comparing the teaching stress of each group, the group of male(sex), under 40(age), under 4 years(period of residence), and under 6 years(teaching experience) felt higher stress than other groups. In case of salary, someone who got more salary felt more stress because salary was in direct proportion with the time of class.

#### 4. Conclusion

Summing up the results, categories of teaching stress could be sorted as student, supervisor, working environment, coworker, treatment, teaching ability, physical and psychological factors by factor analysis. Among the categories, stress of teaching ability showed the highest point(2.68) and stress of treatment showed the lowest point(1.48). Analyzing differences of stress tendency between two groups in each category, which were divided

by the participants' personal background, stress tendency became higher when participants were younger, had shorter residence period, less teaching experience. Only salary was in indirect proportion with the stress.

Inviting native English teacher program requires a lot of costs to provide them with habitation, airline fee and monthly salary. Therefore, to produce decent output of this program, it is needed to carry out studies focused on how to raise efficiency and quality of this program. Based on the results of this study, this researcher set two research goals to expand domain of this study.

1. Investigating satisfaction of native English teachers in their teaching.
2. Suggesting an effective management system for native English teachers.

#### References

- Blase, J. J. (1986). A qualitative analysis of source of teacher stress: consequences for performance. *American Educational Research Journal*, 23(1), 13-40.
- Caplan, R. D., Cobb, S., & French, J. R. P. (1975). Relationships of cessation of smoking with job stress, personality and social support. *Journal of Applied Psychology*, 60(2), 211-219.
- Choi, Y. S. (2008). Practical tips for EFL native teachers and their employers in Korea. *English Language & Literature Teaching*, 14(3), 29-47.
- Davison, M., & Cooper, C. (1983). *Stress and Woman Manager*. Oxford; Martin Robert Son & Company Ltd.
- Kelly, A. L., & Berthelson, D. A. (1995). Preschool teachers' experiences of stress. *Teaching and Teacher Education*, 11, 345-357.
- Kim, B., & Shin, H. Y. (2000). An Exploratory Study on Job Stress of Day Care Teacher. *International Journal of Early Childhood Education*, 20, 253-276.
- Kim, J. I. (1997). *Effects on working stress and the results influenced by negative changes and social expectation*. Unpublished doctoral dissertation. Chung-Ang University. Seoul.
- Kim, J. W. (1991). *Relation between teacher's teaching working stress and psychological and physical symptoms*. Unpublished doctoral dissertation. Chung-Ang University. Seoul.
- Koh, H. H., & Kim, J. H. (1994). *A study on working stress of day-care teachers*. Seoul: Baeyoungsa.
- Korea Immigration Service. (2008). *Monthly report of statistics*, 2, 2008.
- Kyriacou, C., & Sutcliffe, J. (1978). Teacher stress prevalence: Source and symptoms. *British Journal of Educational Psychology*, 48, 159-167.
- Litt, M. D., & Turk, D. C. (1985). Sources of stress and dissatisfaction in experienced high school teachers. *Journal of Education Research*, 78, 178-185.
- Mark, G. B., & Richard, J. R. (1991a). Stress in teaching: a study of occupational stress and its determinants, job satisfaction and career commitment among primary schoolteachers. *Education Psychology*, 11(1), 59-75.
- Mark, G. B., & Richard, J. R. (1991b). Occupational stress and satisfaction in teaching. *British Educational Research Journal*, 17(3), 263-281.
- Ministry of Education & Human Resources Development. (2005). *2005 Plan of EPIK(English Program in Korea)*. Teacher Development Planning Division of Ministry of Education & Human Resources Development, 1, 2005.
- Ministry of Education & Human Resources Development. (2006a). *Plan of EPIK and present situation*. Data for Parliamentary Inspection of the Administration.
- Ministry of Education & Human Resources Development. (2006b). *Annual report for E-2 degree*. Data for Parliamentary Inspection of the Administration.
- Ministry of Education & Human Resources Development. (2007). *2007 Plan of EPIK(English Program in Korea)*. Teacher Development Planning Division of Ministry of Education & Human Resources Development, 1, 2007.
- Shin, H. Y. (2004). *Effects of teachers' job stress and belief of efficacy on the quality of teachers' interaction behaviors in child care*. Unpublished doctoral dissertation. Yon-Sei University. Seoul.

# Partitive-self constructions: Lexico-grammatical resources for constructing sociocultural and individual aspects of the self

Kam-yiu S. Pang<sup>1</sup>

<sup>1</sup>Faculty of Social Sciences & Humanities, University of Macau

kpan@umac.mo

## Abstract

Pang (2005) examines a cluster of English constructions under the rubric *partitive-self constructions*. The primary function of these constructions is to single out a certain aspect of a person's self for comment, evaluation, and/or reflection, whether by the person herself or by others. Two such constructions are the '*N in me V-ing*' and '*my N V-ing*' constructions. This paper argues that these constructions also fulfill their identity-construction function differently. Looking at the types of nouns which typically participate in these constructions and the interaction between the nouns and their determiners, I propose that these constructions are basically complementary. They function as resources for constructing and asserting *different* aspects of a person's self. '*N in me V-ing*' tends to assert/construct more socioculturally based aspects of a person's self, whereas '*my N V-ing*' tends to assert/construct more individual-based aspects of a person's self.

## Keywords

partitive-self constructions; self-talk; sociocultural and individual identities; Construction Grammar

## 1 Introduction

Pang (2005) identifies a cluster of constructions in English which he calls **partitive-self constructions**. Constructions being symbolic items, in that they are not meaningless syntactic patterns, but have symbolic properties inherently associated with them (Lakoff 1987; Langacker 1987; Goldberg 1995; Taylor 2002; inter alia), the primary pragma-semantic function of this

cluster of constructions is to draw attention to a certain part of a discourse participant's self (typically the speaker herself<sup>1</sup>) which the speaker considers relevant to the communication at hand. Examples are:

- (1a) It is his arrogance speaking.
- (1b) Grrrr. (Yes, that was my stomach talking.)
- (2a) That's the purist in me talking.
- (2b) That's the fangirl in you speaking.

The constructions can be broadly divided into two main types, according to the structure of the nominals, i.e., whether it consists of a possessive determiner such as *my*, *her*, etc., as in (1), or a postnominal locative with a personal pronoun such as *in me*, *in him*, etc., as in (2). These are given the rubrics of the '*my N V-ing*' and '*N in me V-ing*' constructions (Pang 2005).<sup>2</sup>

By focusing the hearer's attention on that part of the discourse participant's self (in particular the speaker herself), the speaker signals to the hearer that it is *that* aspect of the person to which he should predicate what is being said, rather than the person as a *whole*. Any evaluation the hearer might have of what is said by the discourse participant should thus be made with this "caveat" in mind. These aspects of the speaker being attended to can range from the social or professional roles which she instantiates to her emotional or physical states. The invited inferences are thus:

- the discourse only represents what the speaker thinks in the capacity of a certain social or professional persona, or as a result of being under the influence of certain physical or psychological experiences;
- by isolating the focal aspect for reflection, the speaker is indicating that she herself is

<sup>1</sup> In this paper, to avoid sexist language, I have adopted the practice of using the feminine to refer to the speaker, agent, experiencer, etc., and the masculine to refer to the hearer, patient, etc., in the case of indeterminate gender, after Sperber & Wilson (1986).

<sup>2</sup> Other constructions in the cluster having basically the same cognitive and pragma-semantic properties include those instantiated by utterances such as *The devil in me made me do it* and *It was my head that said, "Don't be silly!"*

aware of the possible ramifications of the discourse, and

- if it can be perceived as potentially disagreeable to the hearer, her awareness includes that as well.

One basic pragma-semantic function of these constructions is thus to distance the discourse participant from what is being said.

Further to Pang (2005), which discusses the overall constructional, cognitive-grammatical, and pragma-semantic properties of the construction-cluster, this paper focuses on the types of nouns which tend to participate in the two constructions, i.e., the types of nouns which are differentially *attracted* (in the sense of Stefanowitsch & Gries 2003) to the N-slots in [DET:POSS N] and [*the* N in PRO:PER]. It turns out that [DET:POSS N] tends mostly to attract nouns designating psychophysical states, e.g., *my hangover*, *your paranoia*, *my maternal instinct*, while [*the* N in PRO:PER] tends mainly to attract nouns designating sociocultural/professional roles and constructs, e.g., *the cynicism*, *the environmentalist*, *the accountant*. Based on these findings, I will argue that

- the ‘*my* N V-ing’ construction tends to serve the function of asserting/constructing an aspect of the person’s self which is more *individual-based*; whereas
- the ‘N in *me* V-ing’ construction tends to serve the function of asserting/constructing an aspect of the person’s self which is more *socioculturally based*.

That is to say, although both constructions serve the overall purposes of identity-construction and identity-enactment, they tend to do so in different domains and hence attend to different, viz. sociocultural vs individual, aspects of a person’s self. Thus, the two are basically in a quasi-complementary relation.

## 2 The data

### 2.1 Methodology of data collection

The data reported in Pang (2005) were collected from the Internet using the *Google* search engine (www.google.com). Searches were conducted on all the paradigmatic combinations of the variables in the constructions. For example, for the ‘*my* N V-ing’ construction, searches were made with strings consisting of combinations of {*this/that/these/those/it*} {*s/is/are/was/were*} {*my/your/his/her/our/their*} {*\**} {*speak-/talk-/say-/tell-/ing*}. The same procedure was used for the ‘N in *me* V-ing’ constructions. The sampling size for each was 70. For the present paper, this search was

replicated in the first half of 2009. Although content on the Internet has certainly changed since 2005, results from the current collection indicate that, though individual items might appear or disappear, the overall noun-type distribution tendencies have largely remained the same, across all relevant combinations of the variables, i.e., the possessive determiner (*my/your/his/her/our/their*) and the noun in ‘*my* N V-ing’, and the noun and the personal pronoun (*me/you/him/her/us/them*) in ‘N in *me* V-ing’.

Furthermore, a separate search using a larger sampling size of 250 each (consisting of the first 250 unduplicated hits) was conducted for the strings [*this is my* N *talking*] and [*this is the* N in *me talking*]. This is for the purpose of gaining a more comprehensive picture of the noun-type distribution of the constructions. For calculating the respective incidence of the nouns and noun-types, head noun lemmas are counted in the case of complex nominal structures, e.g., *eco-friendly side*. The first-person singular variable was chosen because it is found to be more frequently attested than the other variables (cf. Pang 2005). The ensuing discussion will thus be illustrated with 1sg examples collected from this search.

### 2.2 Operational definitions

The scheme used in Pang (2005) for classifying the types of nouns which participate in this construction-cluster denotes the following domains:

- Mental. These include psychological, emotional, and affective states, e.g., *ego*, *fear*, *intuition*, typically with a more acute sense of subjective experientiality;
- Physical. These include body parts and bodily conditions, e.g. *heart*, *head*, *sickness*, *hangover*;
- Sociocultural constructs. These include reified/personified constructs such as *Irish*, <sup>2</sup> *demon*, *Christ*, *Buddha*, and abstract constructs such as *pessimism*, *values*, *morality*, *hopelessness*;
- Sociocultural roles (which the current paper replaces with the more general term of ‘sociocultural personæ’). These include kinship, ethnic, ideological, institutional, and professional personæ, e.g. *mother*, *realist*, *cynic*, *psychiatrist*;
- Artifacts functioning as metonyms, e.g. *medication*, *suits*, *Unix*.

The current paper adds a sixth domain which pertains to nouns denoting personal experiences, history, and abilities, etc. In establishing operational definitions for these category labels, Pang (2005) notes that there is some difficulty in

categorizing nouns other than those designating physical states, physiology, and artifacts-as-metonyms, for several reasons. On the one hand, nouns designating psychological states/constructs are often ambivalent when they appear in different constructional and/or discursal contexts. For example, *pride* could be referring to an abstract construct when appearing in an utterance like *That was the pride in him talking*, but could arguably be referring to a mental state in *That's his pride talking just now; not him*. On the other hand, an abstract construct and an agent attributed with that construct present a different problem because of their mutually defining relation: “Such an agent (e.g. a *pessimist*) would have to embody, and hence instantiate, the construct (e.g. *pessimism*) to be so deemed, while conversely speaking, the

construct *pessimism* cannot be defined or conceptualized without appealing to a human agent that embodies it” (Pang 2005: 8). To resolve the issue, Pang’s (2005) categorization appeals to foregrounding as a criterion: “Where the designatum as a coherent body of belief(s) about a particular aspect of the socioculture’s worldview is foregrounded, the noun [is] considered a ‘sociocultural construct’ (e.g., *cynicism*); where the type of person as an individual intentional agent who instantiates such a construct is foregrounded, it [is] considered a ‘sociocultural persona’ (e.g., *cynic*); and where the state/experience which an individual intentional agent is undergoing is foregrounded, it [is] categorized under ‘mental’”. (Pang 2005: 8).

Table 1. Incidence of head nouns in [my N]

incidence	nouns
13	<i>heart</i>
7	<i>experience; side</i>
5	<i>gut; brain; self; instinct;</i>
4	<i>inexperience; background; bias; ignorance; ego; intuition; OCD; paranoia; memory;</i>
3	<i>hormones; butt; syndrome; ghost; upbringing; nature; subconscious; anger; logic; mind; emotion; depression; lawyer;</i>
2	<i>flesh; youth; age; degree; sense; anxiety; soul; spirit; disorder; imagination; blood; Americanness; Scorpio; eye</i>
1	<i>head; penis; pregnancy; boner; illness; body; ass; face; anesthesia; beard; sickness; stomach; dying; hangover; taste buds; MBA; roots; MFA; class; knowledge; prowess; creativity; teenybopper; birthday; persona; voice; jealousy; frustration; conscience; grief; kink; naivety; insecurity; unbelief; pettiness; skepticism; cynicism; belief; thinking; idiocy; fetish; crush; stupidity; filter; excitement; consciousness; personality; hindsight; fever; exuberance; shame; curiosity; perfectionism; neurosis; fear; addiction; trekkie; fan; fiancé; behaviorist; Daemon; fanboy; dh; realist; twin; affinity-build; student; hat; pessimist; mom; therapist; dad; fanboyism; bachelorhood; fandom; geekgasim [sic]; Capricorn; house; node; ideology; Americanism; prudishness; child; religion; worldview; practice; Te; ethics; perspective; rights; baggage; partisanship; generation; privilege; Rheumy; pouch; mask; wallet; ball; dolly</i>

### 3 Results

#### 3.1 Noun-type distribution of [my N V-ing]

All six types of nouns mentioned above (2.2) are found for this construction. Table 1 lists the head noun lemmas attested according to incidence. In terms of frequency, ‘mental’, ‘personal history/ability’ and ‘physical’ nouns predominate (31.6%, 26.8%, and 20% respectively). ‘Sociocultural construct’ and ‘sociocultural persona’ nouns are much less frequent by comparison (11.6% and 7.6% respectively), and ‘artifact-as-metonym’ nouns are rare (2.4%) (see Table 2). A Chi-square goodness-of-fit test shows the results to be highly significant ( $\chi^2=97.23$ ,  $df=5$ ,  $p<0.0001$ ).

#### 3.2 Noun-type distribution of [the N in me V-ing]

Table 3 lists the head noun lemmas according to incidence. The predominant noun-type found is ‘sociocultural persona’ nouns: as much as 92%. The other noun-types, viz. ‘sociocultural construct’, ‘mental’, ‘personal history/ability’, and ‘artifact-as-metonym’ nouns are rare by comparison (2.4%, 1.6%, 2.4%, and 1.6% respectively). No ‘physical’ nouns are attested. (Table 4). A Chi-square goodness-of-fit test shows the results to be highly significant ( $\chi^2=1022.1$ ,  $df=5$ ,  $p<0.0001$ ).

Table 2. Noun-type distribution of [*my* N]

	noun type	incidence (N=250)	percentage
A	mental	76	30.40%
B	personal history/ability/trait	67	26.80%
C	physical	51	20.40%
D	sociocultural construct	32	12.80%
E	sociocultural persona	20	8%
F	artifact-as-metonym	4	1.60%

$\chi^2 = 93.34$ ,  $df = 5$ ,  $p < 0.0001$

Table 3. Incidence of head nouns in [*the* N *in me*]

incidence	nouns
7	<i>geek</i>
6	<i>mom/mum</i>
5	<i>cynic</i>
4	<i>teacher; nerd; guy; girl;</i>
3	<i>homer; gamer; American; optimist; idealist; designer; theorist; teenager; director; engineer; fan; lawyer; pessimist</i>
2	<i>snob; fanboy; libertarian; kid; feminist; hunter; fangirl; conservative; newbie; philosopher; Nazi; therapist; boy; person; artist; linguist; New Yorker; developer; student; perfectionist; techie; writer; editor; midwest;</i>
1	<i>animator; Tableeghi; Angeleno; mechanist; filmmaker; trader; dancer; musician; misfit; anthropologist; catcher; pig; pragmatist; CEO; Indian; towel-head; producer; historian; behaviorist; poet; gearhead; proofreader; Republican; bellydancer; archivist; bitch; Pollyanna; educator; consumer; surgeon; Bakla; techno-dork/DJ; publisher; marketer; dreamer; librarian; Michigander; scout; photographer; Gabriel; rescuer; marine; spender; Scrooge; caregiver; naturopath; Canadian; buff; realist; critic; virgin; howard stern; capitalist; supporter; base; hippy; counter; chicken little; planner; Jew; fop; fanatic; Californian; commuter; singer; adult; player; Italian; tree-hugger; reader; mobile-phile; westerner; Buddhist; hand; sadist; luvvy; wife; skeptic; Tim Hortons; chauvinist; druid; whore; romantic; Gen X-er; Chinese; atheist; bartender; primal; yankee; manager; advertiser; Kathy Lee Gifford; brother; dealer; Thespian; freak; man; academic; perv; attorney; psychologist; woman; dad; fogey; gen-Y; diver; rehabilitator; shipper; major; servant; lover; Oklahoma; 21st century; Friday; peteru; PES; pvp; hipster runoff; Google; OCD; comparison; OC; monster; lady; customer; college; sloth; burn-out; ignorance</i>

Table 4. Noun-type distribution of [*the* N *in me*]

	noun type	incidence (N=250)	percentage
A	mental	4	1.60%
B	personal history/ability/trait	6	2.40%
C	physical	0	0%
D	sociocultural construct	6	2.40%
E	sociocultural persona	230	92%
F	artifact-as-metonym	4	1.60%

$\chi^2 = 1022.1$ ,  $df = 5$ ,  $p < 0.0001$

## 4 Discussion

### 4.1 Individual- vs socioculture-based attributes

Schematically, the ‘*my* N *V-ing*’ and ‘N *in me* *V-ing*’ constructions can be represented as follows:

- (i) [DEIC COP NOMINALIZATION[NP[DET:POSS N]  
V<sub>verbal</sub>-ing]]

- (ii) [DEIC COP NOMINALIZATION[NP[*the* N<sub>i</sub> *in*  
PRO:PER<sub>i</sub>] V<sub>verbal</sub>-ing]]

In terms of form, they appear to be in a minimal-pair relation, the contrastive segments being the NPs [DET:POSS N] in ‘*my* N *V-ing*’, and [*the* N<sub>i</sub> *in* PRO:PER<sub>i</sub>] in ‘N *in me* *V-ing*’, e.g., *This is my mother talking* vs. *This is the mother in me talking*. The two, however, also

appear to be contrastive in terms of their meaning-making function which falls out of the difference in the dominant type(s) of head nouns they attract, viz. ‘mental’, ‘physical’, and ‘personal history, etc.’ nouns in the case of ‘*my N V-ing*’, and ‘sociocultural persona’ nouns in the case of ‘*N in me V-ing*’. The contrast becomes even more salient when we further group these noun-types into two megatypes along the line of individual- vs. socioculture-based attributes. That is to say, whether the noun-types pertain to attributes or aspects of the person which can be considered to be more individual, subjective, or specific to the person (i.e., not likely to be shared or replicated by another), or whether they pertain to attributes or aspects of the person which can be considered to be more socioculturally-derived, and hence not as specific to the person (in the sense that others may share or replicate that attribute or aspect as well, *mutatis mutandis*). For example, a person’s anger (disbelief, despair, etc.) is an emotional state which she (and she alone) can subjectively experience. In contrast, anyone can potentially be a geek, and the socioculture that the person is in has a certain idealized cognitive model (Lakoff 1987) of what constitutes a ‘geek’:

typical behaviors, thinking style, certain typical interpersonal shortcomings, etc.

Looking at the six types of nouns, the ‘mental’, ‘physical’, and ‘personal history, etc.’ nouns can reasonably be considered to designate more individual-based attributes. Emotional, psychological, and cognitive experiences, and physiology and physical features are unique to the individual; for the individual, they are private, subjective, and proprietary. With ‘personal history, etc.’ nouns, this is self-evident. ‘Sociocultural persona’ and ‘sociocultural construct’ nouns are by definition socioculture-based.

In summary, therefore, leaving aside ‘artifact-as-metonym’ nouns for the moment, the other five noun-types can be further grouped into the two megatypes as follows: ‘individual-based’ nouns include ‘mental’, ‘physical’, and ‘personal history/ability/etc.’, while ‘socioculture-based’ nouns encompass ‘sociocultural persona’ and ‘sociocultural construct’ nouns. The incidences of these two megatypes in ‘*my N V-ing*’ and ‘*N in me V-ing*’ (together with those of ‘artifact-as-metonym’ nouns) are shown in Tables 5 and 6 respectively.

Table 5. Megatype distribution of [*my N V-ing*]

noun type	incidence (N=250)	percentage
individual-based (A+B+C)	194	77.60%
socioculture-based (D+E)	52	20.80%
artifact-as-metonym (F)	4	1.60%

$$\chi^2 = 234.27, df = 2, p < 0.0001$$

Table 6. Megatype distribution of [*the N in me V-ing*]

noun type	incidence (N=250)	percentage
individual-based (A+B+C)	10	4%
socioculture-based (D+E)	236	94.40%
artifact-as-metonym (F)	4	1.60%

$$\chi^2 = 419.74, df = 2, p < 0.0001$$

The results show that ‘individual-based’ nouns predominate in ‘*my N V-ing*’ (77.60%), while ‘socioculture-based’ nouns predominate in ‘*N in me V-ing*’ (94.40%). Again, a Chi-square goodness-of-fit test shows the results to be highly significant ( $\chi^2 = 234.27, df = 2, p < 0.0001$  and  $\chi^2 = 419.74, df = 2, p < 0.0001$  respectively).

Thus, in addition to being a minimal pair in syntactic form, the two constructions also stand in quasi-complementarity. They are complementary because ‘*my N V-ing*’ primarily attracts individual-based nouns, while ‘*N in me V-ing*’ primarily attracts socioculture-based nouns. However, they are *quasi-complementary* because they attract the megatypes of nouns that

they do only predominantly but not exclusively. There are some ‘socioculture-based’ nouns attested for ‘*my N V-ing*’ (20.8%), and conversely, some ‘individual-based’ ones attested for ‘*N in me V-ing*’ (4%).

I have not considered ‘artifact-as-metonym’ nouns so far, but it is in fact instructive to do so. Although at first blush, artifacts are necessarily products of a culture, these nouns cannot readily be grouped under the ‘socioculture-based’ megatype. Consider the four attested for ‘*my N V-ing*’: *mask*, *wallet*, *crystal ball*, *dollies*. *Mask* stands for something paraphrasable as ‘the speaker’s public self’, *wallet* stands for something like ‘the speaker’s judgment as to

what counts as a good bet', *crystal ball* stands for 'prediction', and *dollies* is what the speaker projects her wishes onto instead of admitting them to be her own. If they were to be reclassified based on what they stand for, *mask* would arguably come under 'personal history/ability/trait/etc.', *wallet*, *crystal ball*, and *dollies* under 'mental', which would add them to the 'individual-based' megatype. On the other hand, the four attested for 'N in me V-ing': *PES*, *pvp*, *hipster runoff*, and *Google*, all stand for sociocultural personæ. *PES* (pro-evolution soccer) stands for 'player/enthusiast of PES', *pvp* (player-vs-player) stands for 'someone engaging in pvp', *hipster runoff* stands for 'host/user of the blog *hipster runoff*', and *Google* stands for 'user of Google'. If reclassified, these would then add to the 'socioculture-based' megatype.

Thus, it can be seen that even 'artifact' nouns, which are ostensibly sociocultural in nature, conform to the quasi-complementary meaning-making propensities and preferences of the two constructions.

## 4.2 Salience of person vs salience of construct/persona

Besides manifesting in the (mega)types of nouns which the constructions differentially attract, such division-of-labor in meaning-making is also evidenced by the different focuses of attention which the two different NP structures effect. This has to do with the interaction between the nouns and their determiners, i.e., how the nouns are grounded (in the sense of Langacker 1987, 1991; inter alia). Recall that the schematic structures of the two NPs are [DET:POSS N] and [*the* N<sub>i</sub> in PRO:PER<sub>i</sub>]. The salient difference between them is that the former is specified by a possessive determiner: the identity and identification of the referent is effected with reference to a person (the discourse participant) who is the referent's Possessor. In contrast, the latter is specified by a definite article in the first instance, its more exact identity and identification only effected through a postnominal locative phrase.

This structural difference has pragma-semantic import on two counts. Firstly, in the case of [DET:POSS N] (specifically [*my* N] in terms of the data examined in this paper), the possessive determiner is the only and obligatory device allowing the hearer to successfully identify the NP-referent. On the one hand, there is no other information which can help the hearer to conceptually "locate" the NP-referent. On the other hand, this sole piece of information—that the NP-referent is a Possession of the

discourse participant—is not optional information. For instance, *That's my headache talking* is grammatically acceptable, but *\*That's headache talking* is not. As information, it is intrinsic and germane to the NP-referent's identity: The entity belongs to nobody else but the discourse participant.

In the case of [*the* N<sub>i</sub> in PRO:PER<sub>i</sub>] (specifically [*the* N in *me*] for this paper), however, the discourse participant is only conceptualized as an abstract or virtual "location" where the NP-referent can be "located". In other words, the NP-referent gets its particular identity by virtue of where it is located in an abstract or, some might say, metaphorical sense, as opposed to the more intrinsic POSSESSION relation which obtains between the discourse participant and the NP-referent in [*my* N]. Because the postnominal locative phrase is a modifier, its designatum is not the only available information available to the hearer for identifying the referent, nor is it obligatory. Thus, grammatically speaking, *That was the cynic talking* is as acceptable as *That was the cynic in me talking*, although they may be propositionally different, to be sure.

Discussing the concept of POSSESSION in language, Langacker (1991) notes that although the concept prototypically pertains to the relations of part/whole, kinship, and ownership, a variety of other relations can also be designated by possessive constructions, e.g., "an unowned possession (*the baby's crib*); something manipulated (*my rook*); something at one's disposal (*her office*); [...]a transient location (*my spot*); [...]and so on indefinitely" (Langacker 1991: 169). POSSESSION, he therefore argues, in its most schematic and abstract form simply represents a conceptualization process whereby the entity designated as the Possession is conceptually accessed via the entity designated as the Possessor, because it lies within the Possessor's 'dominion'. Langacker (1991: 547) defines 'dominion' as "[t]he set of entities (or the region comprising them) that a particular reference point allows one to establish mental contact with". In other words, the Possessor functions as the reference point by which the Possession becomes mentally accessible to the hearer, because the Possession falls within this conceptual region anchored by the Possessor.

Thus, the second pragma-semantic import falling out from the structural differences between the two NPs is this: [*My* N], insofar as it specifies the NP-referent with a possessive determiner, turns on the conceptualization of the discourse participant as the reference point,

which has high cognitive saliency. The discourse participant as an *individual* is what grounds the conceptualization of the NP-referent (Langacker 1991: 168–175). In contrast, with [*the N in me*], mental contact with the noun-referent does not obligatorily depend on having the discourse participant as a reference point. Here, the discourse participant only provides additional (albeit highly useful, informative, and relevant) information for identifying the noun-referent in the form of a search domain (cf. Langacker 1991: 177). What is attended to here, in the first instance, is the construct or persona (which, in a way, is ultimately a construct) which the *socioculture* has an idealized cognitive model (ICM) for.

Furthermore, both Langacker (1991: 169f) and Taylor (1996: 205f) discuss the Possessor in a possessive construction in terms of topicality, and posit that the Possessor has the special status of a *local topic*, i.e., a kind of topicalized element “at the nominal (as opposed to the clausal) level” (Langacker 1991: 177–178). Thus, in [*my N*], the speaker qua individual is the local topic which serves to anchor the conceptualization of the NP-referent. In a similar vein, then, one could also argue that, at the nominal level, not the discourse participant qua individual, but the sociocultural persona/construct is the local topic which serves such anchoring purpose in [*the N in me*]. In other words, the ‘*my N V-ing*’ construction is primarily *about* the discourse participant qua individual subject, while the ‘*N in me V-ing*’ construction is primarily *about* the discourse participant qua instantiation of that sociocultural construct/persona.

### 4.3 Token- vs type-reference

This complementarity as resources for focusing the hearer’s attention on either the speaker’s more individual-based or socioculture-based aspects of her identity is further evidenced by the type of reference that the two NPs make. [*My N*], by virtue of the grounding properties of its possessive determiner discussed above, has unique reference to a singular, definite, and specific token: the discourse participant (or more precisely, that aspect of the discourse participant’s identity). When the speaker says *That was my paranoia talking*, it is *her* particular paranoia that she is referring to, and no one else’s. It is a subjective psychophysical state of hers as an individual subject.

In contrast, [*the N*] without the postnominal locative has non-unique, non-specific, generic reference to a type; unique and specific reference in this case needs to be facilitated by

the addition of the locative. When the speaker says *That was the disciplinarian in me talking*, she is referring to the sociocultural persona of ‘disciplinarian’ in the first instance; it happens that one aspect of her self instantiates that persona. Pang (2005) discusses in greater detail how [*the N*] is best understood to have generic, categorial reference rather than specific and definite reference, as is the case with [*my N*]. One reason is that the noun-referent of [*the N*] does not even have to exist in reality (although of course, it may), but only in virtuality. For example, someone who has never been a mother in actuality can still say of herself, *That was the mother in me talking*, referring to the ‘maternal instinct’ in her. Or for that matter, a man who considers himself a person with a (strong) ‘maternal instinct’ could conceivably also say the same about himself, perhaps in a moment of “camping it up”. Reference is thus made to a virtual entity—what Langacker (2004) calls a *discourse referent*—which because of its virtuality, is most likely to be constructed from the ICM of the sociocultural construct/persona in question.

That [*my N*] refers to a specific token while [*the N*] refers to a generic type is, to a certain extent, also borne out by the difference in requirement made on the hearer of actual knowledge of the NP-referent. With [*my N*], especially with human terms such as *therapist*, successful interpretation of the utterance requires knowledge of the referent on the part of the hearer. For example, the intended implicature of *This is my therapist talking* cannot be *meaningfully* worked out by the hearer if he does not know the speaker’s therapist, what kind of person they are, or how their personality is relevant to the current discourse (other than perhaps the minimal illocutionary point of ‘hedging’ which the hearer would get from knowing the general pragma-semantics of the construction). It requires more specific knowledge because it makes reference to an individual-based aspect of the speaker’s identity.

With [*the N*], because it refers to a generic type in the first instance, knowledge required for successful interpretation is primarily of the generic kind. It is the sociocultural aspect of the speaker’s self that is brought to bear here. Theoretically, to a degree, an interpretation can be worked out using what knowledge the hearer has of the sociocultural construct/persona referred to, even in the case where the hearer might not have specific knowledge of *how* the speaker instantiates the type. To be sure, equipped with such specific knowledge, the

hearer will be able to formulate a richer interpretation.

## 5 Concluding remarks

It is a fundamental tenet of Cognitive Linguistics that any linguistic expression is a symbolic unit (Langacker 1987). Specifically, within Cognitive Linguistics, various theories taking a Construction Grammar approach to syntactic phenomena (Lakoff 1987; Langacker 1987, 1991; Goldberg 1995; Taylor 2002; inter alia) argue that syntactic patterns are more often than not constructions in the sense that they have very specialized selectional, distributional, structural, and pragma-semantic properties, which make, for example, the two sentences *That was my painkiller talking* and *That was my dog barking* syntactically “identical” but constructionally very different. Two such constructions are investigated in this paper: the ‘*my N V-ing*’ and ‘*N in me V-ing*’ constructions, and it is argued that the two constitute quasi-complementary resources for constructing, asserting, or affirming more individual, subjective, and “personal” aspects of a person’s self on the one hand (‘*my N V-ing*’), and more sociocultural and “generic” aspects of a person’s self on the other (‘*the N in me V-ing*’) (see also Pang 2005).

Firstly, the two constructions preferentially attract different types of nouns to their N-slots, which can be broadly distinguished as two megatypes according to whether they pertain to more individual-based or socioculture-based aspects of a person’s self. The findings can be summarized as follows: ‘*my N V-ing*’ attracts predominantly ‘individual-based’ noun-types and ‘*the N in me V-ing*’ predominantly ‘socioculture-based’ noun-types.

In addition, the structural properties of the two NPs also contribute to their functional quasi-complementarity. [*My N*] is grounded by a syntactically obligatory possessive determiner which expresses a POSSESSION relation between the speaker and the aspect of self in question. In contrast, [*the N in me*] is grounded by a definite article in the first instance, and only a LOCATION relation is posited between the aspect of self in question and the speaker by a syntactically non-obligatory locative phrase. This results in a difference in topicalization: the “starting point” of the conceptualization process. The conceptualization of [*my N*] is anchored to the speaker qua individual, while the conceptualization of [*the N in me*] is anchored to the ICM of the sociocultural persona/construct concerned.

Finally, successful interpretation of examples of the two constructions draw on different types of knowledge. ‘*My N V-ing*’ requires more specific knowledge of the NP-referent, especially if it is a human one. ‘*N in me V-ing*’, by comparison, requires more generic, cultural knowledge of the construct or persona which the NP-referent instantiates.

## References

- Goldberg, A.E. (1995). *Constructions: A Construction Grammar approach to argument structure*. Chicago/London: University of Chicago Press.
- Lakoff, G. (1987). *Women, fire and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.
- Langacker, R.W. (1987). *Foundations of Cognitive Grammar, vol.1: Theoretical prerequisites*. Stanford: Stanford University Press.
- Langacker, R.W. (1991). *Foundations of Cognitive Grammar, vol.2: Descriptive application*. Stanford: Stanford University Press.
- Langacker, R.W. (2004). Remarks on nominal grounding. *Functions of Language* 11, 1: 77–113.
- Pang, K.S. (2005). “This is the linguist in me speaking”: Constructions to talk about the self talking. *Functions of Language* 12, 1: 1–38.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford: Blackwell.
- Stefanowitsch, A., & Gries, S.Th. (2003). Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8, 2: 209–243.
- Taylor, J.R. (1996). *Possessives in English: An exploration in Cognitive Grammar*. Oxford: Clarendon.
- Taylor, J.R. (2002). *Cognitive grammar*. Oxford: Oxford University Press.

## Extremes of L2 Learning: Neurolinguistic Evidence

Loraine K. Obler<sup>1</sup> and Elizabeth Ijalba<sup>2</sup>

<sup>1</sup>The Graduate Center of The City University of New York & <sup>2</sup>Queens College, CUNY

loraine.obler@gmail.com, ijalbaslp@aol.com

### Abstract

Substantial literature on L2 learning focuses on students' motivation as a major factor in success or failure to master the L2, but less attention has been paid to the brain bases of L2-learning proficiency. Yet even among monolingual language acquirers there is evidence from the fields of Neurolinguistics and Speech-Language Pathology that there are differences in rate and mastery of L1 acquisition which are related to subtle differences in brain structure. A small but growing L2 literature focuses on the relationship between dyslexia--a brain-based difficulty learning to read despite normal intelligence and appropriate teaching and social-economic circumstances-- and difficulties with second language acquisition (e.g., Sparks et al., 2006, Ijalba, 2007). A smaller body of literature focuses on brain bases of exceptionally talented L2 learners (e.g., Novoa et al., 1988, Schneiderman and Desmarais, 1988). In this paper we review how Geschwind and Galaburda's (1985) neuroimmunological theory would explain the extreme ends of L2-learning ability, focussing on data from two projects we have been involved with: a study of immigrant children acquiring L2 English in elementary school (Humes-Bartlo et al., 1988) and a study of immigrant adult Spanish speakers acquiring L2 English (Ijalba, 2007).

**Keywords:** dyslexia, talent, L2-learning

### Introduction

Most individuals believe that every child learns a first language equally well, but speech-language pathologists know that there is a small subset of children who, while they do not have any obvious brain damage, and while they have normal exposure to language in their childhood environment, have markedly more difficulty than other children acquiring the sound-system, or morphological intricacies, or lexical items, or, when they are in school, learning to read. The former group--those with difficulties acquiring language generally, are said to have Specific

Language Impairment; the latter group, those whose language difficulties do not surface until they are confronted with learning to read, are labeled "dyslexic" in the U.S. and 'developmentally dyslexic' in Great Britain.

In countries like the United States and Japan, where many individuals can live most of their lives using only the dominant language, most children who have Specific Language Impairment will acquire their country's dominant language eventually, even if at a somewhat slower pace than their typically-developing counterparts. Some may continue to manifest difficulties as they are taught to read. For these individuals, it will not be surprising if learning a foreign language poses difficulty. However for others who showed no marked difficulty with acquiring oral language yet have difficulty in learning to read -- the dyslexics--it is not immediately obvious that they should have difficulties with learning a foreign language. This is especially true because with the right training, most dyslexics will learn to read in their first language.

There is actually a substantial literature on the brain bases of dyslexia, which will be a focus of interest in our paper. It would appear that there are a number of cognitive causes underlying dyslexia, perhaps interlinked across different subtypes of the syndrome. Individuals with dyslexia, as a group, tend to have difficulty with general processing speed, and/or phonological processing speed and working memory, relative to typically developing readers. They have difficulty with tasks which require them to manipulate phonemes and they have difficulty with rapid lexical retrieval, although in daily life their speech may be normal. The brain regions that have been implicated in dyslexia --by post-mortem studies (Galaburda and Livingstone, 1993) and by imaging studies include areas of the temporal and parietal lobes, the planum temporale and the magnocellular portion of the visual system. That is to say, brain regions involved in processing phonemes, lexicon, and visual features are somewhat different in individuals with dyslexia as compared to those with typical language development. We will elaborate further on this topic later in this paper.

Whereas in normally developing individuals there is an asymmetry favoring the left hemisphere, in individuals with dyslexia there is increased symmetry such that the left and right hemisphere are more proportional (Galaburda and Livingstone, 1993; Galaburda et al. 1994). One such area is the planum temporale, an area in the temporal lobe extending from deep in the Sylvian fissure to the posterior border of Heschl's gyrus. This area is considered to be part of the secondary auditory cortex (Shapleske et al., 1999), a finding which is consistent with the type of slight language processing deficits individuals with dyslexia are known to have, with many believing that phonological processing deficits being at the core of the disorder, even though people with dyslexia have refined enough phonological skills to learn their first language and sound like native speakers.

Two other brain areas also show different processing in dyslexics compared to non-dyslexics. One sees decreased activation of the angular gyrus and a concomitant increase in activation in the inferior frontal gyrus –Broca's area-- while reading (Shaywitz and Shaywitz , 2001). This finding is consistent with individuals with dyslexia having difficulty mapping the graphic components of language (letters) to the sounds of the language (phonemes), which result in both phonological and word identification deficits. At the same time, the overactivation in Broca's area seen when people with dyslexia are reading may be an attempt for these individuals to compensate for their phonological deficits and word identification problems. So the general belief that "everyone learns their first language just fine", one must conclude, is not true for a subset of children, perhaps as many as 20 percent.

By contrast with the general belief concerning how 'everyone' learns their first language, when it comes to learning a second, or 'foreign' language, however, at least in North America, it is generally understood that, just as with any course (math, or science, or art), some students will perform better and others worse. Most teachers would assert that the students who do not do well in a given scholastic class are just not working hard enough. For a non-academic class like art or physical education, however, the same teachers would easily agree that some children are simply inherently more talented than most, and, perhaps, that some are inherently less talented than most. We will argue that there are brain-related reasons to believe that the same is true of individuals at the extremes of second-language learning, and that these brain-related reasons are related to dyslexia, at least in the case of those individuals

who have particular difficulty learning a foreign language.

Our data come from a set of studies carried out by ourselves and others. In particular, we consider three group studies that together suggest that there is a particular brain configuration or set of configurations that can make learning a second language difficult, and perhaps as well, another set that may make it particularly easy.

Consider the abilities that are required for learning a second language: picking up a new sound system of distinctions and ways of ordering phonemes in words, linking a new set of words to meanings, picking up ways in which syntax and morphology signal words' roles in sentences, perhaps quite differently from the ways they do in the first language, perhaps only minimally differently, but differently nevertheless.

### Talented L2 Learners

Talented L2 learners seem to undertake such tasks effortlessly. They have a relatively good 'ear' for distinguishing L2 phonemic contrasts and speak with a relatively good accent early on. Vocabulary items are able to stick in their minds with little practice, as are idiomatic phrases. Of course most children can do these things without thinking – that may be their advantage—in a second or third language if they are exposed to it naturally. What is interesting is that even post puberty –the age Lenneberg (1967) posited that the brain loses the plasticity to acquire a new language –there are individuals who can learn a new language exceptionally well.

In the data from the doctoral dissertation study of the second author (Ijalba, 2007), we can see that such individuals who are better than average L2 learners share a subset of cognitive abilities that underlie their excellent foreign-language learning. Namely, in a group of 60 adult Spanish-speaking learners of English as a second language, six individuals provided self-ratings and were rated by their teachers as relatively "good English language learners" (rather than "average or poor English-language learners"). When these individuals, whom she called Good English Language Learners, were tested on a variety of language and reading tasks in **their native language**, they proved to perform better than the otherwise-matched poor English-language learners. For example, they were better in receptive vocabulary in Spanish and English, in the reading of nonwords, in the reading of sentences and in the recognition of correctly spelled words. The Good English Language Learners were also better than Poor English

Language Learners in their performance on rapid automatized naming tasks. For example, they retrieved familiar words in their native language more rapidly than the Poor English Language Learners when presented with the visual stimuli (i.e., they were to name as rapidly as possible a series of blocks of color that were presented in rows).

The overall differences in reading and language performance in their first language, Spanish, that were found between these two groups were always significant for the TIME to complete the task: decoding (reading aloud of nonwords), reading of sentences (reading silently and completing the sentences by selecting the best target word), orthographic decision (picking out correctly spelled words) and rapid automatized naming tasks (colors, objects and letters). In other words, Good English Language Learners read faster and more accurately than Poor English Language Learners. In sum, in this set of Good Language Learners, we were able to see differences from Poor Language Learners who had never been diagnosed as dyslexic on precisely those abilities that distinguish dyslexics from non-dyslexics.

In another study (Novoa et al., 1988), we worked within a different model, asking about the components of “foreign” language-learning aptitude as they might relate to a theory of the neurobiology of talent. In this case study, the first author and her colleagues demonstrated that a particularly talented L2 learner did not perform any better than normal on a battery of musical abilities (thus demonstrating that the lay-person’s belief that those with a ‘musical ear’ are the better L2 learners is false) nor on a task that requires making explicit one’s understanding of grammar (the Words in Sentences test of the Modern Language Aptitude Test (MLAT)). He did perform in the top 10% of the normal scale on several tests of abstraction abilities, for code-like tasks, verbal and visual spatial abilities. Interestingly, we demonstrated that in a number of characteristics, such as being a twin and a left-hander, he fell into the cluster of phenomena that Geschwind and Galaburda (1985) had linked to a variety of talents, albeit in that series of articles they made no mention of L2-learning talent.

One further study that must be mentioned concerning particularly good L2 learners is that of Humes-Bartlo (1989). She studied immigrant children entering the U.S. and being placed in so-called “bilingual” classrooms to learn English along with their other subjects. Each year they were tested at the end of the year to see if they had acquired sufficient English to be moved into

English-only classrooms. Some children took many years to pass the test: the best language learners, however, passed out of the bilingual classes to enter the monolingual classes in fewer than three years. So even in childhood there appears to be substantial inter-individual variability in second-language learning. Interestingly, those who were best at L2 learning scored less well, as a group, on learning mathematics. Thus one might posit that those brain structures useful for learning a second language and those useful for learning basic mathematics are distributed differentially in the population (or, one could argue, though we doubt it, that the two abilities rely on the same brain substrate which, if used for one of the two is then less available for the other.)

### **Those who Learn L2 with Difficulty**

The literature on individuals who learn a second language with difficulty is growing. One sizable body of literature comes from the series of studies by Ganshow and Sparks. They focus on young-adult students in U.S. colleges and universities who are required to take a “foreign” language, but have great difficulty passing the course. They have concluded that such students have difficulty with phonological tasks, English spelling, syntactic measures (both oral and written) and in math calculation (Ganschow et al., 1991) when compared with students who do not experience FL learning difficulty. In a recent paper, Sparks and colleagues (2006) reported on a longitudinal study of younger L2 learners. In this study, 54 students were tested over a 10 year period, beginning in 1<sup>st</sup> grade, to determine best native language predictors of FL proficiency. Their findings suggest that at the beginning of 1<sup>st</sup> grade, the Reading Readiness Test and the Formal Reading Inventory were the best predictors of FL proficiency eight or more years later. However, by the end of 1<sup>st</sup> grade, vocabulary also contributed as a predictor. Similar findings were found at the end of the 2<sup>nd</sup> grade, Reading Readiness, the Formal Reading Inventory and vocabulary continued as the best predictors of later FL proficiency. By the end of 3<sup>rd</sup> grade, the Test of Word Spelling was added as an important predictor. A composite measure that included reading and spelling was found to be the best predictor when children were older. Therefore, by the end of 5<sup>th</sup> grade, the best FL proficiency predictor was the Woodcock Reading Mastery Proficiency Test (WRMT-R), a measure of native language reading that includes decoding of nonwords, word identification and reading

comprehension. The overall findings of this study thus clearly indicate that native language reading and spelling skills in the early years of schooling contributed statistically to the variance in overall FL proficiency eight or more years later (Sparks et al., 2006).

A second major study in this field more specifically indicates dyslexia as the cause, or intermediary cause, of difficulty in L2 learning. That is the same dissertation study mentioned above in which Ijalba (2007) asked what might account for the difficulty some adult Spanish-speaking immigrants report in learning English as a second language. Ijalba divided her 60 participants into those who themselves judged themselves worse (or better) than other English learners and whose English teachers agreed. She then excluded those participants who had had less than 12<sup>th</sup>-grade education, so that “learning opportunity” would not account for the differences among the groups. In this way, she found a population of seven particularly poor learners, six particularly good learners, and 17 average learners who were matched for non-verbal IQ, who had all been in the U.S. for the same number of years (eight years on average) and who had studied English for the same number of years (three years on average). Interestingly, on the neurolinguistic and neuropsychological tests she gave them, her particularly poor L2 learners—though none had ever been diagnosed as dyslexic-- showed problems IN THEIR FIRST LANGUAGE, SPANISH, with precisely those abilities which dyslexic adults, even those who have eventually mastered reading, have: rapid lexical retrieval, spelling, and reading aloud. (The only task they did not show significantly worse performance was on the phonemic awareness task, in which participants had to delete the initial sound in target words and decide what the new word sounded like, perhaps, we speculate, because this task was not challenging enough; the poor learners performed at 92% correct but might have performed worse on a task like Spoonerisms, in which they would have been required to swap the consonantal onsets in two words (e.g., “ball, take” to come up with the words “tall, bake”)

Both the studies of Sparks and Ganshow, and that of Ijalba, then, are consistent with a theory that brain-based problems of dyslexia are linked with poor L2 learning, EVEN IF THE POOR L2 LEARNERS WERE NEVER DIAGNOSED AS DYSLEXIC. The reasons, that they may never have been diagnosed as dyslexic, we posit, are many. They may have grown up in a place, or time, when a diagnosis of dyslexia was not possible, or the schools may have preferred to

avoid such diagnoses to avoid the costs of special tutoring for such students. Or the students may have started school late enough to have caught on to reading without any special remediation. Or they may have been learning to read a language with a relatively transparent, one-to-one letter-to-phoneme correspondence system, like Spanish, that might pose relatively few problems, even for individuals with some degree of dyslexia, though they might remain slower readers than most (as Ijalba discovered her poor-L2-English learners did when reading in their L1 Spanish as adults with a mean age of 35).

A single theory that may link the phenomena we discuss in this paper together we alluded to above: the theory of Geschwind and Galaburda (1985) that noted the startling clustering in some families of a set of neurological factors such as dyslexia and left-handedness (or, more strictly speaking, non-right-handedness). Geschwind and Galaburda were the first to point out differences in brain asymmetry between normal and dyslexic readers, a finding that years later was documented first through the work of Galaburda and is sustained by current research that more specifically points to parietal-temporal-occipital associated areas or a network that may be deficient in dyslexic individuals (Shaywitz and Shaywitz, 2001; Paulesu et al. 2001; McCrory, 2003). These neuroanatomical differences between dyslexics and typical readers are likely linked to the increased incidence of left-handedness that has been reported among dyslexics and their families. While it is unclear precisely how left-handedness is related to the lateralization of language functions, language impairment and dyslexia, the link has been identified in a number of studies. Lamm and Epstein (1999) for example have linked left-handedness —especially in boys-- to POOR foreign-language learning in a sizable sample of Israeli middle-school children. And T. Bever has linked familial left-handedness even in right handers to poor grades in foreign-language learning in U.S. colleges (personal communication, 2009).

How, one must ask, can left-handedness be associated BOTH with exceptionally good and exceptionally poor second-language learning? That is the question that remains to be answered. The theory of Geschwind and Galaburda is most crucially an observation about the clustering in families of certain neurological factors (handedness, dyslexia, talents) with endocrinological ones (such as maleness). In light of such epidemiological data, there is no reason that particularly talented and untalented

individuals would be found in the same families as those with inherited left-handedness. Certain aspects of the links make sense: brain areas responsible for aspects of language crucial for learning a foreign language (i.e., phonology and lexical retrieval) appear to differ between dyslexics and typical readers. People with the symptoms of dyslexia in fact are disproportionately represented among those with difficulty learning foreign languages. Left-handers as a group –and if Bever is right, even the 40% of right-handers with left-handed family members-- have different underlying lateral organization for language compared to right-handers (and, according to Bever, precisely those right-handers with no left-handed family members).

Geschwind and Galaburda (1985) hypothesized that the phenomena they reported to be clustered in families arose during the fetal stage of development, perhaps particularly in the 3<sup>rd</sup> month when lateral dominance and endocrinological systems are developing. How, precisely, the brain differences we've reported between dyslexics and typical readers play out when it comes to learning a foreign language, however, remains to be explained. What is clear, however, is that we must continue to search for the brain-based explanations for how it is that "everyone" (eventually) learns to speak their first language, and virtually everyone learns to read their first language –eventually—but some people have great difficulty learning a foreign language, particularly after early childhood.

## References

- Galaburda, A.M., & Livingstone, M. (1993). Evidence for a magnocellular defect in developmental dyslexia. *Annals of the New York Academy of Science*, 682, 70-82.
- Galaburda, A.M., Menard, M.T., & Rosen, G.D. (1994). Evidence for aberrant auditory anatomy in developmental dyslexia. *Proceedings of the National Academy of Sciences*, 91, 8010-8013.
- Ganschow, L., Sparks, R., Javorsky, J., Pohlman, J., & Bishop-Marbury, A. (1991). Identifying NL difficulties among foreign language learners in college: A "foreign" language learning disability? *Journal of Learning Disabilities*, 24, 530-541.
- Geschwind, N., & Galaburda, A. (1985). Cerebral lateralization: Biological mechanisms, associations and pathology. *Archives of Neurology*, 42, 428-459.
- Humes-Bartlo, M. (1989). Variation in Children's Ability to Learn Second Languages, in K. Hyltenstam and L.K. Obler (Eds.). *Bilingualism across the Lifespan: Acquisition, Maturity, and Loss*. Cambridge: Cambridge University Press, 41-54.
- Ijalba, E. (2007). Markers of Dyslexia in Spanish-Speakers who Report Severe Difficulties Learning English. Unpublished CUNY Doctoral Dissertation.
- Lamm, O. & Epstein, R. (1999). Left-handedness and achievements in foreign language studies, *Brain and Language*, 70, 504-517.
- McCrory, E. (2003). The neurocognitive basis of developmental dyslexia. In R.S.J. Frackowiak, K.J. Friston, C.D. Frith, R. Dolan & C.J. Price (Eds.), *Human Brain Function, Second Edition* (pp. 301-328). San Diego, Academic Press.
- Novoa, L., Fein, D. and Obler, L.K. (1988). Talent in Foreign Languages: A Case Study. In L.K. Obler and D. Fein (Eds.), *The Exceptional Brain: Neuropsychology of Talents and Special Abilities* (pp. 294-302). New York: Guilford.
- Paulesu, E., Demonet, J.F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., Cappa S.F., Cossu, G., Habib, M., Frith, C.D., & Frith, U. (2001). Dyslexia: cultural diversity and biological unity. *Science*, 291, 2165-2167.
- Schneiderman, E. and Desmarais, C. (1988). A neuropsychological substrate for talent in second language acquisition. In L.K. Obler, and D. Fein, (Eds), *The Exceptional Brain: Neuropsychology of Talents and Special Abilities* (pp. 103-126). New York: Guilford.
- Shapleske, J., Rossell, S.L., Woodruff, P.W.R. & David, A.S. (1999). The planum temporale: a systematic, quantitative review of its structural, functional and clinical significance. *Brain and Research Reviews*, 29, 26-49.
- Shaywitz, S.E., & Shaywitz, B. (2001). The neurobiology and reading and dyslexia. *Focus on basics, Connecting Research and Practice*, 5, A. From <http://ncsall.gse.harvard.edu/fob>
- Shaywitz, B.A., Shaywitz S.E., Pugh, K.R., Constable, T., Skudlarski, P., Fulbright, R.K., Bronen, R.A., Fletcher, J.M., Shankweiler, D.P., Katz, L., & Gore, J.C. (1995). Sex differences in the functional organization of the brain for language. *Nature*, 373, 607-609.
- Sparks R.L., Patton, J., Ganschow, L., Humbach, N., & Javorsky, J. (2006). Native language predictors of foreign language proficiency and foreign language aptitude. *Annals of Dyslexia*, 56, 129-160.
- Sparks, R., Patton, J., Ganschow, L., Humbach, N., & Javorsky, J. (2008). Early L1 reading and spelling skills predict later L2 reading and

spelling skills. *Journal of Educational Psychology*, 100, 162-174.

# An Investigation into the Prevalence of Voice Strain in Chinese University Teachers

Zhou Gang and Niu Xiaochun

School of Foreign Languages, Dalian University of Technology  
zhougang@hotmail.co.uk, nxchun@yahoo.com.cn

## Abstract

The aim of the present study was to assess the prevalence of voice problems in the general population of Chinese university teachers, and explore whether their voice problems affected their daily life, their social life and their work. A voice strain and voice handicap index questionnaire was administered to university instructors of English (N = 156) in six Chinese universities. Results indicated that voice strain is prevalent among Chinese university instructors. The respondents' self-perceptions revealed that voice strain was significantly correlated with their job and their daily activities.

**Key words:** voice strain; vocal disorder; factor analysis; multiple regression; MANOVA

## 1. Introduction

Vocal disorders are very common occupation-related disease in teachers, though it has never been given enough attention in China. According to Sonnenberg (2005), teachers represent the largest group of professional voice users, and are among those individuals at greatest risk for developing vocal problems. Orr et al. (2002) state that "One of the problems confronted in the teaching profession is the maintenance of a healthy voice. This basic pedagogical tool is subjected to extensive use, and frequently suffers from overload, with some teachers having to give up their profession altogether" (p. 106). However, the occupational health care of professional voice users is surprisingly undeveloped compared to the attention given to occupational hearing disorders or many other occupational symptoms (Rantala,

Vilkman & Bloigu, 2002).

Teachers are considered an at risk population for vocal problems. This risk has been attributed to those factors such as vocal abuse and misuse due to the vocal demands of teaching, the poor acoustic environments in which teachers work, environmental conditions of school facilities, the lack of vocal education and training, length of career, types of teaching, and even factors related to the individual's emotional state such as stress and anxiety (Mattiske, Oates & Greenswood, 1995). Sapir et al. (1993) also argue that several factors contribute to this risk: the professional demands placed on the voice, unsuitable acoustic working environments, lack of voice training, individual voice characteristics and psychological factors such as stress and anxiety. Vocal misuse refers to voicing behaviors that contribute to the development of laryngeal pathologies, such as nodules, cysts, and polyps (Stemple et al., 2000). Often vocal misuse behaviors become vocally abusive, resulting in a greater likelihood of trauma to the laryngeal mucosa (Colton & Casper, 1996). Some examples of vocally abusive behaviors are: shouting, loud talking, screaming, persistent coughing and throat clearing, etc. (Stemple et al., 2000).

Unfortunately, vocal education and training are not a standard part of most curricula in normal universities or teacher-training programs. Teachers often enter the workforce and begin their careers with limited knowledge of the vocal mechanism, vocal hygiene, and effective voice use. According to Smith et al. (1997), teachers are continually exposed to upper respiratory infections, putting them at great risk for developing illnesses that adversely affect their vocal mechanism. Here the

development of laryngeal pathologies includes those illnesses like nodules, cysts and polyps (Sonnenberg, 2005). The need to develop a more extensive understanding of voice strain among university instructors during loading (talking at work), and the need to develop occupational voice care among teachers have motivated this study. Thus, the aim of this study will address the following questions as research questions: 1. How closely is voice strain associated with teachers' social life and their job? 2. Do gender and different age cohorts affect teachers' self-perceptions of voice problems?

## 2. Method

### 2.1 Participants

Teachers from six research universities in China (They are: Dalian University of Technology, Dalian Maritime University, Dongbei University of Finance & Economics, Dalian Medical University, Dalian Nationalities University and Changchun University of Technology) participated in the study. The participants ( $N = 156$ ) were all instructors of English, who were recruited based on the fact that university instructors of English are assumed to be taking up heavy loadings of teaching responsibility, and they have to teach all years round. Therefore, they are representative of university instructors in China. Among the 156 instructors, 35 were males, and 120 were females (with one data missing). The average age was 35.81 ( $N = 155$  with one data missing,  $SD = 7.819$ ), the youngest was 24 years and the eldest was 56 years. In addition, the average year of teaching was 12.21 ( $N = 156$ ,  $SD = 8.458$ ), the shortest was 1 year of teaching, and the longest was 43 years of teaching. The average teaching hour per week ( $N = 147$  with nine data missing,  $SD = 3.722$ ) was 13.84, the minimum was 4 teaching hours per week, and the maximum was 24 teaching hours per week. Finally, 19% ( $N = 30$ ) of the respondents ( $N = 156$ ) reported that they had received voice training, and 20% ( $N = 31$ ) of the respondents ( $N = 155$  with one data missing) reported that they had received voice therapy.

### 2.2 Procedure

Babbie (1995) observes that questionnaires need to be carefully developed, tested and debugged before they are administered on a larger scale. Initially, therefore, a pilot study was carried out among 10 university instructors of English from Dalian University of Technology. This was principally to determine the reliability of the questionnaire and to reveal any difficulties or ambiguities in question wording. The reliability test showed that *Cronbach's* coefficient alpha on Voice Strain scale was 0.67 ( $N = 10$ ), which was judged adequate. Based on Hinton's criteria (2004), it was regarded as internally moderately reliable for the purpose of this research. The Voice Handicap Index scale consists of thirty items. *Cronbach's* coefficient alphas on the scale were 0.93 ( $N = 10$ ), which was regarded as internally highly reliable for the purpose of this research.

Thus, each scale of the questionnaire has a very good internal reliability and consistency. The findings indicate that the internal reliability of each of the scales is acceptable, and the items that finally make up each scale are internally consistent, which shows that the items measure the two dimensions.

### 2.3 Measures

The participants were given the structured self-completion questionnaire (Voice Strain and Voice Handicap Index Questionnaire: VSVHIQ), which consisted of two scales. In Voice Strain Scale, the researcher designed all the 13 items. In Voice Handicap Index Scale, 30 items were drawn from Jacobson et al.'s (1997) VHI Scale. Respondents were asked to circle one number to correspond with each statement, indicating how frequently they have the same experience in a 7-point Likert-type format (1 = *never*, 7 = *always*).

After the questionnaire was administered on a large scale, the reliability test was conducted again. The reliability test showed that *Cronbach's* coefficient alphas on Voice Strain scale were 0.85 ( $N = 147$  valid), which was judged internally highly reliable for the purpose of this research according to Hinton's criteria (2004). The Voice Handicap Index scale consists of thirty items. *Cronbach's* coefficient alphas on the scale were 0.96 ( $N = 151$

valid), which was regarded as internally highly reliable for the purpose of this research. Therefore, the items that finally make up each scale are internally consistent, which shows that the items measure the two scales.

### 3. Results

In order to summarize and analyse the data collected, some specific statistical techniques were employed (using SPSS14.0). Factor analysis was conducted first to organize the variables on each of the two scales by determining which variables were related and which were not, so that the variables that are related are regrouped together, under a factor. Thus, the variables were reduced to a number of factors representing underlying dimensions of the two scales. Simple Regression analysis was run in order to work out how closely voice strain was associated teachers' social life and their job. Finally, MANOVA was carried out in order to find out any characteristic differences between the two groups' (group 1: males; group 2: females) self-perceptions of the variables on each of the two scales (Voice Strain and Voice Handicap Index).

#### 3.1 Summarizing Data

A principal components factor analysis with varimax rotation method to reduce and summarize the scales, especially to determine which of the items were most useful in measuring each of voice strain scale and voice handicap index scale. The factor analysis conducted on each of the two scales resulted in three-factor solution based on visual inspection of eigenvalues with the scree test (Cattell, 1966). The factor analysis that was run on voice strain scale with factors rotated to orthogonal structure, had produced three factors, and scree test (Cattell, 1966) clearly indicated the presence of three factors in the matrix as well. As can be seen from Appendix I, three factors were indicated as an overall summary of the results across the sample ( $N = 156$ ), relating to perceptions of the voice strain dimension. *Factor 1*, "Voice Strain", explained 38.90% of the variance (eigenvalue = 5.05) and was composed of items indicating "voice strain". *Factor*

2, "Job Impact", explained 12.05% of the variance (eigenvalue = 1.56) and was composed of items indicating "job impact". *Factor 3*, "Strain Reduction", explained 7.86% of the variance (eigenvalue = 1.02) and was composed of items indicating "strain reduction". Items loading  $> 0.45$  are listed under their tentative factor labels.

The factor analysis that was run on voice handicap index scale was a three-factor solution as well because the original scale was composed of three factors (Jacobson & et al., 1997). The first factor, "Functional Issues" measuring the impact of the respondent's voice disorder on daily activities, explained 50.97% of the variance (eigenvalue = 15.29). The second factor, "Physical Issues" measuring self-perceptions of laryngeal discomfort and voice output characteristics, explained 8.24% of the variance (eigenvalue = 2.47), and the third factor, "Emotional Issues" measuring the respondent's affective responses to the voice disorder, explained 4.13% of the variance (eigenvalue = 1.24). Items loading  $> 0.45$  are listed under their tentative factor labels.

#### 3.2 Multiple Correlation Analysis on Perceptions of Voice Strain and Voice Handicap

In answering Research Question1 (How closely is voice strain associated with teachers' social life and their job?), a multiple regression analysis was run. Summary statistics from the stepwise multiple-regression analysis on the sample are presented in Tables 3.1 and 3.2 (see Appendix B). Results from the regression using Voice Strain (labeled as *Voicestrain*) and Job Impact (labeled as *Jobimpact*), and Strain Reduction (labeled as *Strainreduction*) on Voice Strain scale, Physical Issues (labeled as *Newhandicap2*) and Emotional Issues (labeled as *Newhandicap3*) on Voice Handicap Index scale as the independent variables, and Functional Issues (labeled as *Newhandicap1*) as the dependent variable, indicated that in Model 3 *Newhandicap3*, *Voicestrain* and *Newhandicap2* explained 71% of the variance in the teachers' self-perceptions of *Newhandicap1* (functional issues). The results of the ANOVA (see Appendix B) showed that the regression model 3 explained a

significant amount of the variance in the dependent variable *Newhandicap1* with  $F(3, 135) = 112.366$ ,  $p < 0.005$  ( $R^2 = 0.71$ ). Thus, in Model 3 *Newhandicap3* (emotional issues), *Voicestrain* (voice strain) and *Newhandicap 2* (physical issues) ( $p < 0.005$ ) were found to be significant predictors of *Newhandicap1* (functional issues), which made statistically positive contributions to explaining the dependent variable.

Tables 1 and 2 (see Appendix A) illustrate the descriptive statistics and the correlation coefficients for the seven variables entering the equation (1 DV and 6 IDs). The results of the Correlation Matrix also indicated (see Appendix A) that *voicestrain* was significantly correlated with *strainreduction* (drinking water during lectures, or taking breaks more than an hour) ( $r = 0.43$ ,  $p < 0.005$ ), *newhandicap 2* (physical issues) ( $r = 0.51$ ,  $p < 0.005$ ) and *newhandicap 3* (emotional issues) ( $r = 0.40$ ,  $p < 0.005$ ), and *jobimpact* (impact of voice strain on job) ( $r = 0.55$ ,  $p < 0.005$ ). In addition, an examination of the mean scores across the sample (see table 3.3) revealed that the respondents scored high on *voicestrain* ( $M = 4.08$ ,  $SD = 1.405$ ) and *strainreduction* ( $M = 4.77$ ,  $SD = 1.257$ ), which means that voice strain is prevalent among Chinese university instructors, and they tended to believe that drinking water during lectures and taking breaks more than an hour between lectures help reduce their voice strain.

### 3.3 Gender and Age Effects on Perceptions of Voice Strain and Voice Handicap

To investigate research question 2 (Do gender and different age cohorts affect teachers' self-perceptions of voice problems?), a two-way MANOVA analysis was calculated. The MANOVA was a 2 X 4 factorial design comprising the independent variables of gender (males versus females) X age group (group 1: 20-29; group 2: 30-39; group 3: 40-49; group 4: 50-59). The gender group and the age group were both between-subjects factors. It was intended to assess the effects of the two independent variables, one at a time, on the six dependent variables *Voicestrain*, *Jobimpact*, *Strainreduction*, and *Functional*,

*Physical* and *Emotional Issues*. The overall differences were studied first. Results are illustrated in Appendix C. Of interest to this analysis are the results for gender and age group, and gender X age group interaction. The accepted benchmark for significance assessing claims in is Wilks' Lambda with eta being a measure of how strong the effects are. Results (see Appendix C) revealed that there were no significant differences in the dependent variables across the gender groups (males:  $N = 35$ ; females,  $N = 120$ ), and age groups (group 1:  $N = 42$ ; group 2:  $N = 65$ ; group 3:  $N = 36$ ; group 4:  $N = 12$ ), with Wilks' Lambda ( $6, 126$ ) = 0.985,  $p > 0.05$ , and with Wilks' Lambda ( $18, 356$ ) = 0.892,  $p > 0.05$ , respectively. Thus, significant multivariate effects for gender and age group were not obtained (i.e. a significant  $F > 0.05$ ). As indicated in the table, according to the generally accepted criteria (Cohen, 1988), Eta for gender and age group indicate that these are not strong or large effects with Partial Eta Squared = 0.015 and with Partial Eta Squared = 0.038, which explained either 1.5% or 3.8% of the variances in the dependent variables scores explained by gender and age group. Moreover, the overall analysis showed that there was no interaction effect (Wilks' lambda ( $18, 356$ ) = 0.840,  $p > 0.05$ ). Therefore, these effects will not be discussed any further in the following study, since no significant multivariate effects were found for gender and age group.

## 4. Discussion and Conclusions

The aim of the present study was to assess the prevalence of voice problems in the general population of Chinese university instructors, and explore whether their voice problems affected their daily life, their social life and their work, and how closely they were correlated with each other. Results of the multiple-regression analysis on the sample revealed that the university instructors' self-perceptions of *Newhandicap3* (emotional issues), *Voicestrain* (voice strain) and *Newhandicap 2* (physical issues) dimensions were significantly positively correlated with their self-perceptions of *Newhandicap1* (functional issues) dimension. It was found that emotional issues (the respondent's

affective responses to the voice disorder), voice strain and physical issues (self-perceptions of laryngeal discomfort and voice output characteristics) were the best and useful predictors of the functional issues (the impact of the respondent's voice disorder on daily activities), which means that the more often a teacher experiences emotional issues, voice strain and physical issues, the more often he/she experiences functional issues. The findings also revealed that voice strain is prevalent among Chinese university instructors, and they believe that they have to drink water when they lecture, and that taking breaks during lectures more than an hour helps reduce their voice strain. Moreover, the findings showed that voice strain was significantly correlated with job impact, which means that the more often a teacher experienced voice strain, the more impact it exerted on his/her job. These findings are significant because they seem to have provided evidence for answering the first research question.

In answering Research Question 2, we found that there were no significant main effects for gender factor (males vs. females) and age factor (age group 1, age group 2, age group 3 and age group 4). In other words, both gender and age cohorts made no any significant differences in self-perceived voice strain, strain reduction, job impact, and functional, physical and emotional issues based on the scores gained in the respondents. Interestingly, the statistical analyses and results revealed no differences between the four age groups' self-perceptions of voice strain and job impact, which has been one of the research focuses in the present study.

Some findings in this study were significant. For instance, the respondents reported that they often experienced voice strain during/after lectures, which suggested a need for preventive voice programs (such as voice care/hygiene, and voice training) within teacher-training students or in-service teachers training programs, which may help increase teachers' vocal endurance and ability in such a challenging profession. In addition, 20% of the respondents ( $N = 31$ ) reported that they had had previous voice therapy, which further supported

the finding that voice strain in university instructors is prevalent. In short, it may be inferred that voice strain in teachers does affect their jobs and daily activities based on the respondents' self-perceptions. Thus, the research has reached its proposed research purpose.

However, several limitations of the present study should be noted and addressed in future research. For example, all constructs in this study were assessed with paper-and-pencil measures. Although this is defensible in the early phases of a research program, future studies should seek the use of a therapy approach for teachers that includes treatment techniques to improve vocal hygiene habits. Moreover, the sample size should seek to concentrate on the respondents' sample scope, thus it will be more representative of the whole population of Chinese university instructors.

Furthermore, this study strongly suggests that teachers may benefit from vocal educational training and prevention programs to raise their awareness of potential voice problems and to promote vocal health, so that the qualities of their teaching and social life may be improved.

## References

- Babbie, E. (1995). *The practice of social research*. Belmont: Wadsworth Publishing Company.
- Cattell, R.B. (1966). "The Scree Test for the number of factors", *Multivariate Behavioral Research*, 1: 245-276.
- Colton, R.H., & Casper, J.K. (1996). *Understanding voice problems: A physiological perspective for diagnosis and treatment* (2<sup>nd</sup> Ed). Maryland: Williams & Wilkins.
- Hinton, P.R., Brownlow, C., McMurray, I., & Cozens, B. (2004). *SPSS explained*. London: Routledge.
- Jacobson, B.H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M.S., et al. (1997). "The voice handicap index (VHI): Development and validation", *American Journal of Speech-Language Pathology*, 6/3: 66-70.
- Mattiske, J.A., Oates, J.M., & Greenwood, K.M. (1998). "Vocal problems among teachers: A

- review of prevalence, causes, prevention, and treatment”, *Journal of Voice*, 12: 489-499.
- Orr, R., de Jong, F. & Cranen, B. (2002). “Some objective measures indicative of perceived voice robustness in student teachers”, *Logopedics, Phoniatrics, Vocology*, 27: 106-117.
- Rantala, L., Vilkman, E., & Bloigu, R. (2002). “Voice changes during work: Subjective complaints and objective measurements for female primary and secondary schoolteachers”, *Journal of Voice*, 16: 344-355.
- Sapir, S., Keidar, A., & Mathers-Schmidt, B. (1993). “Vocal attrition in teachers: Survey findings”, *European Journal of Voice Disorders of Communication*, 28: 177-185.
- Smith, E., Gray, S.D., Dove, H., Kirchner, L., & Heras, H. (1997). “Frequency and effects of teachers’ voice problems”, *Journal of Voice*, 11: 81-87.
- Sonnenberg, L.L. (2005). “A therapeutic approach for improved vocal performance in individuals in teaching occupations”, unpublished MA thesis, The University of North Carolina at Greensboro.
- Stemple, J.C., Glaze, L.E., & Klaben, B.G. (2000). *Clinical voice pathology: Theory and management* (2<sup>nd</sup> ed). San Diego, CA: Singular Publishing Group.

# Memory Strategy Instruction, Contextual Learning and Vocabulary Recall

Somayeh Kazemi Tari

Department of Language and Literature, Science and Research University, Tehran, Iran

somayehkazemitari@gmail.com

## Abstract

Language learning strategies play a prominent role in learning a second language. Students need to self-direct learning in general and language learning in particular, by using different strategies in the process of learning. To take responsibility for their learning, learners should develop independent language learning strategies and in order to do so they may need some help from their teachers at the first steps. In the last decades there has been an increasing interest in vocabulary strategies because they are found to facilitate second/foreign language vocabulary learning and recall. Using a broad range of strategies through their vocabulary learning process, students can become more proficient. The present study investigates the effects of memory strategy instruction along with learning through context on the vocabulary recall of Iranian EFL learners. The study also explores whether there will be any difference in the use of memory strategies of the learners as a result of such instruction.

## Keywords

Memory, strategy, contextual learning, vocabulary recall

## Introduction

Language learning strategies have long been the subject of research. They help learners to learn independently when the teacher is physically absent (Hurd & Lewis, 2008). Because of the centrality of lexical knowledge to learning a language research into the effectiveness of various types of vocabulary learning methods have been of considerable value to second language learning research. In the last few decades there has been an increasing interest in vocabulary learning

strategies because they are found to facilitate second language vocabulary learning.

According to Nunan (1999) many language learners are not aware of the strategies underlying their learning tasks. Also research shows that more strategy use often leads to higher levels of proficiency in language learners (Oxford and Ehrman, 1995).

According to Oxford (1990) self direction is essential in the active development of adults' abilities in learning. O'Malley and Chamot (1990) acknowledged the link between strategic competence in language learning, learner autonomy and successful outcomes on the physical absence of a teacher. Hence, the notions of independence, autonomy and control in learning experiences play an increasingly important role in language education.

According to Oxford (1990) independent language learning (ILL) reflects a move towards more learner-centered approaches viewing learners as individuals with needs and rights, who can develop and exercise responsibility for their learning.

Concern for the individual learner and for learner choice, control and responsibility has been a pervasive influence on language learning and teaching for more than three decades and is central to the idea and practice of independent language learning. As Benson and Lor (1999) emphasize the expectation that language learners can be independent, and that this is an important attribute and goal, underlies much of the writing on learner autonomy such as works on self-access learning, self-directed learning and different forms of online learning such as tandem partnerships explained in Hurd et al. (2001).

Dörnyei (2005) argues that the most effective

way of developing favorable attitudes towards independent learning is for teachers to prepare language learners to think about their needs and objectives and then to learn how to structure their learning.

Ellis (1994) believes that there is a relationship between independence and autonomy. Ellis (1994) sees autonomy in the context of language acquisition as involving an ability to operate independently with the language and use it to communicate personal meanings in real, unpredictable situations. In the aggregation, researchers consider independent learning as a particular context for learning, and as an approach to learning as a goal of education (Ulitsky, 2000).

### 1. Learning Strategies

Various taxonomies of learning strategies have been proposed by different researchers in the past two decades. O'Malley and Chamot (1990) distinguish between metacognitive, cognitive and socio-affective strategies. Oxford (1990) proposed Strategy Inventory for Language Learning (SILL) which consists of direct strategies (including memory, cognitive and compensation strategies) and indirect strategies (including metacognitive, affective and social strategies).

According to Oxford (1990) language learning strategies are commonly defined as the operations or processes which are consciously selected and employed by the learner to learn the TL or facilitate a language task.

Studies on learning strategies have focused on two broad domains: the strategies learners use, and ways of enhancing independent language learning through strategy training and strategy development. This concept can be analyzed from two perspectives namely: Strategy use in independent language learning as well as strategy development and language learning as explained in Paige et al. (2004).

These studies show how learners develop creative language use strategies to build a rich interface with the learning context, how strategy use functions as a means of matching learning needs with affordances of the context, and changes in strategy use over time (Paul, 1990).

In order to deploy self-management strategies, learners need to know how they learn best, and need to have the necessary

procedural skills to set up optimal learning conditions.

One domain of studies about strategies concerns ways of enhancing ILL (independent language learning) through strategy training and strategy development. In these studies researchers explore ways of best providing strategy instruction for particular independent language learning environments, the role of the teacher in that process and the relationship between feedback and strategy development. The issue of implicit learner training in strategies was touched on in the work of Yang (1999). He concludes that learning a language through the distance-learning mode is a challenging task and the feedback we offer students may be crucial to them in terms of helping them to improve their learning strategies. The overall aim is to provide opportunities for learners to manage their independence in optimal ways, by acquiring a series of strategies and skills that will enable them to work individually.

Research on the use of vocabulary strategies has revealed differences among learners in terms of their strategy use. Successful vocabulary learners were found to be active strategy users who were conscious of their learning and took steps to regulate it, whereas poor learners displayed little awareness of how to learn new words or how to connect new words to old knowledge (Sanaoui, 1995). According to Gu (2005) successful learners intentionally select, consciously monitor and evaluate the strategy they use for the fulfillment of their aim. The unsuccessful, on the other hand, employ learning behaviors similar to their peers without being conscious but also without having an aim.

### 2. Research Questions and Hypotheses

The present study thus aimed at investigating the effects of memory strategy instruction along with learning through context on the vocabulary recall of Iranian EFL learners. The study also explores whether there will be any difference in the use of memory strategies of the learners as a result of such instruction. The following research questions were addressed in this study.

1. Will there be any difference between the students who have memory strategy instruction

along with contextual learning and those who only have contextual learning in terms of their vocabulary knowledge?

2. Will there be any difference in the memory strategy use of the experimental students after the memory strategy instruction?

Based on the above mentioned questions the following null hypotheses were proposed and tested:

1. There will be no difference between the students who have memory strategy instruction along with contextual learning and those who only have contextual learning in terms of their vocabulary knowledge.
2. There will be no difference in the memory strategy use of the experimental students after the memory strategy instruction.

### 3. Method

#### 3.1. Participants

One hundred forty five B.A. students studying EGP (English for General Purposes) at Ghaemshahr University participated in this study. The ages of the participants ranged from 17 to 27. Following randomized matching method based on their scores on a teacher-made vocabulary test the students were assigned to experimental and control groups. Based on this procedure 43 matched pairs were found between the two groups. These matched pairs of students were selected for the purpose of the study to assure the homogeneity of the subjects in the two groups with regard to vocabulary knowledge.

#### 3.2. Instruments and Data Collection Procedures

Data for the present study were collected by means of a multiple choice vocabulary knowledge test prepared by the researcher. The test consisted of 30 items which aimed at assessing the level of vocabulary knowledge of students. The words or phrases to be tested were selected from among the target vocabulary course book. The vocabulary was selected from the 7 units to be covered during the study. The test was piloted on 30 students who took the same EGP course and the reliability of the test was estimated to be 0.79.

In order to find out the memory strategies used by the students at the beginning and end of the study, students in the

experimental group were given the Memory Strategy Questionnaire (Appendix) adapted from Atay and Ozbulgan (2007) as a pretest and posttest. Students were told they could choose more than one strategy on the Memory Strategy Questionnaire.

#### 3.3. Materials and Instruction

Both groups used the same course book, i.e., General English Through Reading (Baradaran and Mosallanezhad, 2003). The typical flow of the instruction of a unit was as follows (see Table. 1 for a sample of one week's instruction): first students read a text then they answered different post-reading activities, e.g., answering comprehension questions, matching. Both groups had 1.5 hours of English a day, two days a week.

Table 1. A Sample of one Week's Instruction

Week 1	Control group	Experimental group
Day 1	Introduction to the course	Introduction to the course Introductory lesson on vocabulary learning, modeling of all memory strategies
Rest of the week		
	1.5 hour a day Reading the text, guessing meaning of unknown words, doing post-reading activities, e.g., answering questions, and matching	1 hour a day Reading the text, guessing meaning of unknown words, doing post-reading activities, e.g., answering questions, and matching
		Half an hour a day Memory strategy instruction

On the first day of the program students in the experimental group were given an introductory lesson about vocabulary learning. The researcher first talked about the importance of vocabulary knowledge in language learning and discussed different ways of vocabulary learning, i.e. implicit and explicit. Then she moved on to vocabulary strategies and focused on the memory strategies in Atay and Ozbulgan's (2007) list. First the researcher modeled each strategy

for the students giving different examples for it. For example the teacher asked one student to carry instructions such as follows:

1. Try to imagine the word that I say in English and its Persian meaning.
2. Try to imagine yourself interacting with the object while saying its name.
3. Show a kind of physical movement that represents that object or concept.

On the following day, after having regular instruction, the teacher distributed the handout of memory strategies to students and practiced the strategies with in class. For the rest of the study, i.e., 12 weeks, the experimental group students looked at the target words from each unit and practiced applying the strategies in class to memorize the words and their meanings. They discussed the strategies that most suitably helped them to learn and retain the words in memory with the class. For example one student said that he learned a word by relating it to a word in her mother language that had a similar pronunciation. Another student said that he memorized a list of unrelated words by grouping them together in a sentence which seemed to be ridiculous or strange. Sharing and discussing their strategies, they helped each other with vocabularies.

### 3.4. Data Analysis

To determine whether there is any difference between the two groups in terms of vocabulary knowledge at the end of the study, a paired t-test was applied. The MSQ (memory strategy questionnaire) was analyzed by means of taking frequency counts of each strategy in the pre- and post-tests.

## 4. Results

Analysis of the MSQ (Memory Strategy Questionnaire) administered in experimental group both before and after the treatment showed the frequency of each strategy used by the students. Although there was a little difference between the number of strategies reported to be used by individuals in pretest and the number of the strategies used in posttest, but the frequency of the strategies used by the whole experimental group remained noticeably unchanged. These strategies are arranged in table 2 according to their reported frequency in a descending order (i.e. strategy 14 in

Appendix is the most frequently used strategy while strategy 6 in Appendix is the least frequently used strategy):

Table 2. Frequency of the Strategies Used by Experimental Group

Strategies Arranged Based on Their Frequency	
14. Imaging word's form	
12. Saying new word aloud when studying	
13. Imaging word's meaning	
9. Studying the spelling of the word	
2. Associating the word with its coordinates	
1. Connecting the new word to a previous personal experience	
11. Connecting the word with its synonyms and antonyms	
7. Paraphrasing the word's meaning	
4. Using physical action when learning a word	
10. Grouping words	
3. Using semantic maps	
8. Underlining initial letter of the word	
5. Grouping words together within a storyline	
6. Using scales for gradable adjectives	

This indicates that explicit teaching of memory strategies does not always lead to more frequent use of the strategies.

The results of the paired t-test between experimental and control group posttest did not show any significant difference between the groups in terms of their vocabulary knowledge gain scores (see Table 3).

Table 3. Experimental and control group Posttest and Pretest Means\* (M), Standard Deviations (SD), and their Differences (t)

	Pretest		Posttest		t	df	p
	M	SD	M	SD			
Exp.	10.86	5.70	11.71	5.47	0.77	84	<0.05
Cont.			10.69	6.73			

\* The maximum total number was 30

Although mean of the experimental posttest shows an improvement over the mean of the experimental pretest (11.7143 vs. 10.8604 respectively) and although the mean of the experimental group is higher than the mean of the control group posttest (11.7143 vs. 10.6933 respectively), but the computation of t-test ( $t=0.7716$ ,  $p<0.05$ ) revealed that this difference was not so significant to be attributed to the effect of experimental treatment.

## 5. Discussion

The present study aimed to investigate the effects of memory strategy instruction along with learning through context on the vocabulary knowledge of EFL students and compare these with those of contextual learning alone. The result of the study showed that there was no significant difference between the vocabulary recall of experimental and control group students at the end of the instructional period. During the instructional period, both groups studied the same course book which exposed learners to the target vocabulary in several contexts. While the control group learned the vocabularies only through context, experimental group received memory strategy instruction in addition to contextual learning. During the instructional period the teacher provided the students in the experimental group with different memory strategies and encouraged them to use these strategies to memorize new vocabularies. Experimental group students were shown explicitly how the strategies work. The teacher herself modeled each strategy providing student with different examples.

Although it seems that strategy instruction in the current study was not successful but might be due to a number of intervening and uncontrolled variables such as learners' cognitive style or their preferred stabilized strategies that they have adapted and used for a long period of time in their leaning behavior.

One of the interesting features of this study was the requirement of practical use of strategies inside the classroom because when one student shared his memory strategies with the class others learned how to apply that strategy.

Results from the MSQ showed that 'imaging word's form' and 'Saying new word aloud when studying' were the most frequently used strategies among Iranian EFL learners, unlike learners in other studies such as Atay and Ozbulgan (2007) who regarded 'connecting the new word to a previous experience' and 'using semantic maps' as the most frequent strategies.

One of the factors that caused control group students to perform as good as experimental group students might be that they were highly motivated to learn new word. Moreover the present study assessed

the short term effects of memory strategy instruction on learners' vocabulary recall and the result of the same study might be quite different in a longitudinal study investigating the long term effects of memory strategy instruction on vocabulary recall. Future research is needed to clarify these points.

## References

- Atay, D., & Ozbulgan, C. (2007). Memory strategy instruction, contextual learning and ESP vocabulary recall. *English for Specific Purposes*, 26, 39-51.
- Baradaran, A. & Mosallanezhad, P. (2003). General English through reading. Tehran: SAMT publications.
- Benson, P., & Lor, W. (1999). Conceptions of language and language learning. *System*, 27, 459-472.
- Dörnyei, Z. (2005). *The Psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Gu, P. Y. (2005). Learning strategies: Prototypical core and dimensions of variation. Retrieved December 12, 2005, from [http://www.crie.org.nz/research\\_paper/Peter\\_Gu.pdf](http://www.crie.org.nz/research_paper/Peter_Gu.pdf).
- Hurd, S., Beaven, T., & Ortega, A. (2001). Developing autonomy in a distance language learning context: Issues and dilemmas for course writers. *System*, 29 (3), 341-355.
- Hurd, S., & Lewis, T. (2008). *Language learning strategies in independent settings*. London: Cromwell Press.
- Nunan, D. (1999). *Second language teaching and learning*. Boston, MA: Heinle and Heinle.
- O'Malley, J., & Chamot, A. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.
- Oxford, R. (1990). *Language learning strategies: What every teacher should know*. New York: Newbury House.
- Oxford, R. & Ehrman, M. (1995). Adults' language learning strategies in an intensive foreign language program in

- the United States. *System* 23, 359-386.
- Paige, R., Cohen, A., & Shively, R. (2004). Assessing the impact of a strategies-based curriculum on language and culture learning abroad. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 10, 253-276.
- Paul, R. (1990). Towards a new measure of success: Developing independent learners. *Open Learning*, 5 (1), 31-38.
- Sanaoui, R. (1995). Adult learners' approaches to learning vocabulary in second language. *Modern Language Journal*, 79, 15-28.
- Ulitsky, H. (2000). Language learner strategies with technology. *Journal of Educational Computing Research*, 22 (3), 285-322.
- Yang, N.D. (1999) The relationship between EFL learners' beliefs and learning strategy use. *System*, 27 (4), 515-535.

## Appendix

Please, identify the memory strategies you have used in learning the new vocabulary during the EGP course.

1. Connecting the new word to a previous personal experience
2. Associating the word with its coordinates
3. Using semantic maps
4. Using physical action when learning a word
5. Grouping words together within a storyline
6. Using scales for gradable adjectives
7. Paraphrasing the word's meaning
8. Underlining initial letter of the word
9. Studying the spelling of the word
10. Grouping words
11. Connecting the word with its synonyms and antonyms
12. Saying new word aloud when studying
13. Imaging word's meaning
14. Imaging word's form

# Do Visual Supports Enhance EFL Listeners' Self-confidence? An Empirical Study

Hsiao-fang Cheng<sup>1</sup> & Nina Chiulan Lin<sup>2</sup>

<sup>1</sup>National United University, Taiwan

<sup>2</sup>Takming University of Science and Technology, Taiwan

<sup>1</sup>hfcheng@nuu.edu.tw; <sup>2</sup>ninachiulan@yahoo.com.tw

## Abstract

The integration of technology into education has expanded methodological options for EFL teaching and learning. Providing visual supports, which is more relevant to what people usually do in their daily communications, does significantly enhance EFL listening comprehension. However, do visual supports promote EFL listener's self-confidence in their own listening ability? This study was done to investigate whether the listener's self-confidence increased as the auditory stimuli were accompanied with visual supports. A listening test presented with three different modes, audio-only, listening stimuli with still pictures, and video listening were utilized to collect data. Listener's self-confidence was measured by their responses on a 6-point scale after completing every comprehension question. The results showed that no significant difference on self-confidence was found as the visual aids were progressively increased; i.e. visual supports did not necessarily promote listener's self-confidence in his listening ability. However, listener's self-confidence was highly correlated with their listening performance. The more confidence they had in answering the questions, the more questions they answered correctly. Drawing on the results, some pedagogical implications are proposed.

## Key words

self-confidence; mode of presentation; motivation; listening performance

## 1. Introduction

Motivation is a concept that explains why people behave as they do rather than how successful their behavior will be. According to Csizer and Dornyei (2005), L2 motivation contains seven components: integrativeness, instrumentality, vitality of the L2 community, attitudes toward the L2 speakers/community, cultural interest, linguistic self-confidence and milieu. Those motivational factors affect the effectiveness and success in the classroom to a great extent (Horwitz, 1986). Self-confidence, one of the motivational constituents, reflects a confident, anxiety-free belief that the mastery of a L2 is well within the learner's capabilities. Learners believe that they have the abilities to reach goals or finish tasks successfully (Csizer & Dornyei, 2005). Drawing on Clement's (1980) theory, the main antecedents of self-confidence are the quality and quantity of social contact. It is particularly a function of the beliefs conveyed by one's environment and one's attitude toward the L2 culture.

Clement, Dornyei, and Noels (1994) argued that attitudinal factors were an important motivational basis for L2 acquisition and behavior. A self-confidence process becomes the most important determinant of attitude and the amount of effort exerting on L2 learning. Positive attitudes would orient the individual to seek contact with the target language. To the extent that the contact is relatively frequent and pleasant, self-confidence in

using the L2 would develop as anxious affect would decrease. Research shows that students with little anxiety are satisfied with their current level of English proficiency, and thus with high self-perceptions of L2 competence, they become positive and frequent contact with English. Research has suggested that anxiety and self-perceptions of L2 competence may be determinants of L2 achievement even in contexts where opportunity to use the L2 with members of the L2 community is lacking. To sum up, self-confidence, leading to motivated behavior, would motivate learners to learn the target language and find the course easy (Clement, et al, 1994). In this study, it is more like a basic learner disposition of trust in one's ability to answer the listening question accurately.

Feyten (1991) stated that forty-five percent of a person's total communication is spent in listening, which is critical in general human communication and in academic settings. Getting students ready to deal with natural, communicative situations is an important function of classroom-listening activities (Krashen, 2002). With the development of high technology, the widespread use of computer-aided multimedia has facilitated foreign language learning, especially listening, to a great extent. Multimedia in its simplest form is the combination of two or more media such as text, images, and sound. Numerous studies have documented that static visual aids reduce cognitive load and enhance the learners' listening comprehension (Cheng, 2003; Mousavi, low, & Sweller, 1995; Park, 1998; Ruhe, 1996). Videos with full motion pictures, exposing students to authentic context and providing cultural contexts for the language, promote listening comprehension (Swaffar & Vlatten, 1997). To sum up, visual supports did significantly improve listening comprehension. Drawing on the research result, the listening items in the Test of English as a

Foreign Language (TOFEL) include visual accompaniments--still photos, drawings or pictures to verbal stimuli. It is hoped that the listening test can accurately assess listener's listening proficiency in the real-life communication context.

As stated above, visual supports enhance listening comprehension which might lead to the improvement of self-confidence in mastering the target language. It is generally believed that the more comprehension listeners obtain, the more self-confidence they will have in the course of listening. Therefore, it is worth pondering if visual aids, such as facial expressions or speech-related gestures, improve listener's self-confidence. This study was designed to investigate if visual supports would enhance listener's self-confidence. The three different modes of presentation employed in this study were audio listening without visual support, audio listening in conjunction with multiple still pictures indicating the turns of the speakers, and video listening with full-motion images. Since self-confidence in the language learning context is usually assessed with measures of perceived proficiency at the time of testing, participants were required to indicate their confidence level after completing each listening question.

## **2. Method**

### **2.1 Participants**

One hundred and sixty-five seventeen to nineteen-year-old non-English majors served as the subjects, all of whom enrolled in the fall semester, 2007. They were divided into three groups based on their academic majors. According to the homogeneity of variance assumption, the listening proficiency of all the groups had to be equal prior to treatment; otherwise, the t-statistic would be meaningless (Peers, 1996). Therefore, an intermediate level listening pretest was administered before the experiment. These three

groups were homogenous in listening comprehension proficiency at the  $p < 0.05$  ( $df = 2$ ,  $F = 107$ ;  $p < 0.899$ ) level of significance. Their English ability was at intermediate level with their reading better than their listening.

## 2.2 Procedure

A printed test of nine multiple-choice items with questions and alternatives, phrased in the target language, was given before viewing or listening. After completing the test items, participants were required to indicate the confidence level in their response. The statistical analyses for the data were carried out with SPSS X. The repeated-measures one-way ANOVA was applied to analyze the three main variables, using Scheffe's test comparing the mean score gained from the three different modes of presentation.

## 3. Results

No significant difference was found for self-confidence when the presenting modes were different ( $df = 2$ ,  $F = 19.89$ ,  $p < 0.75$ ). Though the difference was non-significant, participants were somewhat more confident when the auditory messages were presented with still pictures (Mean 2.96). They had less confidence in listening to audio-only texts (Mean 2.68).

In terms of listening performance, participants performed the best in still picture listening, then listening with full-motion images whereas audio-only listening came the last ( $df = 2$ ,  $F = 23.378$ ,  $p < 0.05$ ) (Cheng, 2003). The mean scores of both the still picture listening and the video listening were significantly higher than that of the audio-only listening, indicating that listening in conjunction with pictures was easier for the subjects.

The mean score of the still picture listening, however, was somewhat higher than that of the video listening in terms of listening performance

and self-confidence. Such finding indicated that full-motion pictures, providing authentic communication context and paralinguistic aspects of communication (Altman 1989), do not necessarily promote learner's achievement (Cheng, 2003) and self-confidence.

This study also found that self-confidence positively correlated with listening outcome. The higher the participants scored, the more confident they became. A negative correlation between self-confidence and listening anxiety was found. The less anxious the participants were, the more confidence they had. Even though videos with full motion pictures give redundant and additional information, they did not substantially promote learner's confidence in their own listening proficiency.

## 4. Conclusion and Suggestions

To summarize, the results of this study show that modes of presentation significantly influence the learner's listening performance but slightly affect his/her linguistic self-confidence. Audio-only listening is the most difficult presenting mode and visual aids somewhat enhance listener's self-confidence. There was only one group of Chinese technical college students participating in this study. The results presented in this paper cannot be taken as conclusive and definitely do not allow for generalization. This study should be repeated with other groups of different age, language proficiency, or language background to examine whether the conclusions are the same as those in this study.

## 5. References

- Altman, R. (1989). *The Video Connection: Integrating Video into Language Teaching*. Boston, MA: Houghton Mifflin.
- Cheng, H. F. (2003). Effects of modes of

- presentation on listening comprehension. Paper presented at Teaching and Assessing Language Proficiency 38<sup>th</sup> RELC International Seminar, 03-05 November.
- Clement, R. (1980). Ethnicity, contact and communicative competence in a second language. In H.M. Giles, W.P. robinson, & P. M. Smith (Eds.), *Language : Social Psychological Perspectives*. pp 147-154). Oxford, UK: Pergamon.
- Clement, R., Dornyei, Z., & Noels, K.A. (1994). Motivation, self-confidence, and group cohesion in the foreign language classroom. *Language Learning*, 44 (3), 417-448.
- Csizer, K. & Dornyei, Z. (2005). The internal structure of language learning motivation and its relationship with language choice and learning effort. *The Modern Language Journal*, 89(1), 19-36.
- Feyten, C. (1991). The power of listening ability: An overlooked dimension in language acquisition. *Modern Language Journal*, 75, 173-180.
- Horwitz, E.K. (1986). Preliminary evidence for the reliability and validity of a foreign language anxiety scale. *TESOL Quarterly*, 20, 559-562.
- Krashen, S. D. (2002). *Explorations in language acquisition and use: The Taipei lectures*. Taipei: Crane Publishing Co., Ltd.
- Mousavi, S. Y., Low, R. & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, 87(2), 319-334.
- Park, Ok-choon. 1998. "Visual Displays and Contextual Presentations in Computer-based Instruction." *ETR&D*, 46 (3) :37-50.
- Peers, I. S. (1996). *Statistical analysis for education and psychology researchers*. London: The Falmer Press.
- Ruhe, V. 1996. "Graphics and Listening Comprehension." *TESL Canada Journal*, 14 (1): 45-60.
- Swaffar, J., and A. Vlatten, 1997. "A Sequential Model for Video Viewing in the Foreign Language." *Modern Language Journal*, 81:175-188.

# SEL-Hi: Current and Future English in Japan

Hideyuki Kumaki

School of Education, Waseda University

t-kumaki@mua.biglobe.ne.jp

## Abstract

Concerning English education, there have been three buzzwords recently. One is “EIL”<sup>1</sup>, which means English as an International Language, and another one is “ELF”<sup>2</sup>, meaning English as a Lingua Franca. Also, there is “World Englishes”, which was suggested by Kachru (1992)<sup>3</sup>. As these terms suggest, it has been said that English is not only for native speakers of English but also for non-native speakers. The number of L2 speakers will soon reach and surpass the number of native speakers of English, and most students will be studying English not to interact with native speakers, but to access information in English and to interact with other non-native speakers.

Recently, in Japan there has been a great increase in the number of English immersion schools called “Super English Language High Schools (SEL-Hi),” and interested parties have started to emphasize the importance of English as a communication tool. In addition, the Japanese government is planning to introduce English classes to elementary schools from 2011. Furthermore, in December 2008, the Ministry of Education, Culture, Sports, Science and Technology in Japan (MEXT) announced a draft version of new teaching guidelines for senior high schools, which would require for the first time that English classes, in principle, be conducted in English. This will be implemented in the 2013 scholastic year. In this paper, based on this current movement, I want to look at the problems of English education in Japan today, mainly focusing on a SEL-Hi and consider what the future of English education in Japan holds for English as an International Language.

## Keywords

SEL-Hi, MEXT

## Introduction

In Japan, the government has promoted English in the national school curriculum. Since 1989, Oral

Communication (O.C.) classes in senior high school English classes have been conducted by a native English speaker or a native English speaker and a Japanese teacher. These classes, which are held only once a week, have not been as successful as the government had planned.

In January 2000, an advisory commission to the late Japanese Prime Minister Obuchi published a report in which it suggested that English should be adopted as the second official language of Japan. In July 2002, with the progress of globalization, the Ministry of Education, Culture, Sports, Science and Technology (MEXT) presented a five-year strategic plan aimed at cultivating “Japanese with English Abilities”<sup>4</sup>. To promote the creation of schools which could serve a leading role in English education, senior high schools focusing on English education have been designated, since 2002 as “Super English Language High Schools (SEL-Hi).” Consequently, “to develop students’ basic practical communication abilities such as listening and speaking, deepening the understanding of language and culture, and fostering a positive attitude toward communication through foreign languages”, a Course of Study for Foreign Languages was announced in 2003.

## 1. SEL-Hi

One of the major changes of the English education system in Japan is the designation of a significant increase in the number of English immersion schools called “SEL-Hi”. Academically, teachers and parents have started to emphasize the importance of English as a communication tool. In SEL-Hi’s, schools are given the resources and permission to prioritize English. In this environment, students are more likely to learn not only English itself but also get a wider array of international experiences and ideas. Let me describe one of the SEL-Hi curricula here (Table 1, see the next page):

---

<sup>1</sup> See McKay (2002)

<sup>2</sup> See Jenkins (2007)

<sup>3</sup> See also Jenkins (2003)

---

<sup>4</sup> See MEXT (2002)

Table 1 Curriculum of a SEL-Hi in Kanagawa<sup>5</sup>

Class name	1st yr	2nd yr	3rd yr
English I	4		
English II		4	
O.C. I	2		
O.C. II		(3)	
International Understanding		2	
International Relations		(2)	
Traditional Culture		(2)	
Reading			4
Writing			3
Step-up English			(3)
Rapid Reading			(2)
Current English			(2)
Total	6	6 (+7)	7 (+7)

NOTE: The classes in parenthesis are electives.

There are about 30 students in a class and all the students learn English depending on their English level. There are 19 mandatory English classes. In addition, there are 14 more electives. This is not a university affiliated school; therefore, elective English classes specifically focus on university entrance exams. In advanced level classes, English is the primary language to be used when teaching English.

Compared to other non-SEL-Hi schools such as a female girls' school which offers many English classes as electives, there are more English classes offered in the former school.<sup>6</sup> The total number of English classes students can take at school is dependent upon the school they attend. For example, in a private boys' school in Tokyo, there are fewer compulsory classes and no electives.

All the students in this SEL-Hi must take the TOEIC BRIDGE<sup>7</sup> every year. This test was developed for beginning level to lower-intermediate level learners of English. In 2005, before they were not designated as a SEL-Hi, the average score in the test including listening and reading was 57.5 and 56.6, making the total average score 114.1 out of 180.

And the average score of the same students who were taught for three years under the SEL-Hi project was 124.4, which means that the students' average score was 10 points higher. On the other hand, according to a 2007 DATA & ANALYSIS

report supplied by TOEIC, the average scores for the TOEIC BRIDGE over the last three years (FY 2005 to FY 2007) were 116.5, 116.4, and 117.1, respectively. Tables 2.1 and 2.2 below show the details.

Table 2.1 SEL-Hi Average Score

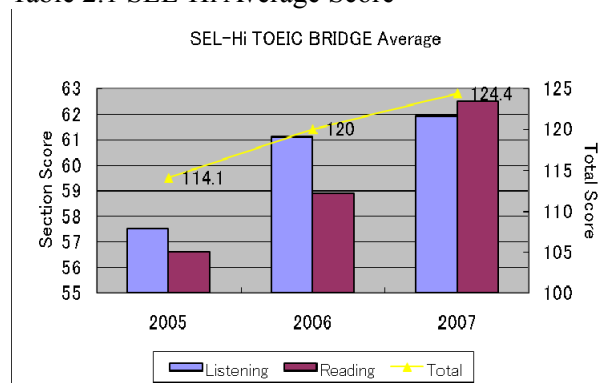
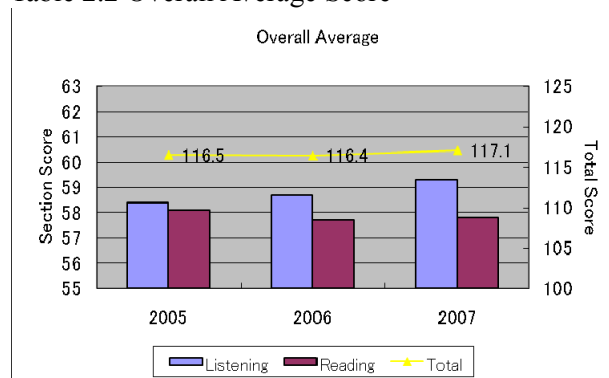


Table 2.2 Overall Average Score



NOTE: Table 2.2 shows all the examinees', including company workers, government and municipal officers, average score.

As you can see from these results above, the average score for this particular SEL-Hi is not much different compared to other examinees. Although we can see that both listening and reading scores did go up, I believe their scores, especially listening, should be much higher since in SEL-Hi, English classes are mainly conducted in English. Therefore, in the next section, based on the data above, I want to think about the prospects for future English education in Japan.

## 2. Prospects for English Education in Japan

The Japanese government has been interested in language policy recently. It is planning to introduce English classes to elementary schools from 2011. Furthermore, in December 2008, the MEXT announced a draft version of new teaching guidelines for senior high schools, which would require for the first time that English classes, in principle, be instructed in English. This will be implemented in the 2013 scholastic year. Altering

<sup>5</sup> Numbers in the table represent classes for week.

<sup>6</sup> See Kumaki 2009

<sup>7</sup> This test was developed to measure the basic communicative English language proficiency of beginning learners.

the Japanese way of English education, which has been primarily focused on grammar, English-to-Japanese translation and Japanese-to-English translation will take a long time, but it is not an unrealistic fantasy. Using English in everyday life in Japan can be very difficult and there is a widespread belief that Japanese people have struggled to master the English language.

Based on the guidelines that the MEXT set, I am going to focus on the methods of teaching English to senior high school students in English in Japan. To begin with, let me compare and contrast the current English curriculum with the future English curriculum guidelines, which start from the year 2013 for senior high schools in Japan (Table 3):

Table 3 New and Current English Curricula

New curriculum (draft)	no. of classes	Current curriculum	no. of classes
Comm. E Basic	2	Oral Comm. I	2
Comm. E I	3	Oral Comm. II	4
Comm. E II	4	English I	3
Comm. E III	4	English II	4
E Expression I	2	Reading	4
E Expression II	4	Writing	4
E Conversation	2		
Total	21	Total	21

This table shows that the number of subjects that the MEXT plans to implement from 2013 is unchanged from the current system. Instead, there are more classes whose titles contain the keyword “communication.”

A listening exam was introduced in 2006 for candidates of the National Centre Test for University Admissions; however at present, teaching styles are still very teacher-centered. As Wray (1999) describes, there are almost no opportunities for students to discuss, provide their opinions, or do group problem solving. It is mainly because English classes at senior high schools focus on helping students to pass the university entrance examinations. The majority of university entrance exams focus on grammar, reading, and writing skills. Furthermore, there are a large number of students in a class. The average student / teacher ratio is 40:1. This ratio does not allow the teachers to interact with the students individually. As a result, while students can leave school able to read and write English to some extent, few end up being able to speak the language.

### 3. Future English Education in Japan

In this section, based on what was discussed thus far, I want to consider what Japanese English scholars and teachers can do for future English education in Japan. The table below (Table 4) indicates some suggestions for English Education in Japan:

Table 4 Suggestions for English Education in Japan

Classroom Language	Besides the grammar points and translation, English should be used more in the school.
Students' Options	Each school should offer more optional English classes which students can select in terms of their own purposes and their English level.
Connection with Other Subjects	Link English classes to other subjects.
Focal Point	In addition to grammar knowledge, communication skills should be encouraged.
Teacher-student Interaction	1) It is essential to reduce class-size. 2) Students should be taught in groups of 15 to 20.
Teachers' English Proficiency	Teachers need to improve their language ability.
Teachers' L1- L2 Carryovers	1) With ALTs, students have opportunities to use natural English with NSs of English. 2) Utilize more Team-Teaching classes.
Government Policy Change	Have long-range targets / goals.
Curriculum	An integrated English curriculum from primary to senior high school is necessary.
University Entrance Examination	Need to alter the dominating influence of written university entrance exams.

(1) In Japan as of today, teaching English entirely in English is a difficult goal to achieve. However, besides grammar points, English can be used more to teach classes.

(2) Besides studying English as a compulsory subject, as some of the schools have already done, from senior high schools, each school should offer more optional English classes which students can select in terms of their own purposes. Currently, it seems to me that there is no connection between subjects. One possible solution to making English meaningful for students is to link it to other subjects, such as history and science in English.

(4) Teachers should be given more space and time to think through improvements in what they do and to be able to engage with students individually. To do that, it is essential to reduce the class size. Teaching about 40 students at the same time is far from satisfactory. Students should be taught in groups of 15 to 20 depending on their levels.

(4) If students want to learn more, there should be a greater range of options. Offering only 3 or 4 English lessons a week is not nearly enough. At such frequency, it is difficult for the students to remember what they learned in the previous lesson, making it difficult for them to feel confident when speaking in or listening to English. Once they leave school (in fact, the English classes), students in Japan do not have many chances to interact in English. Mostly, it is because the majority of Japanese people can live in Japan without using English at all. This status-quo implies that this society offers little or no English support outside of class time to help reinforce what children learn in school.

(5) Expose students to the “native speaker produced” sounds. Scovel (2000) states that the CPH (Critical Period Hypothesis) is a myth; he claims that in the areas of lexicon (or vocabulary) and pragmatics, adults are superior learners to children in a second language (126). He states one advantage younger learners have is in phonology (128). From this point of view, unlike adults, students may be able to achieve “native-like pronunciation”. MEXT addresses that “making use of assistant language teachers (ALTs)” whose native language is English should positively be adopted (MEXT, 2003a). MEXT hired ALTs and reported that approximately 8,400 ALTs worked at public schools between 2001 and 2002.

With ALTs, students have opportunities to use natural English with native speakers of English. Also, ALTs can introduce their cultures by cooperating with Japanese English teachers in classes and may help to motivate students who see language as more than grammar-oriented instruction.

However, there is no requirement of having a TEFL or TESL certificate to be an ALT. I am not saying that having a teaching certificate is essential, but speaking English as a first language should not be sufficient to be a teacher of English. As a matter of fact, it is clearly stated that “TEFL qualification is helpful, but not required” in the Japan Exchange and Teaching (JET) programme website ([http://www.mofa.go.jp/j\\_info/visit/jet/index.html](http://www.mofa.go.jp/j_info/visit/jet/index.html)). Moreover, as the coffers of local governments have

begun to dwindle in recent years, many schools have switched from JET programme ALTs to outsourced cheaper private companies, which makes it even more difficult to find a qualified language instructor in the public and private school system.

(6) In Japan, on the other hand, as for teacher quality, a high level of English proficiency is not necessarily required. Teachers need to have basic grammar knowledge of English, but having skillful conversation abilities is not tested.

MEXT (2003a) would like to upgrade the teaching abilities of English teachers. For example, MEXT (2002) offers domestic training for 2,000 teachers per year and short and long overseas training to approximately 150 people. It also establishes the teaching qualifications: 550 points of paper TOEFL, 730 points of TOEIC. In addition, Kawamura (2004) states that MEXT would like to introduce a teacher certificate renewal system. I believe that this is a crucial step in revitalizing English language education in Japan. Language proficiency is crucial for all language educators.

(7) If more Japanese language teachers were able to speak English at a level where they can conduct a language class mostly in English, team teaching instruction (TT) would be more effective. Unlike native speakers of English, Japanese teachers of English can benefit from drawing on their own experience as language learners. Through careful planning with a native English speaking teacher, collaborative teaching actually saves time by drawing on both individual and collective strengths. English teachers at senior high schools in Japan would teach their classes in English and limit the use of Japanese to the explanation of complicated grammar.

(8) English is going to be an obligatory subject at elementary schools from April 2011. The new classes, officially titled Foreign Language Activities is going to be implemented for fifth- and sixth-graders, or 10- to 12-year-olds, requiring schools to run 35 periods a year.

In addition, the MEXT released the draft of revised curriculum guidelines for senior high schools, which will go into effect from fiscal 2013. The new guidelines have two goals: coping with the diversification of high schools and improving students’ scholastic ability. In the new teaching guidelines it is said that the high school English should be taught primarily in English.

On the other hand, so far there have not been many changes in the junior high school English curriculum. What is necessary for the future English

education in Japan is not the separated educational system but the integrated one.

(9) Although the National Centre Test for the University Admission implemented the Listening test to their exam from 2006, oral tests are mostly absent. Unless the current university entrance exam is modified to a more communicative test, the focal point of the school curriculum will not be changed.

## Conclusion

English education in Japan is in transition. Everyone knows the importance and necessity of English; however there are very few people who can master and use English effectively after learning it for many years in school, even after the Listening Test was implemented in the National Centre Test in 2006. The focal point of English education in Japan is to memorize as many sentences and words as possible for the university entrance exam. But the essential point of learning English is to use it as a tool to communicate with others.

Learning from what the Japanese government has done so far, there are still steps that can be taken in the Japanese education environment. First of all, Japanese teachers of English need to improve their English speaking ability. Secondly, students should be given more opportunities to immerse themselves in an English environment. Each school should upgrade their teaching system and offer students more opportunities to use English. Teaching English in English can be one solution for that. For those who want to further their English studies, there should be more diverse elective English classes.

Learning English in English could broaden their perspectives, but I feel it is pointless to teach in English if the format of university entrance examinations remains unchanged. Unless the requirements of entrance exams are changed, in fact unless the whole system of English education is changed, I do not see the point of trying to teach in English.

An integrated English curriculum from primary to senior high school is necessary to produce students who can speak English.

## References

- Funabashi, Y. (2000). *Aete eigo koyogo-ron*. (Tokyo: bungei Shunju).
- Jenkins, J. (2007). *English as a Lingua Franca: Attitude and Identity*. Oxford: Oxford University Press.
- Jenkins, J. (2003). *World Englishes: A resource book for students*, Routledge.
- Kachru, B. (2005). *Asian Englishes Today: Asian Englishes Beyond the Canon*. Hong Kong: Hong Kong University Press.
- Kachru, B. (1992). 'Teaching world Englishes', in B. Kachru(ed.) *The Other Tongue, English across Cultures*, 2nded. University Illinois Press.
- Kawamura, T. (2004). Reforming compulsory education. Ministry of Education, Culture, Sports, Science, and Technology. Retrieved October 21, 2004, from: <http://www.mext.go.jp/english/topics/04091701.htm>
- Kubota, R. (2002). The impact of globalisation on language teaching in Japan, in David Blocks and Deborah Cameron (eds), *Globalisation and language teaching*, pp.13-28 (London: Routledge).
- Kumaki, H. (2009). *Future Prospects for English Education in Japan: Based on the Language Policy in Singapore*. Paper presented at the meeting of International Cultural Expression Studies, Shizuoka, Japan.
- McKay, S. (2002). *Teaching English as an International Language*. Oxford: Oxford University Press.
- Ministry of Education, Culture, Sports, Science and Technology (MEXT) (2003a). *Regarding the establishment of an action plan to cultivate Japanese with English abilities*. Retrieved June 28, 2009, from: <http://www.mext.go.jp/english/topics/03072801.htm>
- Ministry of Education, Culture, Sports, Science and Technology (MEXT) (2003b). *The course of study for foreign languages*. Retrieved June 28, 2009, from: <http://www.mext.go.jp/english/shotou/030301.htm>
- Ministry of Education, Culture, Sports, Science and Technology (MEXT) (2002). Developing a strategic plan to cultivate Japanese with English abilities: Plan to improve English and Japanese abilities. Retrieved March 21, 2009, from: <http://www.mext.go.jp/english/mews/2002/07/020901.htm>
- Modiano, M. (1999a). 'International English in the global village', *English Today*, 15/2, pp22-34.
- Modiano, M. (1999b). 'Standard English(es) and educational practices for the world's lingua franca', *English Today*, 15/4, pp 3-13.
- Scovel, Thomas. (2000). "'The Younger, the Better' Myth and Bilingual Education," Roseann D. Gonzalez (ed.), *Language Ideologies: Critical Perspectives on the Official English*

- Movement, vol.1. New Jersey: Lawrence Erlbaum Association, Inc. 114-136.
- Smith, L. (1976). English as an International Auxiliary Language. *RELC Journal* 7/2: 38-53
- TOEIC Bridge DATA & ANALYSIS 2007. Retrieved June 23, 2009, from [http://www.toeic.or.jp/toeic\\_en/pdf/data/TOEIC\\_Bridge\\_DAA2007.pdf](http://www.toeic.or.jp/toeic_en/pdf/data/TOEIC_Bridge_DAA2007.pdf)
- Wray, H. (1999). Japanese and American education: Attitudes and practices. Westport, Connecticut: Bergin & Garvey.

## Bilingual Education: What can Malaysia learn from the US Experience?

Juliana Binti Othman

Faculty of Education, University of Malaya  
juliana@um.edu.my

### Abstract

This paper provides a brief comparative perspective of the bilingual education policy and practices in the United States and Malaysia. Particular issues examined include the following: the social context of bilingual education in the two countries, challenges faced by teachers, their classroom practices and support provided. A key assumption was that such a comparison could provide us with a better understanding of conditions and factors that might make for more effective bilingual education practice in the Asian context.

### Introduction

In many countries, bilingualism is implemented through the education system which gives rise to varied forms of bilingual education. Reasons for the implementation of bilingual education may vary from one country to another. In the United States, bilingual education has been implemented to meet the educational needs of children from immigrant families. In Malaysia, the recent change in language policy in Mathematics and Science education is seen as an effort to meet challenges posed by globalization. In this paper, I would compare the bilingual education policy and practices in the United States and Malaysia. Particular issues examined include the following: the social context of bilingual education in the two countries, challenges faced by teachers, their classroom practices and support provided. It is my hope that such a comparison could provide us with a better understanding of conditions and factors that might make for more effective bilingual education practice in the Malaysian context.

This paper describes briefly the bilingual education in the United States and the different types of programs offered. The paper then presents the background on the

Malaysian language policy and its impact on the education system. The final part of the paper discusses the challenges faced by teachers and their classroom practices.

### *Bilingual Education in the United States*

In the USA, both English as a second language and bilingual education programmes have been historically and politically connected to immigration and the integration of immigrants into a predominant English-speaking environment. One of the earliest bilingual education programmes in the USA began during the 1840s as a result of non-English speaking immigrants entering the country. However, in the late 19<sup>th</sup> and early 20<sup>th</sup> centuries, restrictive laws prohibiting instruction in other than English brought this educational practice to a halt due to anti-immigrant sentiment. However, renewed interest in bilingual education developed with the 1960s civil rights movement. In 1968, Congress passed the Bilingual Education Act, also known as the Title VII of the Elementary and Secondary Act. Under the act, English language learners must be kept in an adequate program until they can read, write and comprehend English well enough to participate meaningfully in mainstream classes.

Bilingual education in the United States is manifested in a variety of programs. All bilingual program models use the students' native language, in addition to English, for instruction. Students in bilingual programs are grouped according to their first language, and teachers must be proficient in both English and the students' native language. The most popular form of bilingual education is the *transitional bilingual education program*. In this program, the medium of instruction is in both the native language and English for a period of three years. The levels

of instruction in English increase while the native language decreases and students are exited from the programme when they reach a given criteria. The main goal of this model is to mainstream students to English-only classrooms.

Another favoured program in the United States is the *Structured English Immersion*. In this program, English is taught through the content areas. Structured immersion teachers have strong receptive skills in their students' first language and have a bilingual education or ESL teaching credential. The teacher's use of the children's first language is limited primarily to clarification of English instruction. Most students are mainstreamed after 2 or 3 years.

In *two-way bilingual programs*, language minority students from a single language background are placed in the same classroom with language majority (English-speaking) students. Instruction is provided in both English and the minority language. In some programs, the languages are used on alternating days. Others may divide the use of the two languages by academic subject. Native English speakers and speakers of another language have the opportunity to acquire proficiency in a second language while continuing to develop their native language skills. Students serve as native-speaker role models for their peers. Two-way bilingual classes may be taught by a single teacher who is proficient in both languages or by two teachers, one of whom is bilingual. The goal of this program is to develop proficiency in the students' native language (L1) and in a second language (L2).

### ***The Malaysian Context***

Malaysia has a different language-in-education policy than the one in the United States. The Malaysian Education system has undergone major changes in the last few decades. Malaysia has a long tradition of English medium education as a result of the British colonial influence. Prior to the Education Enactment Bill (1971), English was the primary language and Malay, the

secondary language. With the passing of the Education Enactment Bill (1971), English medium schools were phased out and Malay became the medium of instruction in all national schools, and of university education by 1983. This change in language policy has been driven by a nation-building agenda. It has been argued that through the medium of a common language national unity would be fostered.

However, 20 years after the implementation of Malay as a medium of instruction in schools, concerns were expressed about the English language proficiency of school leavers and graduates entering the workforce. In Malaysia, English language is often associated with economic competitiveness and employment. More than 44,000 public university graduates were unemployed largely due to their poor command of English (The Straits Times, Singapore, 2 June 2002). The root of this situation could be attributed to the fact that when a language like English is not fully used as a medium of academic discourse, the students' mastery of it cannot be expected to attain a level it would have reached had the language concerned be used as a medium of instruction (Asmah Hj Omar, 1996). Although English was retained as a second language in the national education system, it is only viewed as a 'subject' in the school curriculum as the rest of other subjects, with the exception of languages, is conducted in Malay.

As the Malaysian government aspires to achieve developed nation status by the year 2020, a large-scale national-level policy initiative involving a shift in the language of mathematics and science instruction in national schools from Malay to English has been carried out. This policy shift is a "pragmatic response to the inevitable challenges of technological advances and globalisation, and linked to the fear in the falling standards of English" (Martin, 2005). Hence, in January 2003, Malaysia implemented a major change to its education policy whereby the English language is to be introduced as the medium of instruction for Mathematics and Science. This change in

language policy had an impact on bilingual needs and practices in the education system.

The Malaysian bilingual education program can be categorized as an “enriched second language program” (Genesee, 1987). According to Genesee (1987), “at least 50 percent of instruction during a given academic year must be provided through second language for the program to be regarded as immersion”. Unlike in the United States where the second language is used as a medium of instruction beginning in Grade 3, it starts as early as Primary 1 (Grade 1) in the Malaysian education system. Only Mathematics and Science subjects are taught through a different medium of instruction, i.e. English. Students are expected to gain English language proficiency after 11 years of instruction at primary and secondary level.

In most bilingual programs in the United States, participation is by parental choice, or, as in the case with late immersion, by parental and student choice. Unlike in the United States, bilingual education in Malaysia is prescribed by the government. The education system is centrally controlled by the Ministry of Education where the curriculum is designed to have coherence, balance, progression and continuity. Thus, students are all at the same academic level—usually progress as a cohort group beginning with little or no L2 proficiency. The compulsory nature of the program means that the parents of the students are usually well informed about educational issues, and are generally committed to the importance and relevance of learning English as a second language.

In Malaysia, English is considered as a second language which is seen as prestigious and recognized as valuable by the community. In many parts of Malaysia, proficiency in the English language depends largely on geographical location and socio-economic status. In rural areas, students make little use of English outside school where limited opportunities are available.

In contrast, English is the language of the majority of the population and the medium of instruction in schools in the USA. Students’ native language may be given an inferior status. Students are expected to gain

proficiency enough to enter mainstream classes in one year. English language learners must be provided appropriate help to overcome language barriers until they attain academic achievement equivalent to average native English speakers. The focus is on learning the target language through content teaching rather than on teaching the language.

## *Challenges*

### *Malaysian experience*

The implementation of English as a medium of instruction for Mathematics and Science in the Malaysian context is not without its challenges and problems. One of the challenges faced by the teachers is their lack of competence in English (Pandian and Ramiah, 2004; Chan and Ain Nadzimah, 2005). In addition, these teachers also face the challenge of managing the language development of their students with respect to their subject matter. Teachers have to cope with the double demand of grappling with the language and transmitting content in a new medium of instruction. However, the prevailing language support mechanisms do not meet their needs. Most Mathematics and Science teachers in Malaysia did not have specialized training in L2 methodology. Therefore, the need for effective language enrichment teacher development in Malaysia is crucial. In order to successfully implement the new language policy, it is particularly important to have highly trained teaching staff. Well-trained English medium teachers would have a significant influence on the cognitive and affective growth of students whose primary language is other than English.

One of the efforts that has been carried out by the Ministry of Education is to organise a course known as ETeMS (English for Teaching Mathematics and Science), for teachers of Science and Mathematics. This was the first support step in helping these teachers cope with the new medium of instruction. Nevertheless, it was only an urgent interim measure to ensure that these teachers have some basic capacity to use English as the medium of instruction

(Facilitators Notes, ETeMS Module, 2002). In addition, each school was supplied with a teachers' kit and learning kit consisting of multimedia software in the form of compact discs (CDs) to help teachers improve their language proficiency and help them teach their students in class.

Apart from that, a programme called the "Buddy System" was introduced. Under this programme, mathematics and science teachers are encouraged to meet regularly with a peer who is proficient in the English language, also known as a 'critical friend' in order to strengthen their command of the language. This was deemed important as exposure to the language and feedback on their attempt to use the language is fundamental to the development of their proficiency in the language (Hedge, 2000). In addition, teachers are also expected to educate themselves to a large extent by meeting regularly with their colleagues teaching science in order to help them improve their proficiency in the language of science as well as teaching methodology.

The need for sustained teacher development activities in this context is crucial. Governing curricula bodies such as the Curriculum Development Center, State Education departments and teacher educators must provide further supporting programs to these teachers to acquire the linguistic tools and content specific instructional strategies to help them teach effectively.

### *US experience*

One of the main concerns regarding bilingual education in the US is the lack of qualified teacher. Over the last five years, many EL teachers had little or no professional development designed to help them teach these students, and the quality of training was uneven. Although the number of ESL students in US public schools has increased dramatically in the past decade, teacher education programs have not greatly adapted their curricula to accommodate these demographic shifts. Thus, teachers of ESL students have not received the appropriate training necessary to work with these students.

Along with the availability of qualified teachers comes the question of availability of educational materials and resources. Teachers were challenged by the lack of tools to teach, including appropriate assessment materials and instruments. Many classrooms lack textbooks written in a way that made the material accessible to ELLs and culturally relevant teaching materials (Yeh et al., 2002; Gandara et al, 2005).

Another great challenge faced by teachers in bilingual programs is the varying language proficiency abilities among their English language learners. Teachers expressed frustration with the wide range of English language and academic levels often found in their classrooms. Trying to address the different academic needs of learners in the same classroom can create daunting tasks for the teachers.

In addition, teachers found difficulty to communicate with students' families and communities. Teachers' inability to speak the parents' language and parents' inability to help their children with homework often limits the support that can be given to the students.

However, the US bilingual programs do benefit from the vast number of research and years of experience in implementing a variety of bilingual programs. Among the successful classroom practices utilized are the content based instruction, Sheltered Instruction and Cognitive Academic Language Learning Approaches.

### *Content-Based Language Instruction.*

In this approach, bilingual teachers use instructional materials, learning tasks, and classroom techniques from academic content areas as the vehicle for developing language, content, cognitive, and study skills. The second language is used as the medium of instruction for mathematics, science, social studies, and other academic subjects. Instruction is usually given by a language teacher or by a combination of the language and content teachers.

### ***Sheltered Subject Matter Teaching.***

This approach involves adapting the language of texts or tasks and use of scaffolding strategies (demonstrations, visuals, or cooperative work) to make instruction more accessible to students of different English proficiency levels. This type of instruction is given by the regular classroom or content teacher, or by a language teacher with special expertise in another academic area (Brinton, Snow, & Wesche, 1989).

***Sheltered Instruction.*** Here, a content curriculum is adapted to accommodate students' limited proficiency in the language of instruction. This model was originally developed for elementary foreign language immersion programs to enable some portion of the curriculum to be taught through the foreign language (Genesee, 1987). It is commonly used in immersion and two-way bilingual programs (Met, 1991) and has been adapted for use in second language programs with large numbers of limited English proficient students of intermediate or advanced English proficiency.

***Cognitive Academic Language Learning Approach (CALLA).*** This approach integrates language, content, and learning strategy instruction into a transitional ESL approach for students of intermediate or advanced English proficiency (Chamot & O'Malley, 1987). CALLA emphasizes active learning, in which students are given the skills and opportunities to take an active role in their own learning.

### **Summary and Conclusions**

Analyses of the challenges and experiences of bilingual teachers in both countries provide strong implications for professional development. For example, teacher development efforts must address how to seek effective ways to accommodate varying students' academic experiences and language abilities.

Teacher training endeavours may be tailored to discuss many specific pedagogical strategies and practices when working with bilingual students, such as how to facilitate

students' simultaneous acquisition of academic content and English language and literacy.

Staff development programs could offer teachers encouraging support, such as providing them with appropriate teaching material and resources, promoting collaboration among teachers as well as with parents, and developing techniques for teaching the English language more effectively and efficiently. These training efforts will not only enhance teaching practice and knowledge, they will also provide teachers with the appropriate support and services necessary to improve the quality of bilingual education.

Parental involvement is instrumental in contributing to the academic success of bilingual students. Support and encouragement of several people including teachers and parents could help students to succeed in their academic program. However, the language barriers of the parents could be the reason that hinders them from getting involved in their children's learning. Thus, bilingual teachers are in a unique position of being able to communicate with these parents in their native language about the importance of parental involvement and the ways in which they can participate in their children's education.

To summarise, immersion programmes were originally established to fill a need in the United States society for better English as a second language programmes. Assuming similarities between first and second language learning in children, parents and educators believed that by providing young children with a rich and comprehensible input of the target language in school, high levels of proficiency in that language would result.

Of course, Malaysian's solutions will be different because its starting point is different. However, the cycle of discovery of successful approaches to second language teaching might well be the same. Like any innovation, it takes time and a lot of work remains to be done to ensure the success of the policy implementation.

*Practice* (pp. 75-99). Clevedon : Multilingual Matters Ltd.

## References

- Asmah Hj Omar (1996). Nationism and exoglossia : The case of English in Malaysia. In Abdullah Hassan (Ed.), *Language planning in Southeast Asia*. Selangor: DBP.
- Brinton, D.M., Snow, M.A., & Wesche, M.B. (1989). *Content-based second language instruction*. New York: Harper & Row.
- Chamot, A.U., & O'Malley, J.M. (1987). The cognitive academic language learning approach: A bridge to the mainstream. *TESOL Quarterly*, 21, 227-249.
- Chan, S.H. and Ain Nadzimah, A. (2005) After an English Language Policy Initiative: Listening to the Voices of Classroom Implementers. *Pan Pacific Association of Applied Linguistics* 9 (1), 139-149.
- Gandara, P., Maxwell-Jolly, J., and Driscoll, A. (2005). Listening to Teachers of English Language Learners: *A Survey of California Teachers' Challenges, Experiences, and Professional Development Needs*. The Regents of the University of California. The Center for the Future of Teaching and Learning.
- Genesee, F. (1987). *Learning through two languages: Studies of immersion and bilingual education*. Cambridge, MA: Newbury House.
- Hedge, T. (2000). *Teaching and learning in the language classroom*. Oxford: Oxford University Press.
- Martin, P. (2005) Safe Language Practices in Two Rural Schools in Malaysia: Tensions between Policy and Practice. In A. Lin and P. Martin (eds) *Decolonisation, Globalisation: Language-in-Education Policy and Practice* (pp. 75-99). Clevedon : Multilingual Matters Ltd.
- Met, M. (1991). Learning language through content; learning content through language. *Foreign Language Annals*, 24, 281-95.
- Ministry of Education. (2002). *English for teaching mathematics and science (ETeMS) facilitator's notes*. English Language Teaching Centre, Teacher Education Division.
- Pandian, A. & Ramiah, R. (2004). Mathematics and science in English: Teacher voice. *The English Teacher*, XXXIII, 50-61
- Yeh, C. J., Chen, J., Kwong, A., Chiang, L., Wang, Y., and Pu-Folkes, F. (2002). Educators of Asian Bilingual Students: Pedagogical Techniques, Strategies and Challenges *Journal of Multilingual and Multicultural Development* Vol. 23, (4), 296-315.

# Suggestion for Teaching English Based on the Concept of Internalization

Naomi Kakihara

Hosei University

naomiruri.msn.com

## Abstract

In this paper we try to clarify the concept of internalization and indicate the gap between teaching and learning. Two adult classes and primary fourth, fifth and sixth graders participated the researches. We assume the results indicate the students' inner knowledge.

## Keywords

internalization, interlanguage internal knowledge

## Introduction

Teachers are always concerned with the result of teaching. It is not often the case that teachers are satisfied with the result. Some may regret the materials they chose. Another may complain the surroundings such as class size, hours of instruction, variety of the academic ability of students and so forth. These factors are indispensable to teaching. Furthermore, we must not miss the concept of internalization. To begin with, we define the idea of internalization and try to show the result of the internalization.

## 1 Internalization

Piaget (1970/2007) classifies people's conduct in two ways, actions in narrow sense and other actions such as blinking and yawning. Actions in a narrow sense includes two different actions, internal action (=thinking) and external action (=behavior). According to Piaget, thinking is construed as internalized external action. When people can do one thing without any help, those actions are assumed to be internalized.

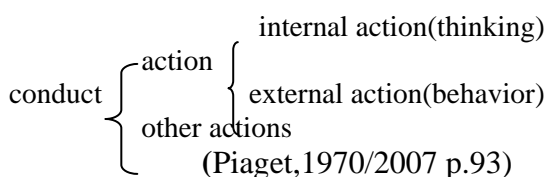


Figure 1: analysis of conduct

## 1.1 Interlanguage

Selinker(1972) introduced the term 'interlanguage' which is used to refer to both the internal system and to the series of inter connected system (Ellis, 1994). If the interlanguage shows the learners' internal knowledge, the result of the same test may show the different state of internalization.

## 1.2 Basic Questions

Two adult groups (foreign students-A , university students-B) answered fifty basic questions based on the grammar in the authorized textbook for junior high first graders. A are from China and studied English for one year as beginners using the authorized textbook. The result of A indicates that forty-five- hour instruction for one year is not enough to have a good foundation in English (Figure 1). Meanwhile, it seems to be natural that almost all B students could get full marks (Figure 3).

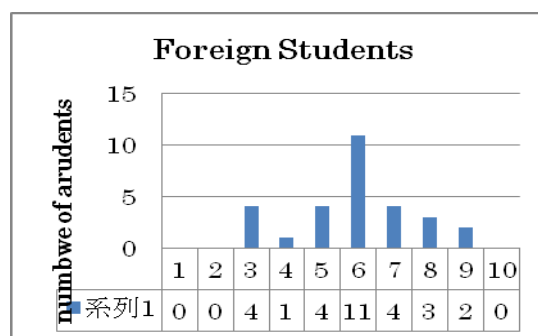


Figure 2: Frequency Distribution-A

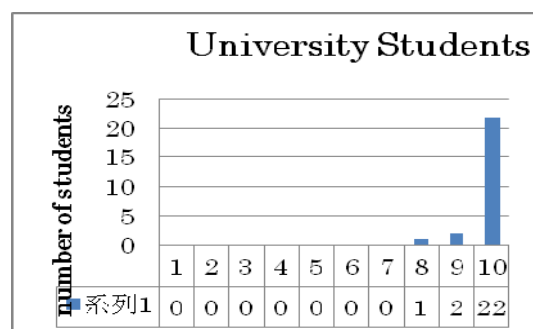


Figure 3: Frequency Distribution-B

## 2 English Quiz

To compare primary students' inner knowledge of English, fourth, fifth and sixth graders answered more than forty questions. According to the homeroom teachers, they do not learn how to read at school. Therefore those questions can be answered easily if they are able to read words in English. Others are listening questions. They are from two different tapes based on two different picture books which are not familiar with them. The purpose of the quiz is to examine how the students respond to unfamiliar factors.

### 2.1 Reading

The first question is to fill the blanks with capital letters alphabetically. Next question is to match capital letters to small letters. Others are to match pictures and words. Figure 4 indicates that the students increase the scores (fourth graders-23.14, fifth graders-26.3, sixth graders-32.48) as they get older. Compared to the differences in detail, the students could answer better as they get older (Figure 5).

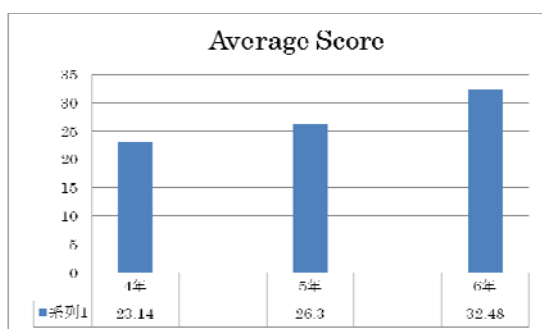


Figure 4: Average Scores-primary Students

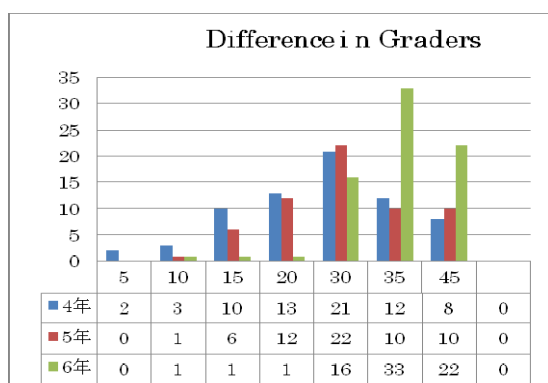


Figure 5: Frequency Distribution-primary students

### 2.2 Listening

Listening questions are to check whether or not they can catch the words in the story. The first one is to number the word along the narrative line. The

second one is to choose the word that is not included in the story. Tapes they listened were made for native speakers of English. They were not modified for learners. It is likely that difficulty is similar both for fifth graders and sixth graders (Figure 6).

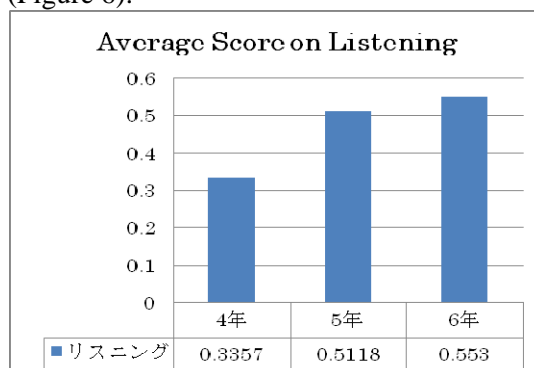


Figure 6: Average Score on Listening

## 3 Discussion

The results of basic questions show that repeated practice is essential. Even though teachers and students think they finished certain factors, it does not mean those factors are not internalized. In this respect, A-students seem to require more time to realize what they learned.

The results of primary students indicate that there are differences among different ages. Although students are not familiar with questions, the number of sixth graders who got better scores is bigger than younger students. One of the reasons may be the amount of time of English lessons. Older students have more time to be exposed to English not only at school but at their homes. The other reason may be due to their development. Children's behaviors are influenced by their cognitive development. In this respect, older students can understand the meaning of the questions better and guess what they are supposed to do.

## 4 Suggestions

In this paper, we reveal that immediate understanding does not always lead to true learning. On the other hand, primary students tell us that they acquire something that they do not learn explicitly. The results suggest that teachers are supposed to notice the concept of internalization and try to find alternative way of teaching

### 4.1 References

Conform to the styles of the Publication Manual of the American Psychological Association. Citations in the text appear in parenthesis as (Author, year) or

(Author, year: page). If the author's name appears in the text, as Author (year) or Author (year: page).

Full citation of literature referred to should be given in **References**. Arrange the references alphabetically by first author's name, rather than by the order of occurrence in the text. Punctuate and capitalize as in References of this document.

## 4.2 Appendices

Any appendix should be appeared directly after the reference, and should be given an informative title:

**Appendix A. Title of Appendix A.**

## References

Ellis, Rod. (1994) *The Study of Second*

*Language Acquisition*. Oxford University Press.

Piaget, J. (1970). Piaget's theory. P.H. (Ed.). *Carmichael's manual of child psychology (3rd ed.)*: Vol.1. N New Yoek: John Wiley & Sons.

Slinker, L. (1972). 'Interlanguage' *International Review of Applied Linguistics* 10: 209-31.

Nakagaki, A. (2007). Translation of *Piaget's Theory*. Kitaoji Shobo.

## Appendix A

### 基礎問題

名前:

2. 次の文の( )のなかから正しい語を選び、その語を○でかこみなさい。

- (1) You (aren't, don't, doesn't) like coffee.
- (2) Mike can (play, plays, playing) basketball.
- (3) Do you (get, gets, getting) up early?
- (4) Does she (get, gets, getting) up early?
- (5) (Is, Does, Are) your mother a teacher?
- (6) This is (I, my, me) bag.
- (7) (Me, We, Our) study English.
- (8) (You, His, He) name is Jack.
- (9) I (do, does, am) playing soccer.
- (10) She's (cooks, cook, cooking) now.

## Appendix B

### 英語クイズ

年\_\_\_\_\_組\_\_\_\_\_

1) アルファベットが順にならんでいます。抜けているところに【 】の中から順番にふさわしい文字を選び ( ) の中に書き入れましょう。

A B C ( ) E F G ( ) I J K ( ) M N O ( ) Q R S ( ) U V W X Y Z  
[H D T L P]

5) テープの英語を2回聞きましょう。絵をよく見て、聞こえた話にふくまれる言葉の順番を(1～6)の中に書き入れましょう。



( )



( )



( )



( )



( )



( )

# **The Development and Implementation of Task-based Writing Performance Assessment for Japanese Learners of English: (3) Main Experiment 1**

Yoshihito Sugita

Faculty of Nursing, Yamanashi Prefectural University

sugita@yamanashi-ken.ac.jp

## **Abstract**

This research examines the main data of the second task-based writing performance test. The same task-based writing test used in the pre-testing was conducted on the participants comprised of 20 university students. Each of forty scripts was scored by five identical teacher raters who displayed acceptable levels of consistency in the pre-testing.

Along with the findings of the pre-testing, the results suggest that the students ability was effectively measured using these tasks and raters. The FACETS analysis showed that the difficulty of the two tasks and the impressionistic scoring were equivalent. The interrater correlation coefficients between pairs of raters were high, and the raters displayed acceptable levels of consistency with themselves. There were, however, significant differences between raters in terms of severity. A bias analysis research was conducted in rater-subject interactions and in rater-task interactions. These analyses indicated that three of the five raters were significantly biased towards certain types of subjects, and these raters' bias patterns were unique. The equivalence of task difficulty may indicate that task development based on the construct-based processing approach could be reliable and valid to estimate students' writing ability. The modified scales associated with the five rating categories and their specific written samples were shown to be mostly comprehensible and usable by raters, and demonstrated acceptable fit. However, there is still room for argument about the reliability and validity of assessment tasks and rating scales.

## **Keywords**

writing performance, task-based assessment, bias analysis

## **Introduction**

In Japan English language has been traditionally taught with a focus on accuracy, and indirect measurement is widely used in the field of assessment. There seems to have been a paradigm shift from accuracy-oriented to fluency-oriented writing instruction, but no significant changes have occurred in assessment of writing. Judging from the present state of teaching and assessing writing in Japan, it would be meaningful to develop scoring procedures for writing performance assessment in place of traditional indirect tests of writing. This study is motivated by such an urgent need for improved assessment of writing, which is conducted in order to develop a task-based writing test for Japanese learners of English.

## **1. The Study**

### **1.1 Purposes and research questions**

In order to examine the degree of reliability and validity of the task-based writing performance test, the following are focused on: raters' severity, interactions with writers' abilities and task difficulties, the reliability of elicitation tasks and rating scales, and the measure's validity.

The specific research questions are as follows:

- 1) Is student ability effectively measured?
- 2) Are teacher-raters equally severe?
- 3) How much do tasks that are designed to be equivalent actually differ in difficulty?
- 4) How well do scales conform to expectations about their use? Do raters use all parts of them, and use them consistently?
- 5) Do individual raters score a particular group of subjects more harshly or more leniently? If so, what are the sub-patterns of ratings in terms of rater-subject interaction for each rater?
- 6) Do the raters score particular tasks more harshly or more leniently than others? If so, what are the sub-patterns of ratings in terms of rater-task interaction for each rater?
- 7) To what extent, statistically, is the task-based

writing test a reliable and valid measure?

## 2.2 Test participants and materials

The data for this study were 40 scripts (20 scripts for each of two tasks) collected from 20 undergraduate students (6 males and 14 females) who took an English teaching methodology course in the first semester of 2008. The subjects were 14 second-year and 6 third-year students from the faculty of global policy management and communications. All of the subjects were native speakers of Japanese with an intermediate level of English language proficiency.

The TBWT was conducted in the computer-assisted learning room. The time limits for Task 1 and Task 2 were 20 minutes and 10 minutes, respectively. After finishing the tasks, each student was required to submit an essay using a web-based essay evaluation service, *Criterion*. They had to finish writing the essay within 30 minutes. The prompt was as follows: People attend college or university for many different reasons (for example, new experiences, career preparation, increased knowledge). Why do you think people attend college or university? Use specific reasons and examples to support your answer.

## 2.3 Scoring materials and procedure

In order to give raters a shared understanding of the construct of writing ability as defined by the test writers and reduce the differences or biases caused by variation among raters, the TBWT scoring guide was edited for this testing. The first section is the background of the TBWT. The second section is the explanation of assessment tasks. The third section is the implementation method of the testing. The fourth section is comprised of the rating scales and written samples accompanied by detailed commentary on each sample at five levels, 1-5.

Each of the forty scripts was scored by five raters, who were all experienced Japanese high school teachers of English. They were all native speakers of Japanese, and they shared similar backgrounds in terms of qualifications of ten or more years of teaching experience. They displayed acceptable levels of consistency with themselves in the pre-testing conducted in January, 2008. Both scripts and scoring guidelines were given to the raters by mail at the end of July, 2008. Each of the five raters rated the entire set of forty scripts and sent them back by the end of August, 2008. They were instructed to rate the 20 scripts of Task 1 first, and then to rate the 20 scripts of Task 2. Finally, they were asked to rate each of the participants' writing proficiency based on the total impression at five levels, 1-5.

## 2.4 Data analysis

Tables 1, 2 and 3 show the descriptive statistics for the scores of the two test tasks and the impressionistic scoring. Since the average of the inter-rater coefficients for each scoring is relatively high (0.76, 0.83, 0.81), the five raters appear to be of acceptable reliability.

Table 1: Descriptive statistics of scoring Task 1

Rater	1	2	3	4	5
Mean	3.20	3.10	3.50	2.85	2.40
SD	0.87	1.30	1.00	1.19	0.91
Min.	2.0	1.0	1.0	1.0	1.0
Max.	5.0	5.0	5.0	5.0	5.0

Table 2: Descriptive statistics of scoring Task 2

Rater	1	2	3	4	5
Mean	2.95	3.10	3.05	2.65	2.85
SD	0.97	1.37	1.20	1.19	0.96
Min.	1.0	1.0	1.0	1.0	1.0
Max.	5.0	5.0	5.0	5.0	5.0

Table 3: Descriptive statistics of impression

Rater	1	2	3	4	5
Mean	3.10	3.20	3.25	2.85	2.50
SD	0.88	1.20	1.13	1.23	1.57
Min.	2.0	1.0	1.0	1.0	1.0
Max.	5.0	5.0	5.0	5.0	5.0

Table 4: Descriptive statistics of different scoring

	Task 1	Task 2	Impression	Criterion TWE
N	100	100	100	20
Mean	3.01	2.92	2.98	2.40
SD	1.12	1.16	1.14	0.94
Min.	1.0	1.0	1.0	1.0
Max.	5.0	5.0	5.0	4.0

Table 4 reports results for each test task, the impressionistic scoring and the scores of *Criterion*, including its mean and standard deviation. The mean scores for all variables are very close, ranging from 2.92 to 3.01. The alpha coefficients for the test tasks and the impressionistic scoring were calculated. Using Davies' cut-off (.90) as an acceptable level of internal consistency on a high-stakes test, each Cronbach's  $\alpha$  would meet the point: .9386, .9582 and .9570 for Task 1, Task 2 and impressionistic scoring, respectively.

The correlation coefficients between the scores provide a preliminary estimate of the parallel-form reliability of each test task. As seen in Table 5, the correlation coefficients between each task and the impressionistic score fall in a range of .778 to .910, which are all significant at the 0.01 level. The correlation between the two test tasks (.778) is, however, slightly lower than the established

estimate of reliability (.80). Table 6 also shows that the two tasks and impressionistic scoring correlate positively with the scores of *Criterion* ( $p < .01$ ). The highest correlation is between the *Criterion* score and Task 2 ( $r = .708$ ), followed by that between the *Criterion* score and Impression ( $r = .703$ ) and finally, between the *Criterion* score and Task 1 ( $r = .621$ ).

Table 5: Pearson correlation coefficients

	Task 1	Task 2	Impression
Task 2	.778**		
Impression	.910**	.903**	
<i>Criterion</i>	.621**	.708**	.703**

Note: \*\*all correlations significant at 0.01 level.

There is a possibility that the test data can be influenced by errors of measurement resulting from variation in rater harshness and test tasks, as well as by the nature of the rating scale used and by the range of ability of the subjects who are being assessed. Therefore, it was necessary to use statistical models which take into account all of the factors that might affect a student's final score.

The analyses for the present study were done using FACETS version 3.63 (Linacre, 2008). To examine the measurement characteristics of this testing, the data was specified as having three facets, namely, the ability of the subjects, the difficulty of tasks and the severity of raters. The partial-credit model was chosen because the scoring criteria for the rating scales were qualitatively different.

### 3. Results

#### 3.1 FACETS summary

Figure 1 shows a summary of all facets and their elements. They are positioned on a common logit scale, which appears as "measure" in the first column. The second column shows the severity variation among raters. The most severe rater (ID: 3) is at the top, and the least severe rater (ID: 5) is at the bottom. The third column shows the ability variation among the 20 subjects. The subjects are ranked with high ability at the top (ID: 9) and low ability at the bottom (ID: 11). The fourth column shows the difficulty variation among tasks. The most severely scored task (Accuracy) is at the top and the least severely scored task (Communicability) is at the bottom. The last three columns graphically describe the three rating scales. Each of the two tasks and the impressionistic scoring has their own scale. The most likely scale score for each ability level is shown.

#### 3.2 FACETS analysis

##### 1) Is student ability effectively measured?

As shown in Figure 1, subject ability estimates

range from a high of 5.94 logits to a low of  $-5.92$  logits, indicating a spread of 12 logits in terms of students' ability. Subject separation value was 6.85, meaning that populations like the students in this study can be spread into about seven levels. The reliability index was .98, which demonstrates it is possible to achieve reliable ability scores.

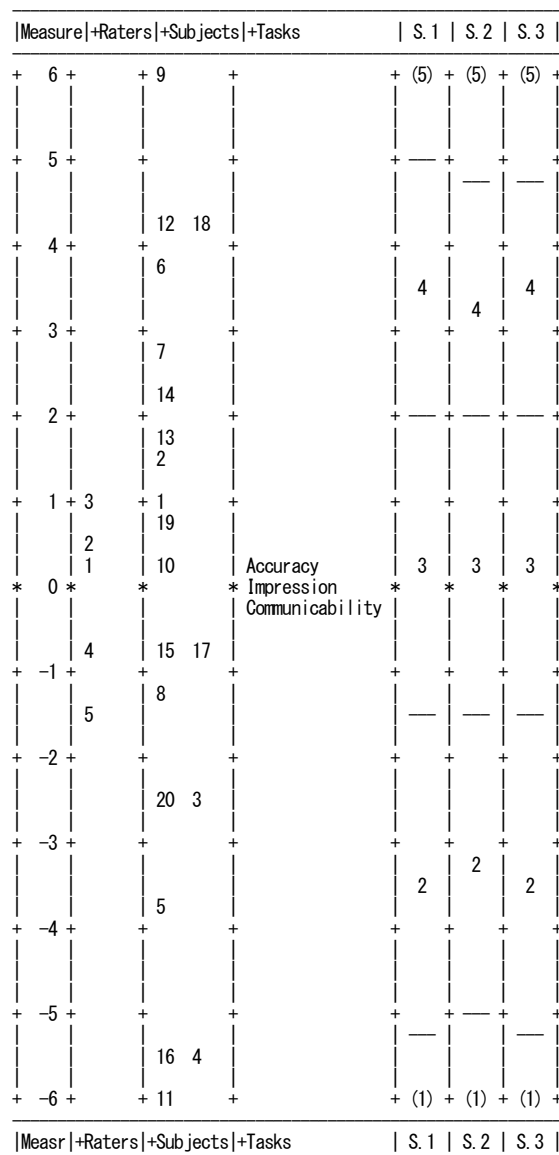


Figure 1 FACETS summary

##### 2) Are teacher-raters equally severe?

The severity span between the most severe rater and the most lenient rater was 2.38 and the difference, based on fair average scores, is 0.65 of one grade in the scale. The reliability of the separation index (which indicates the likelihood to which raters consistently differ from one another in overall severity) was high (.92). The chi-square of 62.8 with 4 df was significant at  $p < .00$  and, therefore, the null hypothesis that all raters were equally severe must be rejected. There was a significant difference in severity among raters. On the other

hand, no raters were identified as misfitting: fit values for all raters were within the range of two standard deviations around the mean ( $0.90 \pm [0.17 \times 2]$ ). In other words, all raters behaved consistently in the scoring.

3) *How much do tasks that are designed to be equivalent actually differ in difficulty?*

The analysis of the two test tasks and impressionistic scoring shows that no significant variation in difficulty exists among them. Raters are considered to be self-consistent in scoring, and the tasks do not appear to separate the subjects to a significant degree meaning that the difficulty of the two tasks and the total impression of the tasks can be considered equivalent. An estimate of the item discrimination was computed according to the "Generalized Partial Credit Model" approach. 1.0 is the expected value, but discriminations in the range 0.5 to 1.5 provide a reasonable fit with the Rasch model (Linacre, 2007, p.132).

4) *How well do scales conform to expectations about their use? Do raters use all parts of them, and use them consistently?*

Linacre (1997) has proposed guidelines for a rating scale: (1) average category measures should advance monotonically with each category, (2) outfit mean-squares should be less than 2.0, and (3) the step difficulty of each scale should advance by at least 1.4 logits and by no more than 5.0 logits.

Table 6 shows the rating scale statistics for accuracy. Since higher category scores are intended to reflect higher measures, the average category measures are expected to rise. All outfit mean-squares are less than 2.0, meaning that each of the five categories has expected randomness in choosing categories. All increases in step difficulty fall within 1.4 and 5.0, which does meet (3).

Table 6: Rating scale statistics for Accuracy

Category Score	Average Measure	Outfit (mean square)	Step Difficulty
1	-5.77	.9	
2	-2.69	1.1	-5.28
3	.26	1.4	-1.54
4	2.87	.8	1.96
5	4.95	1.0	4.86

Table 7: Rating scale statistics for Communicability

Category Score	Average Measure	Outfit (mean square)	Step Difficulty
1	-5.35	1.2	
2	-3.04	1.6	-4.97
3	.26	.9	-1.49
4	2.74	.6	1.98
5	4.97	.7	4.49

Table 7 shows the rating scale statistics for communicability. All outfit mean-squares are less than 2.0, which meet (2). All step difficulty increases fall within 1.4 and 5.0, which does meet (3).

Table 8: Rating scale statistics for Impression

Category Score	Average Measure	Outfit (mean square)	Step Difficulty
1	-5.87	.7	
2	-3.19	.5	-5.13
3	.55	.7	-1.65
4	2.96	.7	2.27
5	5.05	.8	4.51

Table 8 shows the rating scale statistics for Impression. Average measures advance monotonically with each category. All outfit mean-squares are less than 2.0. All step difficulty increases fall within 1.4 and 5.0, which does meet (3).

5) *Do individual raters score a particular group of subjects more harshly or more leniently? If so, what are the sub-patterns of ratings in terms of rater-subject interaction for each rater?*

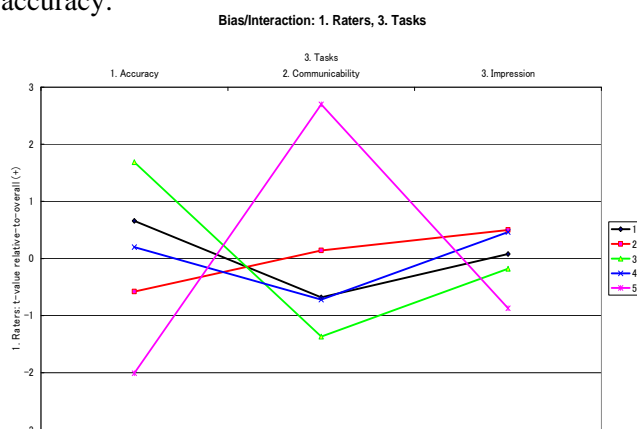
There were a total of five significantly biased interactions among Rater 1, Rater 3 and Rater 5. The three raters' views are summarized, and indicate that each rater had a unique rater-subject bias pattern as follows:

- Rater 1 becomes harsher on accuracy when the number of words are exceeded and the script lacks an organizing principle and development. Rater 1 becomes more lenient when the script demonstrates linguistic accuracy. Rater 1 becomes harsher on communicability when the written items are similar, and its number is limited. Rater 1 becomes more lenient when the script displays adequate communicative effect.
- Rater 5 becomes more lenient on accuracy when there are a number of words and the script demonstrates clear organization with a variety of linking devices. Rater 5 becomes harsher when the script displays a lack of organizational skills. Rater 5 becomes harsher on communicability when the written items are similar, and its number is limited. Rater 5 becomes more lenient when the script displays adequate communicative effect.
- Rater 3 becomes harsher on accuracy when the script displays a lack of organizational skills, but becomes more lenient when the script demonstrates communicative effect.

6) *Do the raters score particular tasks more harshly or more leniently than others? If so, what are the sub-patterns of ratings in terms of rater-task*

### interaction for each rater?

There were two interactions with a significant bias out of the entire 15 interactions. The interactions that displayed a significant bias were distributed in one rater (Rater 5), which indicate that Rater 5 is harsher on communicability and more lenient on accuracy.



### 7) To what extent, statistically, is the task-based writing test a reliable and valid measure?

#### (1) Reliability

All raters displayed acceptable levels of consistency with themselves, but there were significant differences between raters in terms of severity. The difference, based on fair average scores, is 0.65 of one grade in the scale, suggesting that there would be no impact on scores awarded in an operational setting. The analyses of the two tasks and the impressionistic scoring show that no significant difference occurs between the tasks and the impressionistic scoring. The scoring forms do not appear to separate the subjects to a significant degree. This means that in normal operations the three scoring forms can be considered equivalent.

#### (2) Validity

An estimate of the item discrimination was computed according to a "Generalized Partial Credit Model" approach. 1.0 is the expected value, but discriminations in the range 0.5 to 1.5 provide a reasonable fit to the Rasch model (Linacre, 2007, p.132). All the estimates fall in this range (0.90, 1.05, 1.37), which indicates that the randomness in the three sets of data fit the Rasch model. The two tasks and the impressionistic scoring were, therefore, of relevance to dependent data acquisition.

Table 10 Inter-rater correlation coefficients between raters' scores and the *Criterion* score

Rater	1	2	3	4	5	Avg.
T1	.71	.70	.65	.63	.60	.66
T2	.74	.71	.67	.70	.79	.72
IS	.74	.78	.70	.72	.68	.72

Note. T1=task1; T2=task 2; IS=impressionistic scoring

There is also evidence that detracts from the measure's validity. Table 10 shows the resulting correlation coefficients for the relationship between each of three raters' scores and the *Criterion* score, and they were statistically significant ( $p < .01$ ) for Task 1, Task 2 and impressionistic scoring. This result supports the validity of the task-based writing test including these three scores.

## 4. Discussion

The findings suggest that the TBWT scoring guide may have contributed to the reduction of biased interactions, but training for certain raters with his/her unique bias patterns might still be required. It was assumed that the scoring guide gave raters a shared understanding of the construct of writing ability as defined by the test writers, and thus the scoring guide may effectively reduce the differences or biases caused by variation among raters. However, as previous research suggests, training and experience improve agreement among raters (Shohamy *et al.*, 1992; Weigle, 1994). Lumley (2002) suggested that trained teacher raters garner the benefit of training by simply coping with the demanding task, shaping their natural impression to what they are required to do, and using the scale to frame the descriptions of their judgment of a text. This view of the function of training suggests that training plays an important role in influencing raters' behavior, so it may contribute to the variation in frequencies of biased interactions.

From the results of the present study using FACETS, three implications are drawn. First, five teacher raters were found to be self-consistent in scoring 20 different subjects' writing performance. However, there were relatively small but significant differences in overall rater severity. In addition, three of the five raters had a unique bias pattern toward a certain type of text. Fit statistics analysis of the raters in this study suggested that training for a certain rater with his/her unique bias pattern could have a major impact on rating behavior, meaning that the rater facet does not necessarily represent a problematic or validity-threatening part of the testing process.

Second, the 5-point scales were found to demonstrate acceptable fit, and seemed to be a more reliable tool in determining the estimate of subjects' writing ability. The scales associated with the five rating categories and their specific written samples were shown to be mostly comprehensible and usable by raters. However, it must be said that the raters in this study were all participants in the pre-testing. Raters tend to increase their internal consistency in assigning ratings as they gain experience (Weigle, 1998). Whether new teacher

raters are self-consistent in scoring the same writing samples with the rating scales must be observed and confirmed in further studies.

Finally, one source of score variance in the writing performance test, task, was negligible in terms of difficulty. The assessment tasks used in this study provided reasonable fit to the Rasch model. This result implies that task development based on the construct-based processing approach could be a reasonably solid basis to estimate students' writing ability, and those tasks may draw valid inferences to their writing performance.

## 5. Conclusion

In the present study, the results showed that the students' ability was effectively measured using the developed elicitation tasks and five teacher raters, and that all raters displayed acceptable levels of consistency with themselves. There were, however, relatively small but significant differences among raters in terms of severity. The bias analyses also indicated three of the five raters were significantly biased towards certain types of subjects, and these raters' bias patterns were unique. These findings suggest that the TBWT scoring guide may have contributed to the reduction of biased interactions, but training for certain raters with his/her unique bias patterns might still be required.

The FACETS analysis for this study showed that the difficulty of the two tasks and the impressionistic scoring were considered equivalent, which provided reasonable fit to the Rasch model. The equivalence of task difficulty may indicate that task development based on the construct-based processing approach could be reliable and valid to estimate students' writing ability. The rating scales associated with the five categories and their specific written samples were shown to be mostly comprehensible and usable by raters, and demonstrated acceptable fit. However, there is still room for argument about the reliability and validity of assessment tasks and rating scales.

## References

- Alderson, J.C., Clapham, C. and Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Bachman, L.F. and Palmer, A.S. (1996). *Language testing in practice: designing and developing useful language tests*. Oxford: Oxford University Press.
- Bachman, L.F. (2002). Some reflections on task-based language performance assessment. *Language Testing* 19, 453-76.
- Brown, J.D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall Regents.
- Connor, U. and Carrell, P. (1993). The interpretation of tasks by writers and readers in holistically rated direct assessment of writing. In J.G. Carson & I. Leki (Eds.), *Reading in the composition classroom* (pp. 141-160). Boston, MA: Heinle and Heinle.
- Cumming, A. (1989). Writing expertise and second language proficiency. *Language Learning* 39, 81-141.
- Cumming, A. (1990). Expertise in evaluating second language composition. *Language Testing* 7, 31-51.
- Davies, A. (1990). *A Principles of language testing*. Basil Blackwell.
- Hyland, K. (2003). *Second language writing*. Cambridge: Cambridge University Press.
- Linacre, J. (2002). Guidelines for rating scales. Mesa Research Note 2 (Online). Available at <http://www.rasch.org/rn2.htm> (accessed 18 March 2008).
- Linacre, J. (2008). *Facets*, version no. 3.63. Computer program. Chicago, IL: MESA Press.
- Lunz, M., Wright, B. and Linacre, J. (1990). Measuring the impact of judge severity on examination scores. *Applied Measurement in Education* 3, 331-45.
- Pollitt, A. (1990). Response to Charles Alderson's paper: 'Bands and scores.' Alderson, J. C. (1991). In J. C. Alderson and B. North (Eds.), *Language Testing in the 1990s*. (pp. 87-91). London: Macmillan.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.
- Shohamy, E., Gordon, C. and Kraemer, R. (1992). The effect of raters' background and training on the reliability of direct writing tests. *Modern Language Journal* 76, 27-33.
- Sugita, Y. (2008). *Task-based performance assessment of Japanese second language writing*, Paper presented at 42<sup>nd</sup> International Annual IATEFL Conference, Exeter, the UK, April.
- Weigle, S. (1994). Effects of training on raters of ESL compositions. *Language Testing* 11, 197-223.
- Weigle, S. (1998). Using FACETS to model rater training effects. *Language Testing* 15, 263-87.
- Weigle, S. (2002). *Assessing writing*. Cambridge: Cambridge University Press.

# Incorporating Information Gap Activities in EFL Elementary Classrooms

Chiou-hui Chou

National Hsinchu University of Education  
joyce@mail.nhcue.edu.tw

## Abstract

In second language education, researchers indicate that task-based teaching has been recognized as the latest methodological realization of communicative pedagogy. How can this approach be applied in EFL elementary classrooms? What are teachers' perceptions of applying information gap activities in elementary classrooms? This study sets out to explore about the above questions. It reports a project led by the researcher to guide a group of teachers to incorporate information gap activities in their English teaching for their professional knowledge development. It investigates teachers' perceptions and experiences of applying information gap activities in EFL elementary classrooms during their learning to teach in Taiwan. A questionnaire was administered to collect the data. It draws conclusions from the aspects of teaching a second language in terms of linguistic and pedagogical considerations. It finds that teachers' experiences of applying information gap activities in EFL elementary classrooms make them believe that information gap activities can help students reinforce vocabulary and grammatical structures, allow students to use linguistic forms and functions in a communicative way, and increases more student engagement. It offers suggestions for TESOL training programs to adopt a similar training approach to help teachers to gain practical experiences of applying a theory into practice.

## Keywords

information gap activities, EFL, task-based instruction

## 1. Introduction

As Littlewood (2007) points out "educators and governments in Asian countries are intensively addressing the need to increase the number of people in their population who can communicate effectively in English" (p. 243). In order to achieve this goal, national policies and syllabuses have been moving towards

communicative language teaching (Littlewood, 2007; Nunan, 2003). More recently, as Nunan (2003) indicates, teachers are urged to move on from earlier forms of CLT and introduce task-based language teaching in Asia (e.g., Hong Kong, Korea, and China). Task-based teaching has been recognized as the latest methodological realization of communicative pedagogy" (Hu, 2005).

During the past decades, in keeping up with the trend of initiating English programs in elementary education, most Asian countries have implemented one-to-two English instruction periods into their curriculum (e.g., Taiwan, Japan, and Korea). For example, in Taiwan the Grade 1-9 Curriculum was implemented in 2001. Following this new curriculum, English courses, embedded in language arts, were officially implemented for fifth and sixth graders with two periods of instruction per week. Later, in 2003, one 40-minute period of English course per week was officially implemented for third and fourth graders. According to the curriculum guidelines, the goal of English education generally focuses on developing students' basic communicating competences, cultivating their English learning interests and habits, and introducing international culture and social customs. This is just an example of EFL context in Taiwan. Other Asian countries might have different policies.

However, rarely have researchers discussed how communicative approach is applied in elementary education. Literature on communicative approach has been discussed prevalently in high schools and universities. Little is known about how teachers can use communicative activities in EFL elementary classrooms. In order to gain insight into this topic, this study investigates teachers' experiences and perceptions of applying communicative activities in their teaching. It describes a project led by the researcher in an inservice teacher training course, which guided a group of teachers to incorporate information gap

activities in their English teaching for their professional knowledge development.

## 2. Literature review

Why are information gap activities chosen? An information gap activity involves the exchange of information among participants in order to complete a task (Larsen-Freeman, 2000). As Pica (2005) indicates, the information gap task is “characterized by several features: Only one outcome or answer is considered possible, appropriate, or correct, and reaching it requires a verbal exchange of information among task participants” (p. 341). This type of activity came from task-based teaching (Ellis, 2003; Nunan, 2004), which Richards (2005) explains as an extension of the CLT movement and which Littlewoods (2004, 2007) regards as a development within the communicative approach, to develop learners’ communicative competence. It focuses on meaningful interaction while drawing students’ attention to language forms (Long & Crookes, 1993). As Willis and Willis (2007) point out, a focus on form means that “teachers isolate one or two specific forms, specific grammatical structures or functional realizations, and identify these as the target forms” (p. 4). Also, learners know that “by the end of the teaching sequence, often contained in a single lesson, they will be expected to produce these forms with an acceptable level of accuracy” (p. 4). Thus, task-based teaching “aims primarily at helping learners get ready to use the language in real-life situations” (p. 131). In addition, in form-focused work, learners “work on recognizing or manipulating the forms of the language in a number of ways: conscious raising, recall, extension, correction, and exam practice” (p. 133).

Pica (2005) in her review article discusses the contributions of information gap tasks as seen from our learning, teaching, and research perspectives, describes the issues and challenges we have faced in integrating and implementing them, and presents an approach that we have developed for designing information gap tasks both as authentic activities for teaching and learning and as reliable instruments for research. Teachers and researchers have shown that information gap activities have the following advantages: (a) They can reinforce vocabulary and a variety of grammatical structures taught in class; (b) They allow learners to use linguistic forms and functions in a communicative way, (c)

These activities bring the language to life for learners, (d) Learners have the opportunity to use the building blocks of language we teach them to speak in the target language, and (e) Learners have to be involved in their learning, and not as passive instruments who seem to get knowledge from others, but as active individuals who practice what they learn.

Based on the above advantages, in order to actively engage students in speaking activities based on a communicative approach, teachers can design information gap activities for students to practice speaking English. Following Pica’s (2005) perspectives on learning, teaching, and research and other researchers’, this study hopes to add examples to the literature of task-based communicative language teaching.

## 3. The study

The main purpose of this study is to explore inservice teachers’ perceptions of applying information gap activities in elementary classrooms. These teachers volunteered to join an English teacher certificate program offered by the researcher’s university, funded by the bureau of education in their county. In order to be a qualified English teacher, the Ministry of Education in Taiwan requires the bureau of education in every county or city to hire teachers who possess both an elementary teacher’s certificate and an English teaching qualification. This training program was offered to help inservice teachers to be qualified as English teachers in this county. In this program, there were ten TESOL training courses, each of which lasted for 36 hours. The focus group consisted of twenty-one inservice teachers, eighteen females and two males, all coming from the same county as the researcher’s.

The present study used the data from one of the courses taught by the researcher, which was about designing English teaching and learning activities and materials. In this course, participants were required to complete two main assignments, which were to plan two lessons by incorporating information gap activities and designing the worksheets, and then to teach the lessons to a chosen class in their schools. They chose the lessons from the textbooks their students were using. During their teaching, they videotaped their lessons and took snapshots. Then, they shared their teaching video clips, snapshots, and reflections with the peers during their mid and final presentations. After each presentation, they gained insights and comments

from the researcher and peers. Last, they wrote a paper to reflect and report what they had learned from the experiences.

The main portion of the data in this study came from a post-course questionnaire (See Appendix A). The questionnaire was designed by the researcher. It consists of 14 items measuring teachers' perceptions of applying information gap activities, using a 5-point Likert scale (1: strongly disagree, 2: disagree, 3: neutral, 4: agree, and 5: strongly agree). It was administered to the twenty present participants at the end of the course; 18 respondents returned the questionnaires. A response rate of 90% was achieved.

## 4. Results

### 4.1 An example

Figure 1 is an example of an information gap activity designed by one of the participants. It presents a basic format of worksheet A and B, each of which has parts of information missing.

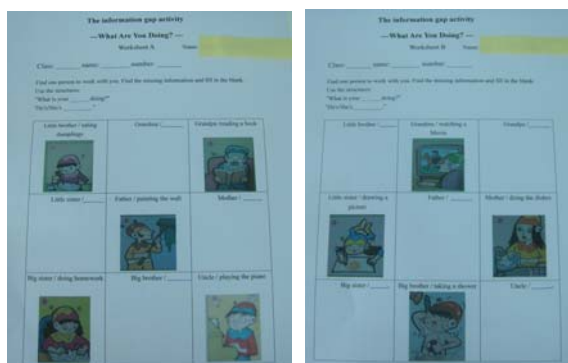


Figure 1: An Example of the Worksheets

The sentence structures and pictures were adopted from students' textbook. The sentence structures are: What is your \_\_\_\_\_ doing? He's \_\_\_\_\_. She's \_\_\_\_\_. Learners have to work with a partner and to use the sentence structures to get the information to fill in the blanks.

### 4.2 The data

Table 1 shows teachers' years of teaching in elementary schools.

Table 1: Teachers' Years of Teaching

Years of Teaching	Number
0-5 years	3
6-10 years	10
11 plus years	5
Total	18

Three participants are novice teachers, teaching less than five years. Ten participants have been teaching for more than 5 and less than 10 years. Five participants have been teaching more than 10 years. The average of the participants' teaching year is 10.

Items 1 to 6 ask about participants' perceptions of how information gap activities can help students' English learning in the aspects of vocabulary, sentences, speaking, meaningful communication, and structures. The results are shown in Table 2. The means of the six items are above 4, indicating that participants agree that applying information gap activities can help students learn English better.

Table 2: How Information Gap Activities Can Help Students' English Learning

Items	M
1. ...learn vocabulary better	4.1
2. ...practice saying sentences better	4.4
3. ...practice speaking more English	4.5
4. ...have the purpose of practice English	4.2
5. ...have meaningful communication	4.2
6. ...learn English structures better	4.2

Note: The elipsis means the information has been omitted.

Table 3 presents the results of Item 7 to Item 14. They show teachers' feelings and experiences of teaching with information gap activities, the difficulties of using pairwork, their opinions about applying this type of activity to teach lower graders, the time they spent in designing the activities, and whether they would value this type of activity in their future teaching. All the results show teachers have positive experiences in applying this type of activities in their teaching. However, they needed to spend a lot of time and efforts designing this type of activity.

Table 3: Teachers' Perceptions of Teaching with Information Gap Activities

Items	M
7. Ss did not like my teaching with IGA	1.8
8. IGA can't attract ss' ... to learn Eng	2
9. ...difficulties grouping ss... 1st IGA	2.4
10. ...difficulties grouping ss... 2nd IGA	2.1
11. ...not suitable for lower graders	3.2
12. I won't use IGA in my future teaching	2
13. ... spend a lot of time designing IGA	3.8
14. I will design IGA in my future teaching	4.4

Item 7 asks, "My students did not like my teaching with information gap activities." The

mean of this item is 1.8, which indicates that teachers thought students liked their teaching with the activities. Item 8 asked, "Information gap activities can't attract students' interests to learn English." The mean of this item is 2, indicating that teacher disagreed with this statement. Examining Items 7 and 8, the study finds that teachers' perceptions of students' feedbacks about the teaching of information gap activities are positive.

Item 9 asks, "I had difficulties grouping students to do the first information gap activity." Similarly, Item 10 asks about the difficulties in the second activity. The means of the two items are 2.4 and 2.1, respectively, indicating teachers did not have difficulties in grouping students.

Item 11 asked, "I think information gap activities are not suitable for lower graders." The mean is 3.2, indicating teachers did not agree with this statement. Item 12 asked, "I won't use information gap activities in my future teaching." The mean is 2, showing they did not agree with this statement. Item 14 asks a similar question, "I will design information gap activities in my future teaching." The mean is 4.4, showing that teachers are likely to design the activities in their teaching.

Item 13 asks, "I need to spend a lot of time designing information gap activities." The mean is 3.8, showing teachers slightly agree with this statement.

## 5. Discussion

Firstly, information-gap activities include a tremendous variety of techniques in which learners are either to convey or to request information. They give learners opportunities to practice using both the language forms of asking and answering. Linguistically, they draw learners' attentions to "focus on forms." They reinforce vocabulary and grammatical structures taught in class, and they allow learners to use linguistic forms and functions in a communicative way (Pica, 2005). As EFL elementary learners have limited exposure to interactive language learning contexts and instruction hours per week, this study shows that teachers find information gap activities not only can help elementary students have the necessity of communicative interaction but also attract their attention to linguistic forms as well (Willis & Willis, 2007). Thus, teachers think that this type of activity can help learners be equipped with linguistic forms better.

Secondly, information gap activities

require pairwork. Did the teachers have difficulties grouping the students? The results of this study show teachers are capable of dealing with this type of situation. One possible reason might be that most of the teachers are veterans and they have more sophisticated skills in classroom management. In this study, teachers did not show they had difficulty using pairwork. As this study engages inservice teachers in the research, other studies that engage novice teachers or pre-service teachers might yield different results.

Thirdly, modern second language teaching approaches emphasize learner-centered classrooms. Using pairwork or groupwork activities to engage learners in learning are common. Task-based instruction also emphasizes learner-centered instruction. As Paul (2003) talks about learner-centered classrooms, he notices that many Asian EFL learners do not have the language ability to work on complex projects in groups, but he suggests that "at least they are able to do simpler activities where they are cooperating together" (p. 30). He thus recommends some possible activities that teachers can encourage children to cooperate with each other. One of the activities is the puzzle project, which I think has the same meaning to promote cooperative learning in doing information gap activities. In an information gap activity, in order to complete each assigned speaking task, learners have to collaborate with the partner to solve the problem. Information gap activities require learners to be active learners, not as passive knowledge receivers. In order to help the young learners to be active participants in speaking activities, teachers have to pay more attention to provide language supports for students. The information gap activity can provide sufficient language supports for learners to practice speaking before they can speak freely. It thus increases more student engagement. In addition, in pairwork, the learner may engage with "more capable others (teachers, advanced peers, etc.), who may provide assistance and guidance" (Oxford, 1997). Thus, this type of activity creates more opportunities for EFL learners to have more real experiences for using English for communication. It may contribute to increasing the number of people who can communicate effectively in English (Littlewood, 2007; Nunan, 2003).

Lastly, for second language teachers and researchers, this study presents an example of

applying and adapting task-based approach in an EFL elementary context (see Littlewood, 2007). It demonstrates how teachers can apply a language teaching approach into a real classroom practically. It advocates that only if teachers are well-equipped with the knowledge of modern language teaching approaches, will they be able to confidently move forward in their teaching career successfully. I thus argue that a solid and fundamental TESOL training program is imperative to promote teacher efficacy. Furthermore, providing an opportunity for teachers to conduct a lesson and demonstrate their teaching in a training course like the ones proposed in this study can help trainees to develop their professional knowledge and learn how to apply theories into real teaching practice.

## 6. Conclusion

This study investigated teachers' perceptions and experiences of applying information gap activities in EFL elementary classrooms during their learning to teach in Taiwan. It has demonstrated an example of putting theory into practice based on task-based instruction in favor of communicative approach. In sum, it draws the conclusions from the aspects of teaching a language in terms of linguistic and pedagogical considerations. It finds that teachers' experiences of applying information gap activities in EFL elementary classrooms make them believe that this type of task can help learners achieve better language accuracy and fluency, creates more real communication opportunities for learners, and increases more student engagement.

In conclusion, second language teachers are continually exploring pedagogical means for educating more successful language learners. This study finds by maximizing the classroom opportunities for learners to practice speaking through communicative activities such as information gap activities, we can get our learners to develop not only linguistic fluency, but also accuracy. Moreover, inviting teachers to apply a teaching method in the training course helps them know how they can apply the language teaching approach to their real teaching context, which also promotes their professional knowledge development.

Thus, this study suggests that TESOL training programs can follow this line of research and training method to help teachers gain authentic experiences of applying theories into practice. In the training, TESOL trainers and researchers can provide opportunities for the

trainees to explore about the language teaching principles, to explore about the process of engaging learners, and to reflect and share their learning results. By setting up a sound training program, we can help more teachers to go smoothly in their teaching career, which will then benefit more successful EFL learners.

## 7. References and appendices

### 7.1 References

- Brown, H. D. (2001). *Teaching by principles: An interactive approach to language pedagogy* (2<sup>nd</sup> ed.). Longman.
- Ellis, R. (2003). *Task-based learning and teaching*. Oxford: Oxford University Press.
- Hu, G. W. (2005). English language education in China: Policies, progress, and problems. *Language Policy*, 4(1), 5-24.
- Larsen-Freeman, D. (2000). *Techniques and principles in language teaching*. Oxford: Oxford University Press.
- Lee, J. F., & VanPatten, B. (1995). *Making communicative language teaching happen*. McGraw-Hill, Inc.
- Littlewood, W. (2004). The task-based approach: Some questions and suggestions. *ELT Journal*, 58(4), 319-326.
- Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. *Language Teaching*, 40, 243-249.
- Nunan, D. (2003). The impact of English as a global language on educational policies and practices in the Asia-Pacific region. *TESOL Quarterly*, 37(4), 589-613.
- Nunan, D. (2004). *Task-based teaching*. Cambridge: Cambridge University Press.
- Oxford, R. (1997). Cooperative learning, collaborative learning, and interaction: Three communicative strands in the language classroom. *Modern Language Journal* 81(4), 443-56.
- Richards, C. (2005). *Communicative language teaching today*. Singapore: RELC.
- Paul, D. (2003). *Teaching English to children in Asia*. Hong Kong: Longman.
- Pica, T. (2005). Classroom learning, teaching, and research: A task-based perspective. *The Modern Language Journal*, 89(3), 339-52.
- Long, M., & Crookes, G. (1993). Units of analysis in syllabus design: The case

for task. In G. Crookes and S. Gass (eds.), *Tasks in a pedagogical context*. Clevedon: Multilingual Matters.

Willis, D., & Willis, J. (2007). *Doing task-based teaching*. Oxford: Oxford University Press.

## 7.2 Appendices

### Appendix A. The Questionnaire

Items	Mean
1. Information gap activities can help students learn vocabulary better.	4.1
2. Information gap activities can help students practice saying sentences better.	4.4
3. Information gap activities can help students practice speaking more English.	4.5
4. Information gap activities can help students have the purpose of practice speaking English.	4.2
5. Information gap activities can help students have meaningful communication with peers.	4.2
6. Information gap activities can help students learn English structures better.	4.2
7. My students did not like my teaching with information gap activities.	1.8
8. Information gap activities can't attract students' interests to learn English.	2
9. I had difficulties grouping students to do the first information gap activity.	2.4
10. I had difficulties grouping students to do the second information gap activity.	2.1
11. I think information gap activities are not suitable for lower graders.	3.2
12. I won't use information gap activities in my future teaching.	2
13. I need to spend a lot of time designing information gap activities.	3.8
14. Once in a while, I will design information gap activities in my future teaching.	4.4

# Greetings and EFL Teachers

James Martin Ronald<sup>1</sup> and Timothy Floyd Hawthorne<sup>2</sup>

<sup>1</sup>Faculty of Humanities, Hiroshima Shudo University

<sup>2</sup>International Exchange Center, Hiroshima International University

ronald@shudo-u.ac.jp, h-timoth@it.hirokoku-u.ac.jp

## Abstract

Greetings are a typical part of many conversations, yet for less confident language learners they may represent a particular challenge, either resulting in a feeling of success or of failure. More proficient learners, however, may feel limited by basic greetings and want to learn and use a wider range of greetings. There were three main reasons for focusing on greetings: they preface much person-to-person interaction; success or failure in their use may directly impact learner motivation; and, finally, greetings may be effectively taught.

This research is conducted with students in two provincial Japanese universities, for most of whom there is little or no contact with English speakers outside the classroom. We decided to focus on two types of learner: false beginners with little ability or confidence in using English and more able and confident students majoring in English. Our study involved the administration of a questionnaire in which students were asked to rate preferences regarding greetings to be used outside class, or to be learned or practiced in class. Following in-class focus on the learning and practice of greetings over a two-three week period, the same questionnaire was administered a second time.

## Keywords

Pragmatic competence, failure, greetings

## 1 Introduction

Especially at beginner or false-beginner level, communicating in a foreign language outside the classroom can be a high risk venture. Attempts at communication, or even simple greetings, may easily result in misunderstandings and confusion, and cause the learner to be discouraged. On the other hand, successful conversations in the target language, however short, may be exhilarating and motivating experiences.

The EFL teacher has two main roles with regard to out-of-class communication: in class, to prepare students to cope with these language needs

and, out of class, either to greet and have short conversations with the students in English or to facilitate such conversational opportunities. This second role is of particular importance in environments in which there are few other English speakers or opportunities to use the target language. In order for teachers to fulfill these roles satisfactorily, we need to know what our students already know, and to know what they need or want to know. We focused on greetings, partly for their crucial role in short conversations, and partly because we felt that awareness raising and teaching in this area were most likely to produce tangible results.

This research is conducted with university students in two provincial universities in Japan, for most of whom there is little or no contact with English speakers other than their teachers outside the classroom. In order to gain some understanding of the range of knowledge and preferences of students at university level, we decided to focus on two types or levels of English learner: false beginners with little ability or confidence in using English; and students majoring in English who are more able and confident about using the language.

Our study involved the administration of a questionnaire in which students were asked to rate preferences regarding greetings to be used outside class, or to be learned or practiced in class. Following in-class focus on the learning and practice of greetings over a two-three week period, the same questionnaire was administered a second time.

## 1.1 Outline

Following the Introduction, in the Literature Review we will consider issues relating to greetings, such as cultural issues, pragmatic failure, adjacency pairs, and the teachability of pragmatic competence in this area. We will then report our research, together with research questions and results, before going on to report our findings. Finally, based on this preliminary study, we will offer suggestions for English teachers' use and teaching of greetings both

in and out of the language classroom.

## 2 Literature Review

Greetings, as many dictionaries implicitly testify, typically do not mean something. “Hello”, for example, does not mean something; rather, it is used in a particular context to acknowledge someone's presence, with intonation and body language often conveying the speaker's attitude regarding that person's presence.

While the speech act of greeting is as familiar to Japanese speakers as it is to English speakers, the range of greetings and responses available in each language, together with the pressure to perform, may render this seemingly simple act complex and confusing.

The choices available in greeting and response types leads us to three related areas of research: language and culture; adjacency pairs, and pragmatic competence or failure. Finally, we will address the question of whether pragmatic competence, specifically relating to greetings, is something that can be taught.

### 2.1 Language and culture

Our starting point in focusing on similarities and differences is similar to the goal stated by Brown and Levinson (1987: 56), “to show that superficial diversities can emerge from underlying universal principles and are satisfactorily accounted for only in relation to them”. Indeed, it is often the failure to identify the universal in differently expressed greetings that causes confusion. Yamada (1997: 49) gives the following example: “When Americans ask each other “What’s happening?” they don’t really mean for you to give them a complete summary of what has happened to you since you last saw them. Unless you are good friends and have a chance to sit down and chat, a greeting like, “What’s up?” is treated as a conventional expression like “How are you?”

As Stephens (2003) points out, learners of Japanese are just as likely to be misled by the surface meaning of greetings in that language: “Like the English greeting “How are you?” the Japanese equivalents *odekake desu ka* (“Are you going out?”), *kaimono desu ka* (“Are you going shopping?”) or *dochira made desu ka* (“Where are you going?”), are not requests for information and do not require a direct answer”. This situation, leading to pragmatic failure, is described by Thomas (1983) as “the inability to understand what is meant by what is said”. Such pragmatic failure with greetings may have dire consequences for the whole conversation.

### 2.2 Adjacency pairs

An issue related to the surface reading of greetings is that of the extent of freedom of choice of response to greetings. Schegloff and Sacks (1973) describe how adjacency pairs, such as greeting and response, are not fixed word-for-word exchanges, but that the response must be chosen from one of a few types of second pair parts. Both Yamada (1997: 49) and Stephens (2003) point out the relative fixity of greeting response types in Japanese as compared with English, and suggest the consequences that it may have for successful communication. Yamada explains this as follows: “The Japanese prefer favor conventional greetings to responses, while Americans prefer ones that are more individualized. This does not mean that American responses are altogether original, but that a variation from “Fine thank you, and you?” is considered more interesting and more hip – in short, preferable” (1997: 49).

### 2.3 Pragmatic failure

Japanese learners’ selection of such dispreferred responses is found not only in direct cultural preferences but also in classroom learning of greetings and responses which may leave many with the perception that “Fine thank you, and you?” is the *only* acceptable response to “How are you?” Stephens, again, addresses the issue of teacher responsibility to help learners overcome these problems: “Pragmatic failure must be addressed by the teacher because of the danger of creating misunderstanding. If the speaker's intent is not communicated accurately [...] misunderstanding may ensue” (2003).

### 2.4 Teachability of greetings

This issue of teacher responsibility, finally, brings us to the question of whether pragmatic competence, specifically that relating to greetings, will either be transferred from the learner’s L1 or acquired through observation of, and interaction with, speakers of the target language. In other words, that it does not require attention in the language classroom. Kasper (1997) has addressed this question in detail. As she points out, “some pragmatic knowledge is universal, and other aspects may be successfully transferred from the learners’ L1”. She also notes that learners often fail to take advantage of this and transfer L1 pragmatic knowledge for application in the L2. This may be especially true of peoples, and languages, which are commonly seen as being very different from each other, as Japanese and Americans or the Japanese and English languages are. The tendency to literal interpretation of greetings rather than a recognition of shared pragmatic value, is a case in point.

As for opportunities for learning through contact with target language speakers, this is only a

realistic option if such contact exists outside the passive media of television and cinema. Where it does not, the responsibility for addressing areas of likely pragmatic failure remains with the teacher. This brings us to the question of whether pragmatic routines such as greetings can be effectively taught. Kasper (1997), following a review of research into teaching areas of pragmatic competence, concludes that pragmatic routines can indeed be successfully taught to foreign language learners with lower levels of proficiency. This conclusion leads us to our study, in which through awareness-raising and direct teaching or practice we aimed to do precisely that.

### 3. Research and Results

Our study was focused on Japanese language learners' knowledge of, and attitudes towards, greetings used in English. We addressed three questions through our research:

1. With which English greetings would our learners prefer to be greeted outside class? What variation is there for different proficiency level classes?
2. Which greetings do they want to learn or practice? What variation is there for different proficiency level classes?
3. What change in responses are there following in-class focus on greetings?

#### 3.1 Instruments, participants, and teaching

The research instrument, a basic questionnaire (see Appendix A), required the students to rate a set of 24 English greetings according to whether they would wish to be greeted with these outside class, and to whether they wished to learn or practice the greetings in class. Specifically, for each greeting, participants were first asked to indicate the frequency with which they would want the greeting to be used to address them outside class, choosing between *Sometimes*, *Often*, and *Never*. Then, for the same set of greetings, participants were asked for each greeting whether they wanted to learn or practice the greeting in class: *Learn*, *Practice*, or *Not do*. To avoid the first question having an undue influence on answers for the second, questionnaires were folded in half so that only one question or the other would be visible at one time. The survey was administered one more time, around four weeks later.

A total of 86 participants, mostly aged 18 to 21, took part in the study: these included classes of Nursing students, of Pharmaceutical Sciences students, and of English Department students. The reason for including groups of students with different majors was not so much concerned with their particular field of study as with the varying

levels of ability, motivation, and interest in English typical of students in each field.

For these participants the study comprised the two administrations of the survey described above, and focused teaching informed by the first survey results in the three weeks between the surveys. Specifically, the teaching involved explicit explanation and practice of response choices both for familiar greetings such as "How are you?" or "How are you doing?" and for greetings less commonly encountered in the language class, such as "What's up?" or "Holding down the fort?"

### 3.2 Results

Sets of data drawn from the two surveys are presented below in Tables 1 and 2. Table 1 shows general participant preferences for out-of-class use of greetings for each group of participants according to their major at university as expressed both prior to the classroom focus on greetings (Pre) and at the end of the study (Post), while Table 2 provides data for preferences for in-class teaching and practice of the greetings.

#### 3.2.1 Out-of-class preferences

As Table 1 shows, there was considerable difference for overall preferences among the groups of participants, with English Department students selecting *Often* or *Sometimes* for a total of over 90% of greetings as compared with around 50% for Nursing students and 57% for Pharmaceutical Science students. Regarding differences between the first and second administration of the questionnaire, there was very little difference overall.

Table 1: Greeting preferences for out of class

			Pharmacy	Nursing	English
Pre	Often	No.	135	245	187
		%	19.4	27.6	43.3
Post	Often	No.	256	331	207
		%	38.1	36.3	47.9
Pre	Some-times	No.	479	516	212
		%	68.8	58.2	49.1
Post	Some-times	No.	384	488	186
		%	57.1	53.5	43.1
Pre	Never	No.	82	126	33
		%	11.8	14.2	7.6
Post	Never	No.	32	93	39
		%	4.8	10.2	9.0

Note: Asterisk for figures not available when submitted

In addition to the figures above, it is worth looking at the most and least popular greetings for teachers to use with students outside class. The top five greeting for each participant group was as follows: Pharmaceutical Sciences: *Hi!*, *Hello!*, *How are*

*you?*, *Nice day, isn't it?*, and *How are you doing?*  
Nursing: *Hi!*, *Hello!*, *How are you?*, *Nice day, isn't it?*, and *How are you doing?*.

English: *Hi!*, *Hello!*, *How are you?*, *How are you doing?*, and *Nice day, isn't it?*.

The least popular greetings were:

Pharmaceutical Sciences: *How's the family?*, *Holding down the fort?*, *Keeping out of trouble?*, *Keeping up with your studies?*, and *Getting much sleep lately?*.

Nursing:

English: *Keeping out of trouble?*, *How's life been treating you?*, *How's the family?*, *Getting much sleep lately?*, and *Holding down the fort?*.

Table 2: Preferences for in-class learning and practice.

			Pharmacy	Nursing	English
Pre	Learn	No.	313	392	*
		%	45.0	46.7	
Post	Learn	No.	306	245	141
		%	51.0	30.0	32.6
Pre	Practice	No.	335	370	*
		%	48.1	44.0	
Post	Practice	No.	257	500	273
		%	42.8	61.3	63.2
Pre	Not do	No.	48	78	*
		%	6.9	9.3	
Post	Not do	No.	37	71	18
		%	6.2	8.7	4.2

Note: 1. Asterisk denotes figures not available at time of submission

The top three greetings for learning were as follows:

Pharmaceutical Sciences: *Holding down the fort?*, *What's new with you?*, *How are you holding up?*

Nursing: *How's life been treating you?*, *Holding down the fort?*, and *How are you holding up?*.

English: *How's life been treating you?*, *Holding down the fort?*, and *How are you holding up?*.

For practicing, these were the most chosen greetings:

Pharmaceutical Sciences: *Hi!*, *How are you?*, and *Hello!*.

Nursing: *Hello!*, *Nice day, isn't it?*, and *How are you doing?*.

English: *Hello!*, *What's new?*, and *How are you?*.

We will now go on to consider these results in detail.

#### 4. Discussion

We will begin by returning to address the three questions raised earlier. As we do this, we will consider why we obtained the results we did, and include any further data that may be relevant to these questions.

1. With which English greetings would our learners prefer to be greeted outside class? What variation is there for different proficiency level classes?

As we saw in the Results section above, the largest number of participants selected the most familiar and basic greetings as those with which they would prefer to be addressed on campus: *Hi!*, *Hello!*, *How are you?*, *Nice day, isn't it?*, and *How are you doing?* The same few greetings were, surprisingly, most popular among all three groups of participants, including the more confident and capable English Department students. However, a closer look at the figures shows that the high response rates of *Often* and *Sometimes* for the majority of greetings for these participants resulted in the near universally *Often* rated being counted as most popular; in other words, less confident participants within this group had the "voting power" to relegate other greetings to the not quite so popular level. Another valid interpretation of the data may be to consider that participants would consider the everyday greetings they selected as suitable to be used, literally, every day, while less familiar greetings were less familiar precisely because they are used less frequently. This impression may not be accurate, and based on classroom or textbook-mediated exposure to English, since, at least for native speakers of their age, greetings such as "What's new?" or "What's up?" may be used at least as frequently as the more prosaic "How are you?".

The gap between *Often* and *Sometimes* was often greater in Pharmaceutical Science and Nursing groups, suggesting that while they felt ready for the most familiar of greetings, they generally did not wish to be greeted with the other greetings about which they were not confident. Finally, regarding the Pharmaceutical Science students, one interesting response was the popularity of the less common greeting "How's your health?". While in the world in general, this might be seen as a less common greeting, or a more specific one, in the world of health care it is certainly an everyday greeting; it is precisely this world for which this group of participants is being prepared.

2. Which greetings do they want to learn and practice? What variation is there for different proficiency level classes?

As is clear from their answers, the majority of participants differentiated between learning and practicing. As a rule, unfamiliar or unusual greetings were identified as ones that most participants wanted to learn:

*How's life been treating you?, Holding down the fort?, and How are you holding up?.*

This raises a couple of issues. It may be perfectly reasonable that language learners are drawn to language that they do not know yet. At the same time, it is important for learners to have a sense of the words and phrases they need to learn out of all the whole target language.

3. What change in responses are there following in-class focus on greetings?

There was surprisingly little change in this area, most probably due to the short length of the learning period of this study. Generally, though, responses suggested increased interest and confidence in using greetings that were previously less well known.

## 5. Conclusion

This study has been of value largely in bringing to our attention the importance of greetings, and in identifying greetings that learners want to know. Teachers need to be sensitive to these questions so as to do the best for their learners both in and out of class.

## References

- Brown, P., & Levinson, S.C. *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.
- Kasper, G. (1997). *Can pragmatic competence be taught?* Hawaii: Second Language Teaching and Curriculum Center.
- Schegloff, E.A., & Sacks, H. (1973). Opening up closings. *Semiotica* 8: 289-327.
- Stephens, M. (2003). Sources of Pragmatic Failure by Japanese EFL Learners: A Discussion of Greetings and Leave-Taking. *Studies in the Humanities and Sciences, Hiroshima Shudo University*, 44, 1, 373-390.
- Thomas, J. (1983). Cross-cultural pragmatic failure. *Applied Linguistics*, 4 (2), 91-112.
- Yamada, H. (1997). *Different games, different rules: Why Americans and Japanese misunderstand each other*. New York: Oxford University Press.

## Appendix A. Title of Appendix A.

学生のみなさんへ

次のアンケートへの協力をお願いします。英語のあいさつに関する 2 種類の質問に答えてください。これはテストではなく、成績に関係するものではありませんが、授業の内外で教師の参考にさせていただきます。協力ありがとうございます。紙を半分に折って始めてください。

Tim Hawthorne

**教室外**(例えばキャンパス内、食堂など)で、あなたは次のどのあいさつを外国人の英語の先生からしてもらいたいと思いますか？

下の欄に O, S, X から選んで書き入れてください。

O=Often

(よくこのあいさつをしてもらいたい)

S=Sometimes

(時々このあいさつをしてもらいたい)

X =Never

(ぜったいこのあいさつをしてほしくない)

\_\_\_\_\_ What's up?

\_\_\_\_\_ Hi!

\_\_\_\_\_ How's it going?

\_\_\_\_\_ What's new?

\_\_\_\_\_ How's life been treating you?

\_\_\_\_\_ How are you doing?

\_\_\_\_\_ How are things going?

\_\_\_\_\_ How have you been?

\_\_\_\_\_ How's the family?

\_\_\_\_\_ Nice day, isn't it?

\_\_\_\_\_ How are you getting along?

\_\_\_\_\_ Staying busy?

\_\_\_\_\_ Keeping out of trouble?

\_\_\_\_\_ Getting much sleep lately?

\_\_\_\_\_ Hello!

\_\_\_\_\_ Holding down the fort?

\_\_\_\_\_ How are things with you?

\_\_\_\_\_ How are you?

\_\_\_\_\_ Keeping up with your studies?

\_\_\_\_\_ How's your health?

\_\_\_\_\_ How are you holding up?

\_\_\_\_\_ How are you getting on?

\_\_\_\_\_ How are your classes?

\_\_\_\_\_ What's new with you?

**教室内で**、あなたは次のどのあいさつを教えてもらいたいですか？または練習したいですか？

下の欄に L, P, X から選んで書き入れてください。

L=Learn

(授業でこのあいさつを習いたい。まだ知らない)

P=Practice

(授業でこのあいさつを使う練習をしたい。もう知っている)

X=Not do

(授業でこのあいさつを習いたくない、練習したくない)

\_\_\_\_\_ What's up?

\_\_\_\_\_ Hi!

\_\_\_\_\_ How's it going?

\_\_\_\_\_ What's new?

\_\_\_\_\_ How's life been treating you?

\_\_\_\_\_ How are you doing?

\_\_\_\_\_ How are things going?

\_\_\_\_\_ How have you been?

\_\_\_\_\_ How's the family?

\_\_\_\_\_ Nice day, isn't it?

\_\_\_\_\_ How are you getting along?

\_\_\_\_\_ Staying busy?

\_\_\_\_\_ Keeping out of trouble?

\_\_\_\_\_ Getting much sleep lately?

\_\_\_\_\_ Hello!

\_\_\_\_\_ Holding down the fort?

\_\_\_\_\_ How are things with you?

\_\_\_\_\_ How are you?

\_\_\_\_\_ Keeping up with your studies?

\_\_\_\_\_ How's your health?

\_\_\_\_\_ How are you holding up?

\_\_\_\_\_ How are you getting on?

\_\_\_\_\_ How are your classes?

\_\_\_\_\_ What's new with you?

# Interaction of Native and Target Phonology in Learner Speech: A Study of Korean English

Yunju Heo, Jin Hyun Kim and Jong-mi Kim

Department of English Language and Literature, Kangwon National University  
excelsior8@hanmail.net, tgjykh@hanmail.net, kimjm@kangwon.ac.kr

## Abstract

In the development of learner speech, we expect both the decreasing interference of native phonology and the increasing acquisition of target phonology. We examine how these two types of phonological interaction occur as the learner proficiency increases.

For this study, we collected 2672 word data of L1 English, L2 English, and L1 English loanwords, which were elicited from 124 Korean adult learners of American English and 23 native speakers of Korean and English. Learners were tested for the listening proficiency in terms of a standardized test TOEIC as a covariance unit, and scaled into five ranks. We measured the acoustic features that are relevant to the phonological phenomena in Korean and English. We focused on *l*-alternation in Korean and deaspiration rule in English.

Our results show that the learner speech resists the native phonology interference in significantly earlier stage of second language acquisition than the stage when the target phonology is acquired. Furthermore, the error rate drops drastically after instruction to indicate that an explicit instruction is effective to phonological acquisition of the learner speech.

## Keywords

English loanwords into Korean, Korean learner speech of English, English rule of de-aspiration, Korean rule of *l*-alternation, learner speech, loanword phonology

## 1 Introduction

Will easier or harder be the elimination of native phonology than the acquisition of target phonology? Previous studies show that there are varying degrees of acquisition rate in L2 speech, for easier elimination of vowel insertion by L1 interference (Park & Kim, 2008), easier acquisition of stress assignment rules of L2 (Kim, 2005), and difficult acquisition of vowel reduction in L2 (Kim, Flynn & Oh, 2007). The discrepancy in results observed between tasks may be resolved by

studying the comparative rules in L1 and L2 alternation. To answer these, we make a working hypothesis that to resist L1 phonology interference is easier than to acquire the L2 phonology of alternation.

(1) HYPOTHESIS 1: Learner speech avoids the native phonology interference faster than the acquisition of target phonology rules.

The answer to the question of native versus target phonology transition may depend on the developmental aspects of learner speech, because studies have shown that learner speech progressively assimilate to the value of the target speech (L2) as the proficiency increases (Han, 2006; Kim et al., 2006). In other words, there is a possibility that learner speech in its early stage is similar to the form of L1 or loanword phonology, and later develops toward the target value as L2 fluency increases. We thus make another working hypothesis that resisting L1 phonology interference is easier than acquiring the L2 phonology of alternation.

(2) HYPOTHESIS 2: According to the increase of the L2 proficiency, the learner speech resists L1 phonology interference and acquire the L2 phonology of alternation.

Hypothesis 2 can be confirmed if a cross-sectional or longitudinal study in learner speech shows a developmental transition of decreased insertion and deletion in L2 words.

The hypothesis on developmental enhancement, however, is controversial in literature, as Scovel (1988) claims that adult learners improve insignificantly in foreign pronunciation. We, therefore, investigate our second question: Whether or not learner speech develops from the form of loanword phonology toward the form of target speech.

We investigate the developmental enhancement by using both a cross-sectional and longitudinal

methods. The longitudinal method is to test cognitive and affective factors in adult language learning -- such as the role of attention and the importance of motivation in learning, as Schmidt (1995) claims, “In order to acquire phonology, one must attend to phonology” (Schmidt, 1995: 17).

The following sections test these two hypotheses by means of a set of acoustic and perceptual phonetic experiments.

## 2 Method

In order to accomplish our objective, we administered the phonetic experiment that comprises four sub-experiments on native and learner speech in cross-sectional and longitudinal studies. The four sub-experiments are on: 1) Korean native speech of English loanwords in Korean context, 2) English native speech of the English loanwords in English context, 3) Korean learner speech of the English loanwords in English context by a cross-sectional study, 4) Korean learner speech of the English loanwords in English context by a longitudinal study. Each sub-experiment used different sets of speech materials, participants, and the data acquisition steps.

## 2.1 Speech materials

Three types of speech materials and the recording lists were used: 1) recorded L1 Korean speech materials of English loanwords in Korean context, 2) recorded L1 and L2 English speech materials of English words in English context, 3) listening test materials in English for Korean learners to serve as co-variance reference to the development in speech production.

The listening test materials were the listening component of *Test Of English for International Communication* (TOEIC). Two TOEIC listening tests were used as a standardized test to recognize the improvement of listening comprehension before and after instruction.

For recorded speech materials, the native and learner participants read words in isolation, and those embedded in English or Korean contexts. Table 1 is the recording list, in which underlined are the segments to contrast for L1 and L2 rule application. The Korean context is romanized according to *the Official Romanization System* of Korean.

Table 1: Recording list for L1 and L2 rule application in L2 English words, spoken by L1 Korean speakers.

L1 and L2 phonology	(a) Words in isolation	(b) Words in English context	(c) Words in Korean context
Korean rules of liquid alternation	1. <u>l</u> ight [r] 2. <u>o</u> nly [nn / ll] 3. p <u>l</u> ate [r]	1. We like in <u>l</u> ine skates. [r] 2. Where are they p <u>l</u> aying tennis? [r]	1. geuneun in <u>l</u> aine seukeiteu seonsuya. (He is an <i>in-line skate</i> player.) 2. oneul pe <u>u</u> lei eottaetseo? (How was <i>play</i> today?)
English rules of aspiration alternation	1. <u>a</u> pple [p=] 2. <u>a</u> pplication [p=] 3. <u>s</u> creen [k=] 4. <u>s</u> plendid [p=] 5. <u>s</u> tyle [t=]	1. I have four <u>a</u> pples.[p=] 2. Fill in the <u>a</u> pplication form.[p=] 3. I <u>s</u> creamed at him.[k=] 4. How <u>s</u> plendid the idea was.[p=] 5. I <u>s</u> tared at her. [t=]	1. eoje ae <u>p</u> eul keompyuteoreul sasseo. (I bought an <i>apple</i> computer yesterday.) 2. yeonghwa “seu <u>k</u> eurim” gaebonghaettae. (The movie “ <i>scream</i> ” was released.)

Words in Table 1 are subject to the phonological alternation of Korean and English segments. For Korean rules of liquid alternation, Korean has a single phoneme for liquid phones, that becomes [r] in onset and [l] in coda of a syllable (e.g., [r] for *l*ight and [l] for *t*ile). The phoneme triggers the assimilation of sonorants, to make the input sound sequence *nl* to either [nn] or [ll] (e.g., *o*nly). The phoneme is expected to become a geminate lateral [ll] in an onset cluster as in the word *s*lide (*Official Loanword Transcription Regulations*, Art.6, sec.1).

In comparison, English rules of alternation in Table 1 show various kinds of aspiration alternation. Stop consonants in English are not aspirated in an unstressed syllable as in *a*pple, or after an onset [s] as in *s*creen.

The recording lists in columns (a) and (c) were used to obtain L1 Korean speech materials of English loanwords in Korean context, while the lists in columns (a) and (b) were to obtain L1 and L2 English speech materials of English words in English context.

## 2.2 Participants

Research participants consisted of 11 Korean native speakers who read English loanwords in Korean context, 12 native speakers of American English and 124 Korean learners of English who read the English loanwords to Korean in English context.

All native speakers spoke the General American English, while learners the Standard

South Korean as their native language. The learner's English proficiency level varied with respect to the scale of five ranks in accordance to TOEIC score. We then putatively used the scaling of score intervals in *TOEIC Can-Do Guide* (2000) as co-variance units for our analysis.<sup>1</sup> This follows the developmental effects that the more proficient the speaker is, the more target-like the perception of the target sound is for the listener.

### 2.3 Data acquisition procedure

The data acquisition procedure varied for all four sub-experiments. For the first sub-experiment, where we acquired Korean native speech of English loanwords, we used the following three stages of eliciting the reading speech. At the first stage, the Korean native speakers were given the English word list in English alphabet, and asked to transcribe the words into Korean alphabet. The transcription was done for the words in isolation. At the next stage, the speakers were given the Korean sentence list where the English loanwords are embedded hidden in empty blanks. The speakers are then asked to fill in the blanks by copying down their own previously transcribed English loanwords in Korean alphabet. At the last stage, the speakers were asked to read the completed Korean scripts for recording the words in isolation and the words embedded in sentences.

For the second sub-experiment, where we acquired English native speech of the English loanwords in English context, the native speakers of American English underwent only one stage of the recording of the production stimuli.

For the third sub-experiment, where we acquired Korean learner speech of the English loanwords in English context by a cross-sectional study, the learners took listening and production tests after a brief listen and repeat practice of the recording stimuli, on a very first week of any class instruction to avoid experimental intervention.

For the fourth sub-experiment, where we acquired Korean learner speech of the English

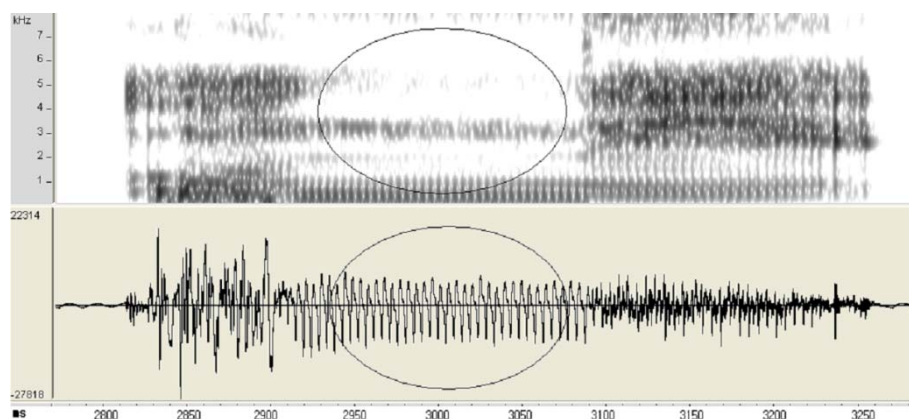


Figure 1: Illustration of an incorrectly produced nasal [n] for the target liquid [l] in L2 English word *only*, spoken by a female Korean learner.

loanwords in English context by a longitudinal study, the procedure consisted of three stages: 1) pre-listening and production tests after a brief listen and repeat practice of the recording stimuli, 2) pronunciation instruction in class for 7 weeks that included many different aspects of pronunciation including, but not focused on, avoiding the Korean l-alternation and applying the English de-aspiration rules, and 3) post-listening and production tests on the same data after 7 weeks, and again on the different data after next 7 weeks.

### 2.4 Analysis

We measured the acoustic features that are relevant to the phonological phenomena in Korean and English. To evidence the application of Korean rule of l-alternation, we identified the expected L2 phones as presented in Table 1: [r] for *light*, [nn] or [ll] for *only*, etc. Figure 1 illustrates an incorrectly produced nasal [n] for the target liquid [l] in L2 English word *only*, spoken by a L1 Korean learner. Nasal consonant quality was determined by a very low first formant centered at about 250Hz, and a large region above the first formant with no energy. This speaker has a second, rather faint, nasal formant just below 2000 Hz. We do not see a formant in the neighborhood of 1100 or 1200 Hz that is typical of most initial laterals for most speakers.

On the other hand, we also identified the expected L2 phones by not applying the English rule of de-aspiration in unstressed onset or after an onset [s]. For example, the English word "screen" in learner speech may have an aspirated [k<sup>h</sup>], unlike the target phone [k=]. The aspiration was detected by the acoustic cues of 1) silence (stop gap), 2) release burst indicated by a strong vertical spike, and 3) a long VOT of more than 30 ms as short frication noise (scattered marks after the release) before vowel formants begin.

<sup>1</sup> This edition of the *Can-do Guide* fits to our learner speech data, as we used the previous version of TOEIC listening test using the General American English. The next edition was published on 2009 to reflect the recent modification in TOEIC to contain different English accents around the world.

For analysis, we compared the learner speech with the native speech. For less clear cases of acoustic features, we used the additional means of the perceptual judgment by researchers and native speakers of English. We counted only the expected errors that are shown in the loanwords for the different L2 phones in question.

### 3 Results

A total of 2,672 word data were acquired from the recorded corpus of native and learner speech in word-level and sentence level production by 124 Korean learners of English and 11 Korean and 12 English native speakers. The results of the study are shown in Figure 2 for word-level production and Figure 3 for sentence-level production.

These figures represent the interference error rate in percentage in which either the native phonology rule applied as in (a), or the target phonology rule failed to apply as in (b). To be specific, the application rates of Korean rule of *l*-alternation are shown in (a), and the failure rates of applying the English rule of de-aspiration are shown in (b). The results in (a) and (b) are from the cross-sectional study, while those in (c) are from the longitudinal study.

In both Figures 2 and 3, the relative frequency of L1 and L2 rule interference is higher in L2, as graphically demonstrated in 2(c) for word-level production, and 3(c) for sentence-level production. At all stages in Figures 2(c) and 3(c), failure rates of L2 rule application are higher than the interference rate by L1 rule application. The results thus support our Hypothesis 1 that learner speech avoids the native phonology interference faster than the acquisition of the target phonology rules.

To move on to the developmental results in relation to Hypothesis 2, Figures in 2 for word-level production clearly supports the hypothesis that error rate decreases as the speaker proficiency increases. In Figures 2(a) and 2(b) for cross-sectional study, learner speech resists L1 phonology interference and acquire the L2 phonology of alternation according to the increase of the L2 proficiency. Figure 2(c) shows the results from the longitudinal study, in which the average error rate before instruction decreases after instruction regardless the test is on the same data or on new data. The acquisition of L2 rule shows more difficulty than the resistance to L1 rule as demonstrated in Post-test with new data. The developmental results are less obvious in Figure 3 for sentence level production, although the average rate in 3(c) shows the similar results, in which the error rate decreases as the learner proficiency increases. The acquisition of L2 rule for new data in 3(c) also shows more difficulty than the resistance to L1 rule, as in 2(c). We thus

find that the developmental results in Figures 2 and 3 support Hypothesis 2.

### 4 Discussions

The results indicate that the learner speech manifests significantly more difficult L2 rule acquisition than L1 rule resistance, in both word-level and sentence-level production. In addition, the error rate decreases according to the speech proficiency. Furthermore, the error rate drops more distinctively after instruction to indicate that an explicit instruction is effective to phonological acquisition of the learner speech.

Additional findings in this study are the following. Firstly, learner speech is far more distant from loanwords, and much closer to target phonology. In other words, learner speech is far more faithful to target phonology than loanword phonology. Secondly, another difference between learner speech and loanword phonology is that learner speech develops toward target phonology, while loanword phonology is rather constant. Thirdly, needless to mention is an additional fact that loanwords conform to native phonotactics, whereas learner speech does not.

Currently, the discussions in this study are based on the average values of a large number of data. The phonological alternation involve only the L1 rules of *l*-alternation and L2 rules of deaspiration. We plan to do a more sophisticated analysis with statistical significance.

### Acknowledgements

This work was supported in part by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (KRF-2006-321-A00933).

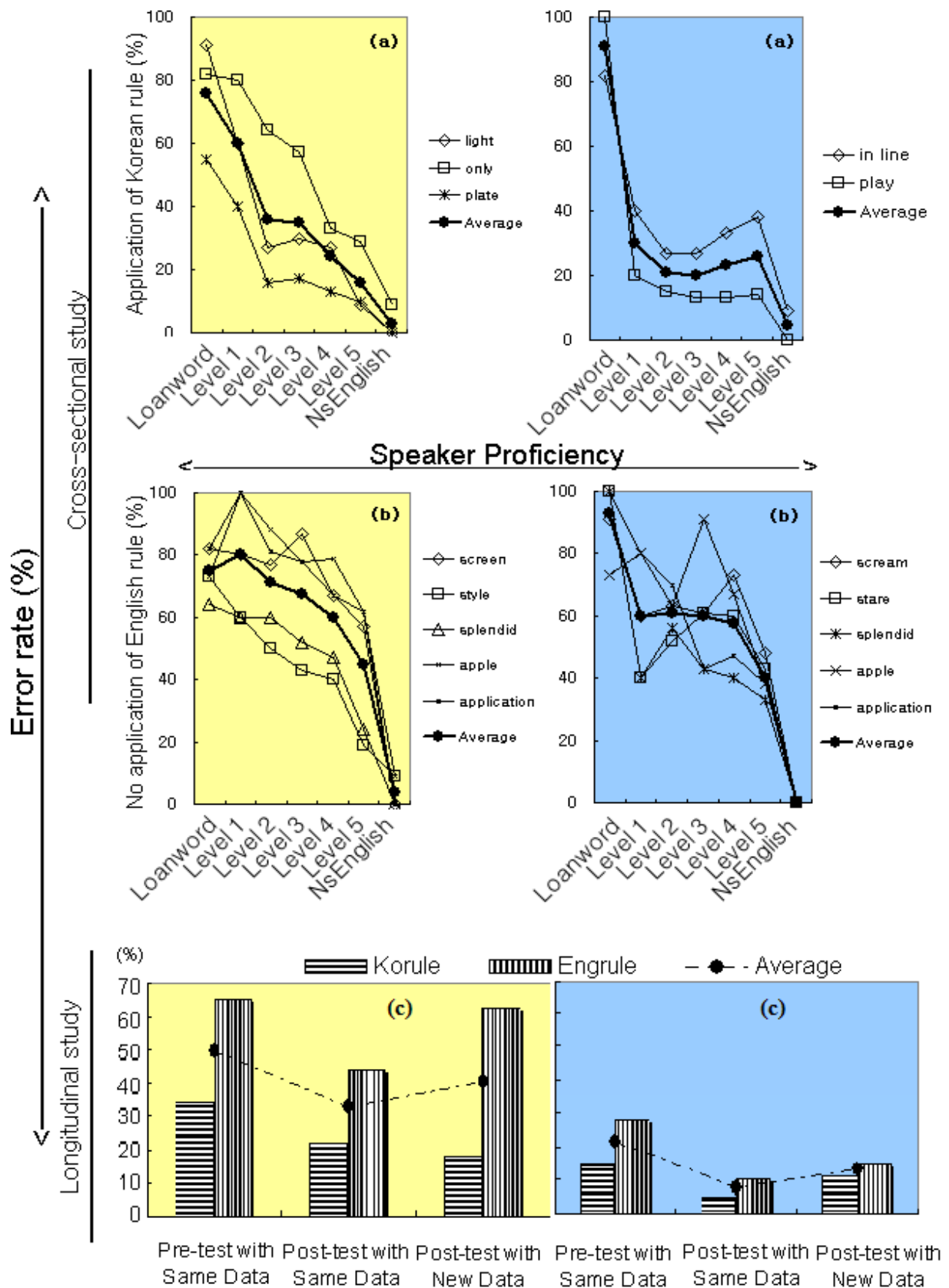


Figure 2: Error rate (%) of word-level production in L2 English speech by L1 Korean learners. L1 Korean rule of l-alternation applied in (a), while L2 English rule of de-aspiration failed to apply in (b). The results of a cross-sectional study are in (a) and (b), and the longitudinal results are in (c).

Figure 3: Error rate (%) of sentence-level production in L2 English speech by L1 Korean learners. L1 Korean rule of l-alternation applied in (a), while L2 English rule of de-aspiration failed to apply in (b). The results of a cross-sectional study are in (a) and (b), and the longitudinal results are in (c).

## References

- ETS. (2000). *TOEIC Can-Do Guide: Linking TOEIC Scores to Activities Performed Using English*. Princeton, NJ: Author. [TOEIC\_CAN\_DO.pdf, www.ets.org]
- Han, J-I. (2006). On the development of anticipatory coarticulation in Korean/English interlanguage, *Studies in Phonetics, Phonology and Morphology*, 12.1, 199-210. The Phonology-Morphology Circle of Korea.
- Kim, J-M. (2005). Stress assignment rules in Korean English, *Studies in Phonetics, Phonology and Morphology*, 11.2, 247-258. The Phonology-Morphology Circle of Korea.
- Kim, J-M., S. Flynn, & M. Oh. (2007). Non-native speech rhythm: A large-scale study of English pronunciation by Korean learners, *Studies in Phonetics, Phonology and Morphology*, 13.2, 219-250. The Phonology-Morphology Circle of Korea.
- Official Loanword Transcription Regulations*. (1995). [The current edition in official use]. Ministry of Culture and Tourism, Republic of Korea. [www.korean.go.kr].
- Official Romanization system*. (2000). [The current edition in official use]. Ministry of Culture and Tourism, Republic of Korea. [www.korean.go.kr].
- Park, M. S. & J-M. Kim. (2008). How different are vowel epenthesis in learner speech and loanword phonology, *Speech Sciences*, 15.2, 33-51. The Korean Association of Speech Sciences.
- Scovel, T. (1988). *A Time To Speak: A Psycholinguistic Inquiry Into The Critical Period For Human Speech*, New York: Newbury House.
- Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in foreign language learning. *Attention and awareness in foreign language learning* ed. by R. Schmidt, 1-63. Honolulu, HI: Univ. of Hawai'i. Second Language Teaching and Curriculum Center.

# A corpus study of passive unaccusative verbs in L2 English produced by advanced Japanese learners

Kazuharu Owada<sup>1</sup>, Victoria Muehleisen<sup>2</sup> and Hajime Tsubaki<sup>3</sup>

<sup>1</sup>Tokyo College of Music, <sup>2</sup> School of International Liberal Studies, Waseda University, <sup>3</sup> GITI / Language and Speech Science Research Laboratories, Waseda University

qwq03702@nifty.com, vicky@waseda.jp, and hjm.tsubaki@asagi.waseda.jp

## Abstract

There has been extensive research on the use of 'passive' unaccusative verbs (e.g., \*the accident *was happened*, \*the mobile phone *was appeared*) in L2 English produced by English learners of various L1s, some using judgment and elicitation tasks and some using corpora (see for example Rutherford 1998; Zobl 1989; Oshita 1997, 2000; Hirakawa 2003).

Several accounts for the passive unaccusatives have been proposed: 1) L1 transfer of a tense/aspect auxiliary verb + a past participle, 2) overgeneralization of the English adjectival passive formation, 3) non-target lexical causativisation, 4) identification of the lack of a logical subject and the be + *en*, and 5) non-target overt marking of syntactic NP movement (Oshita 2000).

This corpus study describes the actual uses of passive unaccusatives by Japanese English learners based on the method of analysis used by Oshita (2000).

## Keywords

learner corpus      passive unaccusatives

## 1 Introduction

The purpose of this study is to describe the actual uses of passive unaccusatives by Japanese English learners based on the method of analysis used by Oshita (2000). Examples of both passive unaccusatives as well as correctly-used unaccusatives are extracted from a corpus of more than 5,600 essays by college-level students, and examined for semantic and contextual factors that explain their use.

## 2 What are passive unaccusatives?

'Passive' unaccusative verbs (e.g., \*the accident *was happened*, \*the mobile phone *was appeared*) are forms which do not occur in native varieties of English, but which have been noticed in L2 English produced by English learners of various L1s. There has been extensive research on the use of

these structures, some using judgment and elicitation tasks and some using corpora (see for example Rutherford 1998; Zobl 1989; Oshita 1997, 2000; Hirakawa 2003).

## 3 Previous research on the acquisition of passive unaccusatives

### 3.1 Unaccusative verbs

It has been widely recognized that intransitive verbs are classified into two classes: unergative verbs and unaccusative verbs in the literature (e.g., Perlmutter, 1978; Burzio, 1986; Levin and Rappaport Hovav 1995).

### Unaccusative Hypothesis

- (1) a. Naomi *worked*. (unergative)  
(Agent)
- b. Naomi *arrived*. (unaccusative)  
(Theme/Patient)

Some linguists make a further distinction between alternating unaccusative verbs and non-alternating unaccusative verbs.

### Two types of unaccusative verbs

A. Alternating unaccusatives (unaccusatives with a transitive counterpart)

- (2) a. The window *broke*.  
b. Ken *broke* the window.
- (3) a. The door *opened*.  
b. Ken *opened* the window.

B. Non-alternating unaccusatives (unaccusatives without a transitive counterpart)

- (4) The leaves *fell*.
- (5) The guests *arrived*.
- (6) Something *happened*.

In English, the alternating unaccusatives can occur in passive structures, but the non-alternating unaccusatives cannot.

- (7) The door *was opened*.  
 (8) \*The leaves *were fallen*.

### 3.2 Passive unaccusatives

Second Language (L2) researchers have been conducting extensive research on why L2 learners of English produce passive unaccusatives, even though they do not normally encounter such forms in English input. Some of them have proposed a UG-based account regarding this issue.

Although most of the previous studies have used such tasks as grammatical judgment and elicitation, only a small number of them have used learner corpora. Among them, Oshita (2000) is one of the pioneering studies using learner corpora for the analysis of the passive unaccusatives produced by L2 learners of English.

## 4 Oshita (2000)

### 4.1 Five accounts for passive unaccusatives

Several accounts for the passive unaccusatives have been proposed. Oshita (2000) categorizes the previous studies on passive unaccusatives into five accounts: 1) L1 transfer of a tense/aspect auxiliary verb + a past participle, 2) overgeneralization of the English adjectival passive formation, 3) non-target lexical causativisation, 4) identification of the lack of a logical subject and the *be + en*, and 5) non-target overt marking of syntactic NP movement (Oshita 2000).

1. L1 transfer of a tense/aspect auxiliary verb + a past participle

In many languages such as Italian, which have a compound tense/aspect system, unaccusatives take a *be*-type auxiliary verbs while unergatives and transitives require a *have*-type auxiliary verb. Therefore, it is said, Italian speakers, for instance, transfer their L1 system, thus producing passive unaccusatives.

2. Overgeneralization of the English adjectival passive formation

In this account, passive unaccusatives occur based on the rule of deverbal adjectives. Therefore, passive unaccusatives should have a 'stative' meaning. However, it is noted that L2 learners rarely produce 'died person' and 'happened people.'

3. Non-target lexical causativisation

According to this explanation, non-alternating unaccusatives are causativised in L1 English. Passive unaccusatives are genuine verbal passives.

4. Identification of the lack of a logical subject and the *be + en*

Grammatical passives and unaccusative passives are the same in that they do not require a logical subject. L2 learners associate the lack of a logical subject, that is, the immediate causer of an event, with the form *be + Ven*. If non-target *be + Ven* indicates simply the lack of a logical subject, the NP does not need to move to the subject position, placing expletives such as *there* and *it* and *pro* in the subject position instead.

5. Non-target overt marking of syntactic NP movement

In this account, L2 learners move an internal argument within the unaccusative to the subject position as native speakers do, but mistakenly and overtly mark *be + Ven*.

### 4.2 The corpus study

To test which account best describes the acquisition of passive unaccusatives, he analyzes the passive unaccusatives using the ten syntactic patterns for classification. The corpus used in his study is the Longman Learners Corpus (Version 1.1., March 1993, henceforth LLC), which consists of 3,362 essays written by four L1 English learners: 684 for Italian, 1,079 for Spanish, 236 for Korean and 1,363 for Japanese.

The target unalternating unaccusatives and the syntactic patterns are show below:

- (9) The target non-alternating unaccusatives

*appear, arise, arrive, die, disappear, exist, fall, happen, occur, rise*

- (10) The syntactic patterns for classification of token sentences with unaccusatives

- a. NP-V
- b. NP-*be*-*Ven*
- c. *there*-V-NP
- d. *it*-V-NP
- e. *pro*-V-NP
- f. *there-be-Ven*-NP
- g. *it-be-Ven*-NP
- h. *pro-be-Ven*-NP
- i. NP<sub>1</sub>-V-NP<sub>2</sub>
- j. *there-be*-NP-V

### 4.3 The results and discussion

The results in Ohita (2000) are summarized below.

1. L1 transfer of a tense/aspect auxiliary verb + a past participle

Since learners whose L1 does not have a compound tense/aspect system produce passive unaccusatives, this account is untenable.

2. Overgeneralization of the English adjectival passive formation

In most cases it is difficult to interpret the passive unaccusatives as 'stative.' Out of 38 passive unaccusatives extracted from the corpus, only a few have 'stative' meanings.

3. Non-target lexical causativisation

The ratio of non-target causatives to passive unaccusatives is 38: 11. Since the frequency of non-target causatives is low, causativisation does not always occur.

4. Identification of the lack of a logical subject and the *be + en*

Grammatical passives and unaccusative passives are the same in that they do not require a logical subject. L2 learners associate the lack of a logical subject, that is, the immediate causer of an event, with the form *be + Ven*. If non-target *be + Ven* indicates simply the lack of a logical subject, the NP does not need to move to the subject position, placing expletives such as *there* and *it* and *pro* in the subject position instead.

5. Non-target overt marking of syntactic NP

The fact that there are a very few instances of 'expletive-V-NP' indicates that in most cases the NP movement occurs.

Based on a study of the LLC, he argues that non-target marking of NP movement provides the most valid explanation for the passive unaccusatives.

Although Ohita (2000) analyzed the data produced by Italian, Spanish, Korean, and Japanese speakers, in the following section, we will restrict ourselves to his Japanese data so that we can make a comparison.

## 5 This study

### 5.1 SILS English Learner Corpus

In this study we use a corpus called the SILS

English Learner Corpus, a corpus compiled at the School of International Studies at Waseda University. It contains more than 5,600 essays by college-level students. In this study first-draft essays written by Japanese students have been extracted for the analysis.

The tokens are 3,600,607. The types are 53,551. The type/token ratio is about 0.015. Although no information is given on the number of types and tokens in the Japanese data which is made up of 1,363 essays in Ohita (2000), the SILS corpus is larger in its scope.

### 5.2 The target verbs

The same non-alternating unaccusative verbs used in Ohita (2000) are analyzed.

- (11) The target non-alternating unaccusatives in this study

*appear, arise, arrive, die, disappear, exist, fall, happen, occur, rise*

### 5.3 The results

#### 5.3.1 Syntactic patterns for the unaccusatives

We can see from Table 1 that out of 4,609 sentence patterns in which the unaccusatives appear, 169 (3.7%) are passive unaccusatives. This is similar to the ratio in Ohita (2000), which is 5.6% (17/304).

Table 1. The frequency of the syntactic patterns in which the target unaccusatives appear

Structural patterns	Oshita (2000)	this study
a. NP-V	269	4387
b. NP- <i>be</i> - <i>Ven</i>	17	169
c. <i>there</i> -V-NP	1	43
d. <i>it</i> -V-NP	1	0
e. <i>pro</i> -V-NP	1	0
f. <i>there</i> - <i>be</i> - <i>Ven</i> -NP	4	0
g. <i>it</i> - <i>be</i> - <i>Ven</i> -NP	0	0
h. <i>pro</i> - <i>be</i> - <i>Ven</i> -NP	0	0
i. NP <sub>1</sub> -V-NP <sub>2</sub>	7	10
j. <i>there</i> - <i>be</i> -NP-V	4	0
Total	304	4609

Next we will focus on the three syntactic patterns: NP-V, NP-*be*-*Ven* and NP<sub>1</sub>-NP<sub>2</sub>, as can be seen from Table 2 and Figure 1.

Table 2. The breakdown of three syntactic patterns of each unaccusative

	NP-V	NP-be-Ven	NP <sub>1</sub> -V-NP <sub>2</sub>
appear	864	51	0
arise	63	2	2
arrive	37	0	0
die	1367	11	0
disappear	203	17	0
exist	189	16	0
fall	208	1	1
happen	582	32	2
occur	555	38	0
rise	319	1	5
Total	4387	169	10

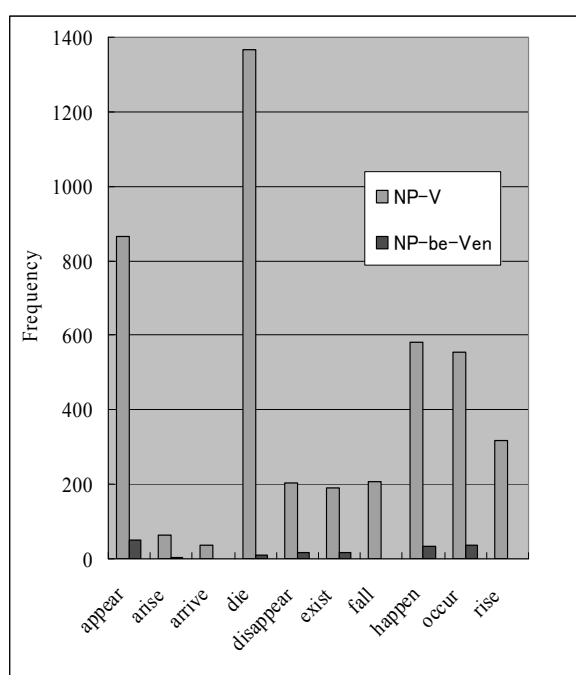


Figure 1. The breakdown of three syntactic patterns of each unaccusative

Among the ten unaccusative verbs investigated, 'appear' was the most frequently used in the passive construction (51 cases), followed by 'occur' (38 cases) and 'happen' (32 cases). We will present some examples of those unaccusative passives extracted from the corpus.

### 5.3.2 Some example of passive unaccusatives and non-target causativisation

Some examples of passives unaccusative are shown below.

#### appear

(12) We human beings, are appeared about 4~5

million years ago. (age: 19, TOEFL CBT: 227, class level: advanced)

(13) TV personalities are appeared in both TV commercials. (age: 18, TOEFL PBT: 503, class level: advanced)

(14) Furthermore, a lot of Latino directors are appeared in Hollywood and it helps to take away stereotypes of Latino. (age: 20, TOEFL CBT: 205, class: advanced)

As shown above, 'are appeared' in (12) is meant to be 'appeared,' while 'are appeared' in (13) and (14) should have been 'appear.'

#### occur

(15) In conclusion, bullying is occurred both in Japan and the West but why students carry out bullying is a bit different. (age: 19, TOEFL PBT: NA, class level: intermediate)

(16) However if Japanese President would speak English by spreading English, meeting could do more smoothly and miscommunication which is occurred by interpreter would be disappeared. (age 18: TOEFL PBT: 505, class level: intermediate)

(17) The tragic case is occurred in all over the world. (age: 19, TOEFL: NA, class level: intermediate)

In both (15) and (16) 'is occurred' can be interpreted as 'is caused,' while in (17) 'is occurred' should mean 'occurs' or 'has occurred.'

#### happen

(18) Children in western country think that bullying means strong one offend weak one, on the contrary, children in Japan think that bullying is happened by group and one person. (age: 20, TOEFL PBT: NA, class level: intermediate)

(19) Others say single-sex education is better than coeducation. This discussion is happened because of some differences between coeducation and single-sex education. (age: 19, TOEFL PBT: 547, class level: intermediate)

(20) This phenomenon is happened because Arab people do not have much power to change the images in Hollywood compared to other racial minorities like blacks. (age: 18, TOEFL PBT: 560, class level: advanced)

In (18) 'is happened' may mean 'is caused' because of the *by*-phrase. In both (19) and (20) 'is happened' should mean 'happens.'

As for the causitivation of unaccusatives, only 10 cases are observed. Among them, that of 'arise' and 'rise' may be due to the confusion with the verb 'raise.' The two cases of the causitivation of 'happen' are shown below:

happen

(21) The only global language will happen innovation. (age: 19, TOEFL PBT: 453, class level: intermediate)

(22) This spread of race may be one of the advantages for Disney, but this situation happened some problems. (age: 18, TOEFL PBT: 477, class level: advanced)

### 5.3.3 The tenses used in 'appear'

Let us look at the tenses used in the most frequently used unaccusative 'appear.' Table 3 shows the tense/aspect of 'appear.'

Table 3. The tense/aspect of 'appear'

Tense/aspect pattern	freq.
appear	469
appears	136
appeared	232
am appearing	0
is appearing	5
are appearing	5
have appeared	0
has appeared	11
have been appearing	2
has been appearing	0
had appeared	3
had been appearing	1
be appearing	0
was appearing	0
were appearing	0
Total	864

Table 4 shows the passive unaccusative use of 'appear' in terms of tense/aspect.

Table 4. The tense/aspect of 'be appeared'

Tense/aspect pattern	freq.
be appeared	12
is appeared	12
am appeared	0
are appeared	14
has been appeared	1
have been appeared	2
was appeared	8
were appeared	2
had been appeared	0
Total	51

Out of 51 instances of the passive unaccusative, 38 instances use non-past simple tense. This result may indicate that the learners use non-past simple tense of 'be appeared' as a cross between non-past simple tense of 'appear' and non-past perfect simple tense of 'have appeared.' This result might also mean that the learners interpreted the non-past simple tense of 'be appeared' as having a 'stative' meaning. This issue needs to be investigated further.

## 6 Conclusion

Although this study does not provide a definitive answer as to which account best describes passive unaccusative, several findings can be drawn from it.

First, even advanced-level Japanese learners produce passive unaccusatives (3.7%; 169 instances out of 4,609 sentence patterns in which the unaccusatives appear. Secondly, among the ten passive unaccusatives, 'appear' was the most frequent (51 instances), followed by 'occur' (38 instances) and 'happen' (32 instances). Thirdly, the learner data allows researchers to interpret passive unaccusatives in more than one way, as we have seen in the case of 'are appeared.' Fourth, the causitivation of unaccusatives is a rare phenomenon. And finally, more detailed research is needed to determine what semantically motivates the learners to produce passive unaccusative.

## References

- Balcom, P. (1997). Why is this happened? Passive morphology and unaccusativity. *Second Language Research*, 13(1), 1-9.
- Burzio, L. (1986). *Italian syntax: A government-binding approach*. Dordrecht: Redidel.
- Hirakawa, M. (1995). L2 acquisition of English unaccusative constructions. *Proceedings of BUCLD 19*, 291-302.
- Hirakawa, M. (1997). On the unaccusative/

- unergative distinction in SLA. *JACET BULLETIN* 28: 17-27
- Hirakawa, M. (2003). *Unaccusativity in second language Japanese and English*. Tokyo: Histuzishobo.
- Levin, B. and M. Rappaport Hovav. (1995). *Unaccusativity: at the syntax-lexical semantics interface*. Cambridge: MIT Press.
- Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago: Chicago University Press.
- Oshita, H. (1997). *The unaccusativity trap: L2 acquisition of English intransitive verbs*. Ph.D. dissertation, University of Southern California.
- Oshita, H. (2000). What is happened may not be what appears to be happening: a corpus study of 'passive' unaccusatives in L2 English. *Second Language Research* 16, 293–324.
- Perlmutter, D. (1978). Impersonal passives and the unaccusative hypothesis. *Proceedings of the 4th Annual Meeting, Berkeley Linguistics Society*, 157-89.
- Rutherford, W. (1998). *A workbook in the structure of English: Linguistic principles and language acquisition*. Oxford: Blackwell
- Shomura, Y. (1996). A study of the acquisition of ergatives by Japanese learners. *Proceedings of the JACET 35th Annual Convention*, 229-32.
- Sorace, A. (1993). Unaccusativity and auxiliary choice in non-native grammars of Italian and French: Asymmetries and predictable indeterminacy. *French Language Studies* 3, 71- 93.
- Sorace, A. (2000). Gradients in auxiliary selection with intransitive verbs. *Language* 76: 859–890.
- Zobl, H. (1989). Canonical typological structures and ergativity in English L2 acquisition. In S. Gass & J. Schachter (Eds.), *Linguistic perspectives on second language acquisition* (pp. 203–221). New York: Cambridge University Press.

# A corpus analysis of intransitive verbs used in junior high school English textbooks in Japan

Kazuharu Owada<sup>1</sup>, Hajime Tsubaki<sup>2</sup>, and Daisuke Yoshizaki<sup>3</sup>

<sup>1</sup>Tokyo College of Music, <sup>2&3</sup>GITI / Language and Speech Science Research Laboratories, Waseda University

qwq03702@nifty.com, hjm.tsubaki@asagi.waseda.jp and

d\_yoshizaki@toki.waseda.jp

## Abstract

There has been a great deal of research in corpus studies of textbooks and learner data (see Ishikawa 2009 for recent developments in this field). Since textbooks used in junior and senior high schools provide the students with crucial portions of input, it is important that the textbooks present verbs so that learners can notice the difference between intransitive and transitive verbs.

In the school year 2009-2010 six kinds of junior high school textbooks published by six publishing companies are currently used in Japan, although the market share of each series widely varies. These textbook series are: *Columbus*, *New Crown*, *New Horizon*, *One World*, *Sunshine*, and *Total English*. Since one company publishes three textbooks for the 1st year, the 2nd year and the 3rd year, we create a junior high school textbook corpus based on 18 textbooks.

In this paper we will first present the frequency data of English intransitive verbs in junior high school textbooks, and secondly to describe software that outputs the frequency information of the intransitive/transitive uses.

## Keywords

junior high school textbooks intransitive verbs corpus analysis

## 1 Introduction

The purpose of this study is twofold: 1) to examine the frequency data of English intransitive verbs such as unaccusative verbs and ergative verbs in junior high school textbooks currently used in Japan, and 2) to show software that outputs the frequency information of the transitive/intransitive uses of verbs.

## 2 English intransitive verbs

It has been widely recognized that intransitive verbs are classified into two classes: unergative verbs and

unaccusative verbs in the literature (e.g., Perlmutter, 1978; Burzio, 1986; Levin and Rappaport Hovav 1996).

## Unaccusative Hypothesis

- (1) a. Naomi *worked*. (unergative)  
(Agent)
- b. Naomi *arrived*. (unaccusative)  
(Theme/Patient)

Some linguists make a further distinction between alternating unaccusative verbs and non-alternating unaccusative verbs.

## Two types of unaccusative verbs

A. Alternating unaccusatives (unaccusatives with a transitive counterpart)

- (2) a. The window *broke*.
- b. Ken *broke* the window.
- (3) a. The door *opened*.
- b. Ken *opened* the window.

B. Non-alternating unaccusatives (unaccusatives without a transitive counterpart)

- (4) The leaves *fell*.
- (5) The guests *arrived*.
- (6) Something *happened*.

## 3 Junior high school textbooks in Japan

### 3.1 Six textbook series

As of the school year 2009-2010, six kinds of junior high school textbooks are used in Japan. These textbook series are: *Columbus* (1-3), *New Crown* (1-3), *New Horizon* (1-3), *One World* (1-3), *Sunshine* (1-3) and *Total English* (1-3).

### 3.2 Intransitive/transitive verbs

None of the textbooks explicitly use the grammatical terms such as ‘intransitive’ and ‘transitive’. Out of the six textbook series, only two series (*Columbus* and *New Crown*) draw the learners’ attention to the distinction between intransitive and transitive verbs. These textbook series briefly mention that there are two kinds of verbs: those that have an object and those that do not.

As for the instruction of passive constructions, all the six series simply state that passives are formed using ‘be + past participle.’ Most of them present *sareteiru* and *sareta* as the Japanese translation for ‘be + past participle’ without explicitly stating that only transitive verbs are converted into the passive voice in English.

Japanese learners of English should be taught that although some intransitive verbs can be passivized in Japanese, this is not the case in English.

## 4 This study

### 4.1 The data

Junior high school textbook series currently used in Japan are: *Columbus*, *New Crown*, *New Horizon*, *One World*, *Sunshine*, and *Total English*. We created a junior high school textbook corpus based on 18 textbooks. Although we made use of the accompanying data files produced by the publishing companies, we edited the text data so that we could use them as a corpus data source. For instance, we deleted the lyrics included in the text files.

We ran the computer program in the *JACET list of 8000 basic words* and found that this textbook corpus contains 3,996 types (indexes) and 94,226 tokens.

### 4.2 The target verbs

The target verbs are categorized into four classes mainly based on the verb classification in *Collins Cobuild intermediate English grammar* (2005) and previous studies (see for example Oshita 1997, Hirakawa 2003). The verbs which do appear in the text corpus are omitted from the list.

#### A. Alternating unaccusative verbs<sup>1</sup> (17 verbs)

*bake, begin, break, close, drop, end, finish, grow, melt, move, open, rest, shake, stand, start, stop, turn*

#### B. Non-alternating unaccusative verbs (11 verbs)

*appear, arrive, come, die, disappear, fall, go, happen, live, rise, stay*

#### C. Unergative verbs (8 verbs)

*cry, dance, jump, laugh, sleep, smile, swim, work*

#### D. Transitive verbs which can omit an object (28 verbs)

*answer, change, choose, clean, cook, draw, drive, eat, explain, forget, help, know, learn, leave, paint, play, read, remember, ride, sing, speak, steal, study, type, understand, wash, watch, write*

Although we are well aware that we need to consider the semantic role or argument structure level when we classify verbs, we will restrict ourselves to the intransitive/transitive distinction in this study.

### 4.3 The results

For the following analysis, we used AntConc3.2.1. The frequency of verbs is shown in the following tables.

Table 1. Alternating unaccusative verbs (17 verbs)

alternating unaccusative verbs	intransitive (freq.)	transitive (freq.)	Total
bake	0	7	7
begin	23	21	44
break	4	11	15
close	4	18	22
drop	1	4	5
end	3	1	4
finish	3	23	26
grow	20	15	35
melt	2	0	2
move	14	1	15
open	7	18	25
rest	3	0	3
shake	0	1	1
stand	21	1	22
start	23	35	58
stop	5	16	21
turn	31	0	31

We can see from Table 1 that some verbs (e.g., *turn, move*) tend to be used intransitively, while some verbs (e.g., *finish, bake*) are likely to be used transitively.

Table 2 shows that ‘go’ is the most frequently used unaccusative verb.

<sup>1</sup> These verbs are called ‘ergative verbs’ in *Collins Cobuild intermediate English grammar* (2005).

Table 2. Non-alternating unaccusative verbs  
(11 verbs)

unaccusative verbs	freq.
appear	11
arrive	31
come	109
die	47
disappear	14
fall	24
go	586
happen	19
live	177
rise	1
stay	73

Table 3 shows that ‘work’ is the most frequently used unergative verb.

Table 3. Unergative verbs (8 verbs)

unergative verbs	freq.
cry	26
dance	13
jump	3
laugh	8
sleep	22
smile	18
swim	37
work	93

As can be seen in Table 4 below, the top five of the most frequently used verbs are *play*, *know*, *study*, *help*, and *eat*. All these verbs tend to be used transitively. Given this level of ample input, learners may not have difficulties with using the transitive form. However, considering the relative low frequency of the intransitive form, learners need to be explicitly taught how to use the intransitive form

Table 4. Transitive verbs which can omit an object  
(28 verbs)

transitive verbs	intransitive (freq.)	transitive (freq.)	Total
answer	4	19	23
change	12	19	31
choose	0	2	2
clean	0	26	26
cook	10	25	35
draw	0	15	15
drive	4	1	5
eat	17	116	133
explain	2	4	6
forget	0	19	19
help	5	129	134
know	59	168	227
learn	23	83	106
leave	9	36	45
paint	0	7	7
play	38	252	290
read	3	111	114
remember	2	16	18
ride	2	24	26
sing	34	33	67
speak	28	72	100
steal	0	4	4
study	36	120	156
type	0	3	3
understand	5	43	48
wash	0	4	4
watch	3	103	106
write	18	99	117

## 5 Developing software

With the results in the section 4 in mind, we will develop software which outputs the frequency data of intransitive/transitive verbs. This software will extract lexical information based on two functions: the frequency of intransitive/transitive verbs in each textbook and the concordance lines in the textbook corpus

The first function is to show the frequency of verbs in each textbook, as in Figure 1.

The interface shows a 'Verb input area' with a text box containing 'xxxxxx' and a 'Word frequency' button. Below this, there are sections for different textbooks: 'NEW CROWN', 'NEW HORIZON', 'Columbus', 'ONE WORLD', 'Sunshine', and 'TOTAL'. Each section has a 'Word frequency' button and a table with rows for 'Total', 'Intransitive verbs', 'Transitive verbs', 'Active', and 'Passive'. The 'ALL TEXTBOOKS' section also has a 'Word frequency' button and a similar table. The tables contain placeholder text like 'xxxxxx', 'xxx', and 'xxx'.

Figure 1. The frequency of verbs in each textbook

When you input a certain verb into the *Verb input area* and then press the enter key, its

frequency is shown in accordance with the classification of verb attributions of each textbook. Additionally, the total frequency is displayed in the *ALL TEXTBOOKS* section.

The second function is to display concordance lines in the textbook corpus, as in Figure 2.

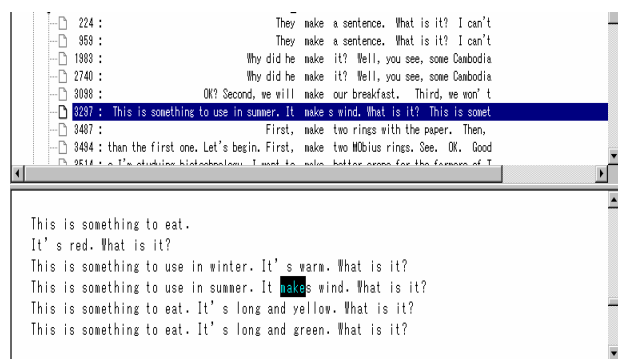


Figure 2. Concordance lines in the textbook corpus

The classification of verbs is based on the rules related to a morphological analysis. This function is to retrieve occurrence lines containing the target verb and display the occurrence line in the textbook corpus. Figure 2 shows the results of the verb *make*.

Only sentence lines which include *make* are displayed on the *Retrieval view* in the concordance format. When you click one line, the corresponding sentence line in the textbook corpus is shown on the *Text view*. This function is executed by a file-retrieval-software with a text viewer called KWIC Finder ([http://www31.ocn.ne.jp/~h\\_ishida/](http://www31.ocn.ne.jp/~h_ishida/)).

## 6 Conclusion

So far we have seen that junior high school textbooks do not provide learners with enough information on the intransitive/transitive distinction.

With software specifically designed for the intransitive/transitive verbs, we can make full use of its output in the areas of material development and grammar instruction.

## References

- Burzio, L. (1986). *Italian syntax: A government-binding approach*. Dordrecht: Redidel.
- Collins Cobuild intermediate English grammar (2nd ed.) (2005). Glasgow: Harper Collins Publishers.
- Hirakawa, M. (2003). *Unaccusativity in second language Japanese and English*. Tokyo: Histuzishobo.
- Ishikawa, S. (2008). *Eigo copasu to gengo kyoiku* [English corpus and language education]. Tokyo: Taishukan Shoten.
- Japan Association of College English Teachers (JACET) Basic Words Revision Committee

(Ed.). (2003). *JACET list of 8000 basic words*. Tokyo: JACET.

Laurence, A. (2007). AntConc3.2.1. <[http://www.antlab.sci.waseda.ac.jp/antcon\\_c\\_index.html](http://www.antlab.sci.waseda.ac.jp/antcon_c_index.html)>

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago: Chicago University Press.

Levin, B. and M. Rappaport Hovav. (1995). *Unaccusativity: at the syntax-lexical semantics interface*. Cambridge: MIT Press.

Oshita, H. (1997). *The unaccusativity trap: L2 acquisition of English intransitive verbs*. Ph.D. dissertation, University of Southern California.

Perlmutter, D. (1978). Impersonal passives and the unaccusative hypothesis. *Proceedings of the 4th Annual Meeting*, Berkeley Linguistics Society, 157-89.

# Effectiveness of Student Perception on Need-based Grouping

Handan Girginer

School of Foreign Languages, Anadolu University

hgirginer@anadolu.edu.tr

## Abstract

This study was carried out in the English Speaking Skills course designed for flight training and air traffic control departments students. The started out as a personal project of mine to research the positive effects need-based grouping. From my own experience I felt positive especially from my own motivation. Therefore, I investigated the validity of these proportions to satisfy personal curiosity. I would like to start my presentation with an explanation of how this study came about. I am teaching an English Speaking Course at the School of Civil Aviation. intelligibility in different social settings necessary for their future careers.

## Keywords

Keywords are your own designated keywords

## Introduction

The focus of the English Speaking Skills Course is improving the students' speaking and listening skills. The aim of the course is to enable the students to learn how to communicate in English in different social settings and in their future careers with a reasonable degree of fluency, accuracy and intelligibility. This course is designed to help students develop their spoken English. It also makes students speak effectively, accurately and fluently in academic situations. The course takes place at the Aviation School. It is a 4 hour- a - week course, and it takes 24 weeks spanning both fall and spring semesters.

The syllabus is designed according to the needs of the students to refresh and improve their English after taking a one year intensive English program at Prep school.

## 1 Purpose of the Study

The study started out as personal project of the class teacher to research the positive effects of need-based grouping. From my own experience I felt positive especially form y own motivation. Therefore, I investigated the validity of these proportions to satisfy personal curiosity for an English Speaking Skills Course for students of the

Flight Training and Air Traffic Control departments at the School of Civil Aviation, Anadolu University.

## 2 Research Questions

The research questions examined in this study were:

1. What are students' perceptions of need-based learning?
2. How do psychological and instructional items compare?

## 3 Data collection

### 3.1 Participants /Informants

20 students – took English Speaking Skills Course in 2008-2009, 10 in Flight Training;10 in Air Traffic Control

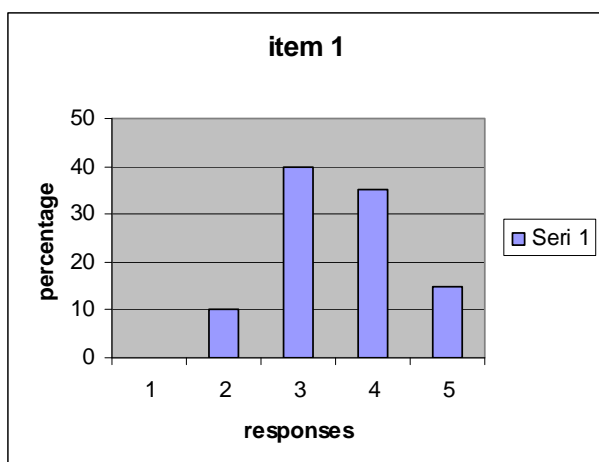
### 3.2 The instrument

A questionnaire was distributed by the researcher and used in the study to gather data.

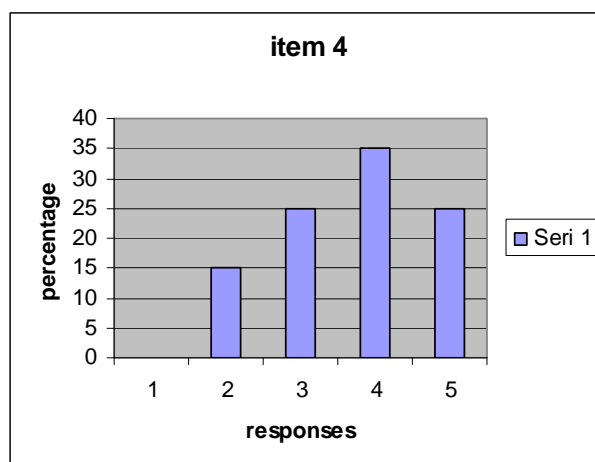
The items of the questionnaire were written by the researcher with the references used in other studies (Lui, 2008; Chen, Lin & Feng, 2004; Chien & Ching, 2002; Yu, 1994) Items from the instrument were mainly designed to measure how students perceive the effects of need-based grouping on two dimensions: psychological factors (their confidence, motivation, interest, academis pressure) and instructional factors (speaking and listening ability, teaching materials and the teacher's instructional method). See Appendix A.

## 4 Findings and discussion

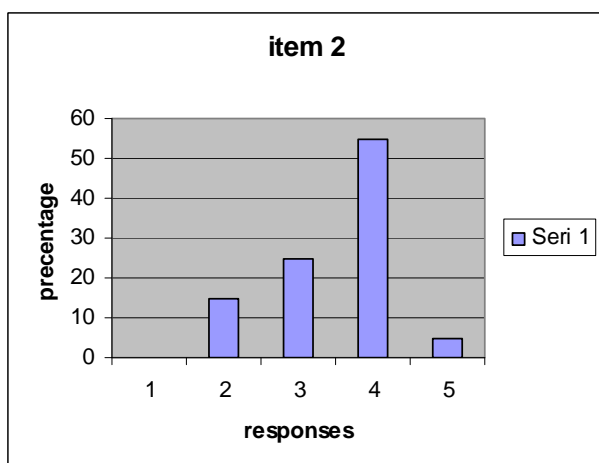
Need-based grouping is beneficial to my English learning



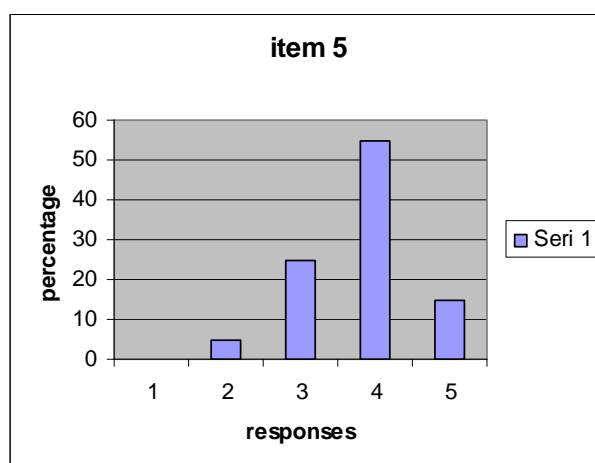
Need-based grouping helps me feel less pressure to learn English



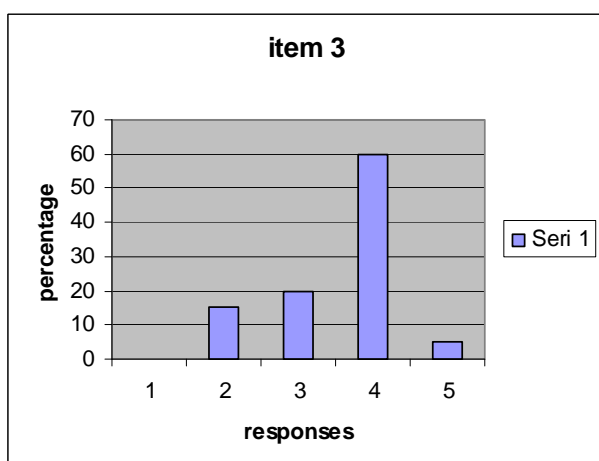
After I have been grouped by need, the teaching materials used are more suitable for me.



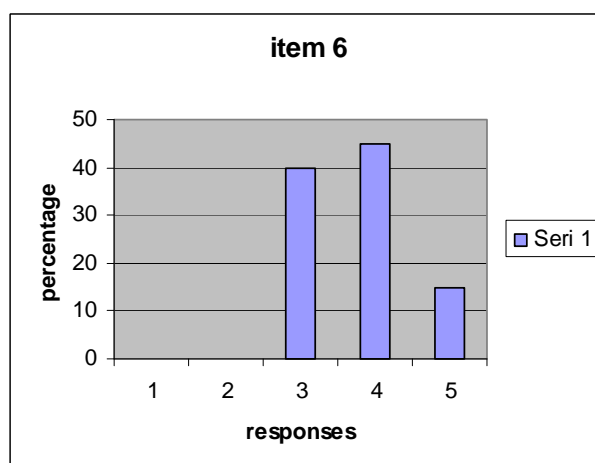
Need-based grouping helps me build more confidence in learning English.



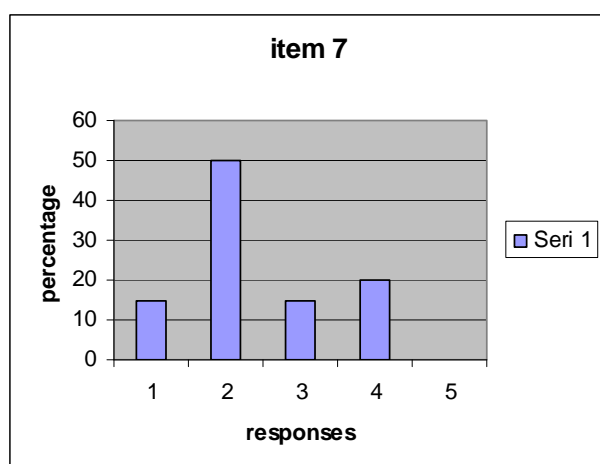
After I have been grouped by need, the teaching materials used help me become more interested in learning.



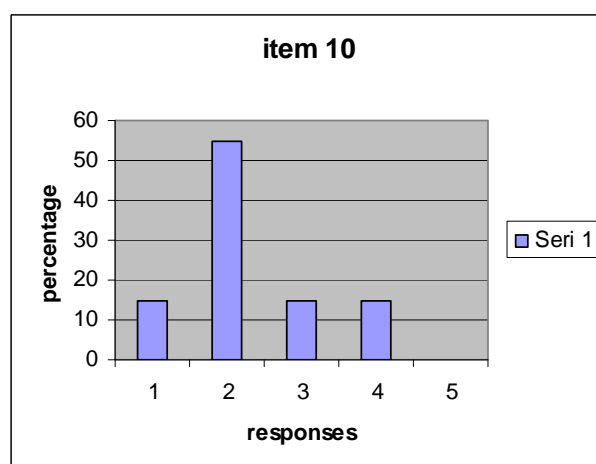
Need-based grouping helps increase my motivation in learning English.



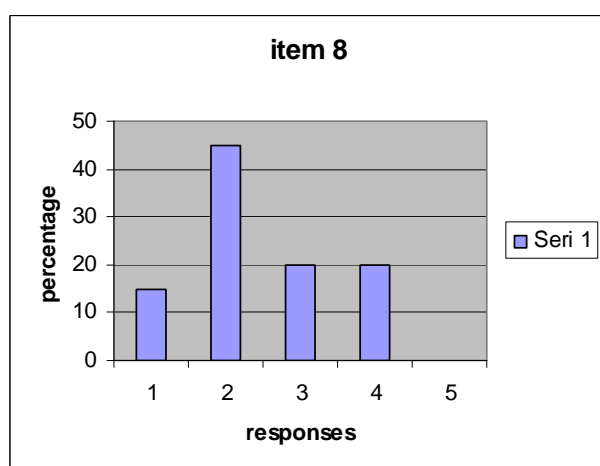
The key to further improvement in English is my learning attitudes, not need-based grouping.



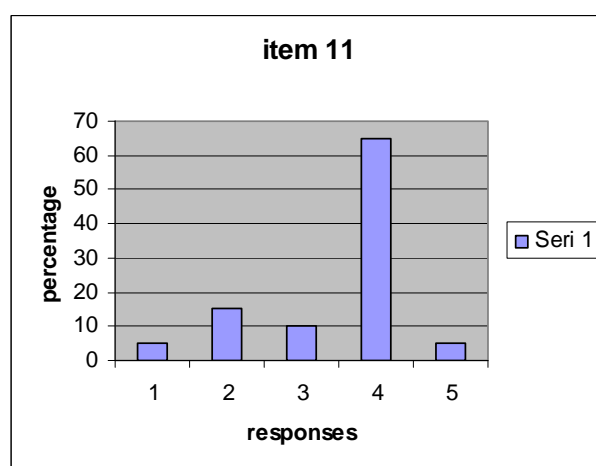
The key to further improvement in English is my learning strategies, not need-based grouping.



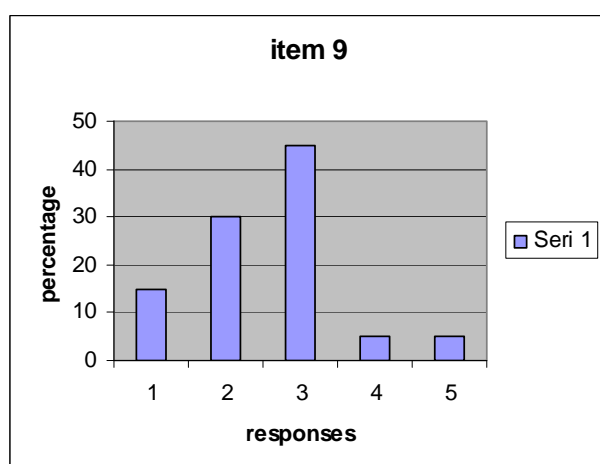
Need-based grouping benefits my English listening proficiency level.



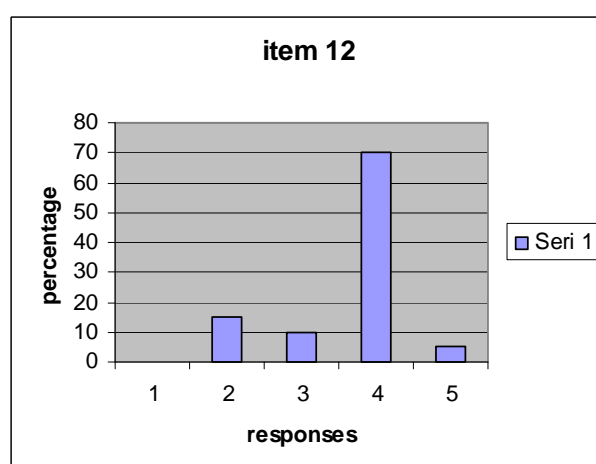
The key to further improvement in English is teachers' instructional methods, not need-based grouping



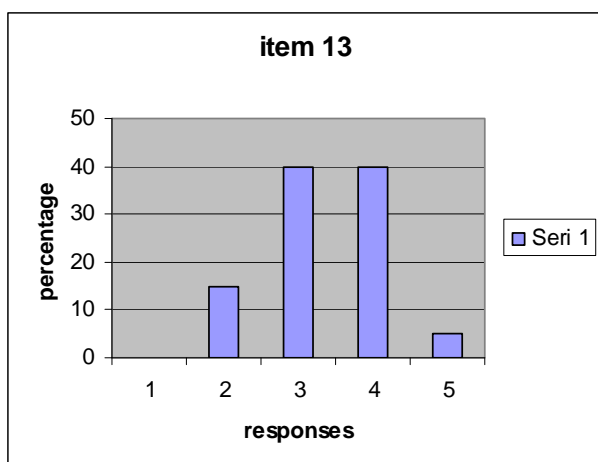
Need-based grouping benefits my English speaking proficiency level



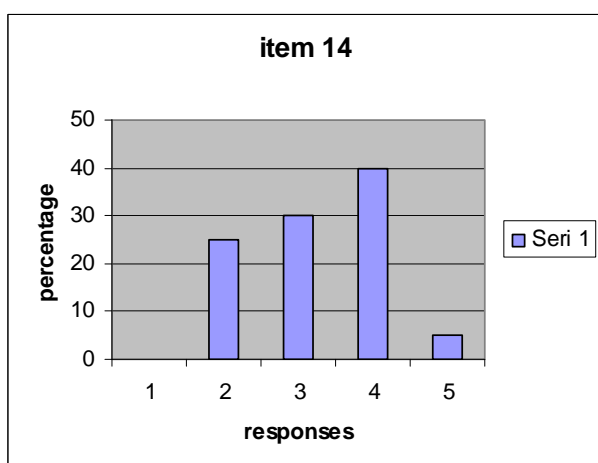
Need-based grouping causes me more anxiety over learning English.



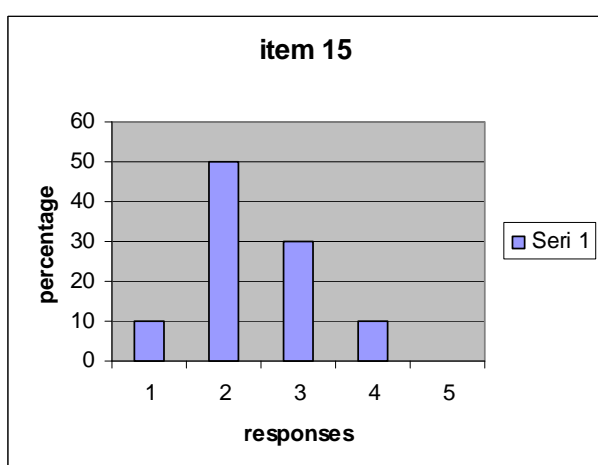
I feel I am appropriately placed in the group that matches my English listening proficiency level.



I feel I am appropriately placed in the group that matches my English speaking proficiency level.



Need-based grouping does not have any impact on the improvement of my English speaking/listening level.



NBG: Need-based grouping

Items 2 and 10 confirms that students tend to feel less pressure and anxiety when grouped by their

needs.

As shown by items 5 and 6, students report a higher suitability of the teaching materials when their needs are taken into account. Moreover, they find the materials more interesting. In brief, the application of NBG seems to serve students much beter vis-a-vis appropriateness and inducing enthusiasim.

In the line with the responses to 3, 4, 2, and 10 (which are related to the student affective factors) items 7 and 8 indicate that NBG is much more important for students's linguistic improvement than both learning attitudes and strategies. Besides, as shown by the responses to item 9, NBG is seen as a more important factor in linguistic proficiency enhancement than teacher's instructional methods.

Analyzed from a skill-based perspective, students consistently report a high degree of positive impact of NBG on their speaking and listening proficiency. Only a small number of students (12%) report some negative effects of NBG

Items 13, 14 and 15 demonstrate that students strongly feel that they are appropriately placed in the right group and NBG helps them improve their English.

Item 1 well summarizes the overall finding of this study: that is, NBG is found by the students very beneficial. This beneft is twofold: the first is affective ( as shown by the responses to items 1, 2, 3, 4, 7, 10), the second is instructional (as shown by the responses to items 5, 6, 8, 9, 11, 12, 13, 14, 15). While NBG id found helpful and effective in both dimensions, affective factors seem to have more useful impact on student learning.

## 5 Implications for teaching

I strongly believe that such queries (questionnaires to determine student perceptions) can also be successfully applied to other language teaching situations as well.

Besides the findings I discussed above I also noticed a higher awereness of the issues and class activities more attention given to the tasks at hand. Through the implementation of NBG, I have also observed a lowered affective filter on the students's part.

## References

Lui, Hui-ju, (2008). An analysis of the effects of the grouping on student learning in university-wide English classes. Feng Chia Journal of Humanities and Social Sciences. 16, 217- 249.

## Elaboration of Tasks as Means to Mediate Classroom Learning and the Outside World: Analyses of Language Learning Activities in Three EFL Classrooms for Children

Hirokatsu Kawashima<sup>1</sup>, Naoki Sugino<sup>2</sup>, and Yuya Koga<sup>3</sup>

Department of English, Nagasaki University of Foreign Studies<sup>1</sup>, College of Information Science and Engineering, Ritsumeikan University<sup>2</sup>, Department of Child Education, Shukugawa Gakuin College<sup>3</sup>

kawashima@tc.nagasaki-gaigo.ac.jp,

gwisno@is.ritsumei.ac.jp, koga@shukugawa-c.ac.jp

### Abstract

Task-based language teaching/learning is commonly seen in EFL classrooms for children, and recent theoretical developments from the sociocultural perspective have suggested that tasks should function as means to link learners' language learning experience in the classroom with their real-world experience (e.g., Samuda and Bygate, 2008, p. 18). In an attempt to grasp how language learning tasks can be transformed and developed to fulfill such a function, the presenters have been conducting a series of descriptive and comparative studies of three Japanese EFL classrooms for children for the past several years.

One of their preliminary studies (Sugino et al., 2008) has shown, for example, that a series of language learning tasks used in a content-oriented classroom were woven together to function as means to mediate between the classroom and the outside world, and that the selection of their topic/content, which had personal relevance with the children's experiences in their immediate environments, played a crucial role in elaborating the series of tasks.

However, it is still unclear exactly how selected tasks can be elaborated in the classroom, because such an elaboration process is complex and multi-faceted, involving collaborative aspects of constantly-changing verbal and non-verbal interactions between the instructors and learners (the instructor's directions induced by the learners, and the learners' utterances and attitudes influenced and changed by the instructor, etc.). In order to disentangle this kind of complexity and to consolidate their argument, the authors, employing the COLT (Communicative Orientation of Language

Teaching) Observation Scheme by Spada and Fröhlich (1995), analyzed and compared different series of language learning activities in three EFL classrooms.

### Keywords

language learning; sociocultural perspective; task sequence; task elaboration/development; media; English language activities for preschool children; classroom; content-oriented; outside world; Communicative Orientation of Language Teaching (COLT) Observation Scheme

### Introduction

For the past 20 years, task-based language teaching (TBLT) has attracted the attention of second language acquisition (SLA) researchers, curriculum developers, educationalists, teacher trainers and language teachers, worldwide (Van den Branden, 2006, p.1), and task-based language teaching/learning is commonly observed in EFL classrooms for children. It can be claimed that one of the possible roles that TBLT may play is to bridge classroom learning and the outside world. The importance of such a sociocultural perspective on language learning tasks is supported by a number of scholars. For example, Skehan (1998, p.95) claims that "a task is an activity in which there is some sort of relationship to comparable real-world activities." The same kind of view is presented by Samuda and Bygate (2008, p.18), who hold that language learning tasks should embrace "the role of personal relevance and purposeful activity in mediating the world of the learner and classroom learning."

Although it is uncertain how much this kind of sociocultural viewpoint is reflected and utilized in actual EFL classrooms for children in

Japan, it can certainly be claimed that relatively little attention has been directed at how language learning tasks are elaborated and developed as means to bridge classroom learning and the outside world, and that there have been few studies which empirically and extensively explore aspects of such elaboration and development. Sugino et al. (2008) is a preliminary study conducted in order to improve this research situation. First, employing the COLT framework, the researchers carried out observations and analyses of three Japanese EFL classrooms for children and quantitatively disclosed a number of distinctive features of each class, in which it is shown, for example, whether the class is content-oriented or form-oriented, and whether the topic is related to the children's immediate environment and experiences.

These classrooms were then examined qualitatively from the perspective of task selection, task use, task sequence, task elaboration and task development, by paying special attention to collaborative aspects of constantly-changing verbal and non-verbal interactions between the instructors and learners. A number of instructive points were revealed with regard to the use of language learning tasks in the classroom. The following summarize some of its main findings (Sugino et al., 2008, p. 560):

- 1) The choice of topic/content plays a crucial role in making a task fulfill the mediating function, and in order for a task to function as a means to mediate between the classroom and the world beyond, its topic/content must be chosen among the children's experiences in their immediate environments.
- 2) When the topic/content is the linguistic form(s) in activities, the task falls into the trap of self-circulation and loses its link with the real world experience.
- 3) For any linguistic form to be consolidated in a learner's knowledge, "repetition" is of utmost significance, and the target linguistic feature needs to be repeated in varied contexts with different participants.

These findings, based upon a close description and comparison of the three classes,

may offer valuable suggestions to instructors in EFL classes for children, but it must be noted that this study represents limited aspects of language learning tasks. It is unclear, for example, how "varied contexts" for "repetition" as in 3) are elaborated and developed in tasks based upon "the children's immediate environments and experiences" as in 1). There is a greater variety of EFL classrooms for children, which may involve more complicated and multi-faceted elements of how language learning tasks are elaborated and developed as means to bridge classroom learning and the outside world. Further studies are necessary to understand the nature of such elaboration and development more closely and comprehensively.

## **1. The Present Study**

### **1.1 Aim**

The present study, which employs the basic analysis framework of Sugino et al. (2008), aims to conduct observations and analyses of different Japanese EFL classrooms for children, focusing mainly upon relationships between the task and its constituent activities and interrelationships among them, and to depict in detail how language learning tasks are elaborated and developed serving as means to interface the classroom with the world beyond.

### **1.2 Data Collection**

Three EFL classes for children with different educational backgrounds were videotaped. One of the classes (Class A) was at a private kindergarten, which had incorporated bilingual/bicultural education since 2001. This class, composed of about 25 children (aged between 5 and 6), was led by a native speaker of English as its main instructor and a Japanese-speaking instructor as her assistant. A day at the kindergarten started from 10 a.m. and finished at 2 p.m., during which the children spent most of their instructional time mainly in English. The videotape recording was carried out from June 2001 to December 2003 on a weekly basis. Another class (Class B), with six first graders (aged between 6 and 7), was taught by a Japanese-speaking instructor. A lesson, of about one hour, was provided weekly, which has been videotaped since 2003. The third class (Class C) was at a nursery school, where college students carried out their practice teaching as part of their college's English language learning

program. They taught English to twenty children (aged between 5 and 6) for about forty minutes once a week for three months. Nine lessons by these teacher-trainees were videotaped in total.

### 1.3 Target Lessons

The accumulated videotaped data were scrutinized, and one lesson from each class was selected for this study, 1) by counterbalancing the children's age ranges among the three lessons, and 2) by examining and verifying the inclusion of one single whole language learning task covering most of the lesson and consisting of several different activities as its constituents. Table 1 summarizes some of the basic description of the three lessons:

Table 1: Basic Description of the Three Lessons

	Class A	Class B	Class C
Type of School	Kinder-garden	Private English School	Nursery School
Number	25	6	20
Age Range	5-6	6-7	5-6
Date	25/06/2003	01/07/2004	19/12/2007

### 1.4 Description and Analysis

The selected lessons were described utilizing the basic concepts of the COLT observation framework (Spada and Fröhlich, 1995), which consists of two parts: Part A for describing "classroom events at the level of episode and activity" (p. 13); and Part B for analyzing "the communicative features of verbal exchanges between teachers and students and/or students and students as they occur within each episode or activity" (p. 13).

As a first step, a general description and analysis of each lesson was carried out employing four basic analytic features of COLT Part A with slight modifications: 1) TIME, which examined durations spent on the various COLT analytic features, 2) RANGE OF REFERENCE, which determined types of activities by three categories, such as BROAD referring to topics going well beyond the classroom and immediate environment, 3) FORM, which dealt with grammar, vocabulary, pronunciation, etc., and 4) FUNCTION, which checked functions and communicative acts, such

as "apologizing". Each lesson was investigated in these aspects: 1) what kind of task was used in each class, including its aim and duration, 2) how many activities the task was composed of, and 3) what range of reference each activity bore, including its duration.

Second, the authors, employing the basic framework of COLT Part B, conducted a more language-based description of the three lessons. Verbal and non-verbal interactions between the instructors and the children were minutely examined using some of its basic analytic features, such as GENUINE REQUESTS FOR INFORMATION, which examined if the information requested was not known in advance by the questioner (e.g., a teacher-to-student question: *Where did your family go last Sunday?*).

Lastly, the authors, based upon all these depictions, characterized each of the three lessons, especially by investigating relationships between the task and its constituent activities and interrelationships among them, and made a number of comparisons of the three lessons.

## 2. Main Results and Discussion

A number of findings have been made by conducting description, characterization, and comparison of the three EFL classes for children. Due to space limitations, some of the major results based upon an analysis of COLT Part A, are reported in this paper.

### 2.1 The Task and its Constituent Activities

In each of the targeted three lessons, a language learning task, composed of its several constituent activities, was used. First, the topic/content of such a task, including its aim, was checked. Class A dealt with a shopping game, in which the children were taught how to do shopping in English. In Class B, a task focusing upon English language itself was used, where the children were taught how to write capital letters from A to Z, for instance. Class C used a task dealing with Christmas, in which the children were exposed to English in a Christmas setting. Next, the duration of the tasks, the number of the constituent activities, and their mean duration were investigated. Table 2 provides a basic description of the task and its constituent activities in each of the three lessons:

Table 2: Basic Descriptive Data of the Task and its Constituent Activities

	Class A	Class B	Class C
Topic/Content	shop-ping game	English sound & letters	Christ-mas & Santa Claus
Duration of the Task (min.)	34.0	52.6	43.2
Number of Activities of the Task	10	3	3
Mean Duration of the Activities (min.)	3.4	17.5	14.4

One point in this table deserves much attention: Class A is composed of far more activities than the other two despite the fact that its duration is the shortest and that each activity lasts for only a short time. Such a task could be interpreted in various ways, such as in terms of the children's English proficiency, but it may be the most reasonable to assume that the constituent activities of the task are elaborated and developed to provide various contexts for the children to learn certain aspects of the target language repeatedly.

## 2.2 Constituent Activities of the Task

Such elaboration and development of the constituent activities of the task are considered to be complex and multi-faceted. The data from the three lessons have been processed in the following four ways:

- 1) a brief description of the content of each activity
- 2) whether the topic of the whole task is more or less referred to or dealt with in each activity
- 3) whether language form-related learning/teaching is performed in relation to the topic of the whole task
- 4) whether language function-related learning/teaching is performed in relation to the topic of the whole task

Table 3 provides a description of Class A:

Table 3: Topic/Form/Function in Activities in Class A

	Topic	Form	Function
A1: getting started	○	○	○
A2: birthday interview	○		
A3: talking about some visitors	○		○
A4: shopping game: role play	○		
A5: how to count, paper money	○		
A6: names of animals to buy	○	○	
A7: expressions for shopping	○	○	○
A8: coloring the animals they bought	○	○	
A9: shopping bag and naming	○		
A10: demonstra-tion of shopping	○	○	○

This table shows in general that although the topic of the whole task is more or less referred to or dealt with in every activity, language form-related and language function-related learning/teaching are not continuously performed in relation to the topic of the whole task. A close look at the table indicates, however, that there are three activities, in which reference of the topic of the whole task and performance of language form-related and language function-related learning/teaching were observed at the same time: in Activity 1, Activity 7, and Activity 10. This view of the activities of

the task in Class A seems to represent more of the nature of task-based language learning, and it may be possible to claim that although topical consistency and reference must be fully maintained in task-based language learning, language form-related and language function-related learning/teaching do not have to be continuously carried out or should be conducted when necessary.

Next, let us look at Table 4, which provides a description of Class B:

Table 4: Topic/Form/Function in Activities in Class B

	Topic	Form	Function
A1: writing capital letters: A to Z	○		
A2: pronouncing the letters of the alphabets	○		
A3: learning words	○		

This table shows that although the topic of the whole task is more or less referred to or dealt with in every activity, as in Class A, language form-related and language function-related learning/teaching are not performed at all in relation to the topic of the whole task. Basically, it must be admitted that a non-language-based task is certainly used as in Class B, but it is quite conceivable, considering the nature of Activity 2, that some kind of language form-related learning/teaching would have been carried out in the activity. An analysis based upon COLT Part B may be able to examine it more closely.

Finally, let us look at Table 5, which provides a description of Class C:

Table 5: Topic/Form/Function in Activities in Class C

	Topic	Form	Function
A1: story time: Santa Claus	○	○	○
A2: decorating: Christmas tree	○	○	○
A3: Christmas Present	○	○	○

This table shows that the topic of the whole task is more or less referred to or dealt with in every activity as in the other two classes, and that language form-related and language function-related learning/teaching are performed in every activity in relation to the topic of the whole task. It may be possible to say that the whole task was well elaborated and developed continuously throughout its constituent activities, although closer examination based upon COLT Part B is necessary to be conclusive.

A comparison in data among these three tables reveals that the task in each lesson may vary a great deal although each of its constituent activities is more or less related to the task topic. This variety is considered to be largely derived from its selection which determines whether the task topic is based upon children's experiences in their immediate environments. For example, Class A, dealing with "a shopping game," offers a number of scenes that involve more or less personal relevance with the children (Normally, all children have done some shopping in their immediate environments.), in which they are encouraged to learn certain aspects of the target language, such the language function of ASKING (the price) and to use them in some quasi situations (using toy money).

### 3. Concluding Remarks

The authors have analyzed and compared task-based language learning in three EFL classrooms for children. In this paper, some of the major results based upon an analysis by COLT Part A, have been reported. In the presentation, employing the results from an analysis by COLT Part B, they will make an attempt to gain a better understanding of how language learning tasks are elaborated and developed, serving as means to interface the classroom with the world beyond.

### Acknowledgment

The present study was supported by Grant-in-Aid for Scientific Research (B) (19320092) from the Japan Society for the Promotion of Science.

### References

Samuda, V. & Bygate, M. (2008). *Tasks in second language learning*. NY: Palgrave Macmillan.

- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford, UK: Oxford University Press.
- Spada, N., & Fröhlich, M. (1995). *Communicative orientation of language teaching observation scheme: Coding conventions and applications*. Sydney, Australia: NCELTR, Macquarie University.
- Sugino, N., Kawashima, H. & Koga, Y. (2008). Tasks as means to mediate between the classroom and the world beyond: qualitative comparison of language activities in three EFL classrooms for children. *Proceedings of CLaSIC 2008*, 554-561.
- Van den Branden, K. (2006). Introduction: Task-based language teaching in a nutshell. In K. Van den Branden (ed.), *Task-Based Language Education: From Theory to Practice*, 1-16. Cambridge University Press.

# A Case Study of the Development of Communication Skills through Cross-Cultural Distance Learning

Akiko Watanabe and Annette Karseras Sumi

Graduate School of Education, Waseda University and Waseda University International Cooperation

aki49@fuji.waseda.jp, asumi@kurenai.waseda.jp

## Abstract

This study investigates advantages of Cross-Cultural Distance Learning (CCDL) for acquiring skills for social and global communication. The results of the study demonstrate a significant increase in reported use of both expressive and receptive communication skills. Both these skills were taught explicitly throughout the course as a part of a broader set of 'facilitation' skills and in conjunction with nonverbal communication skills. The study focuses on discussion between university students based in Japan and Taiwan in which all the students used English as a foreign language (EFL). Small group discussions were held using internet video-conferencing. All Japan-based students from one CCDL class completed a weekly self-evaluation questionnaire during fall 2008. The questionnaire was distributed to students at the end of each class and collected on completion. Four items on communication skills (3 expressive and 1 receptive) were assessed using a five-point scale. Students were asked to indicate the extent to which they had used each skill. The opportunities to practice specific communication skills through a variety of structured activities, prior to using them in discussion situations with overseas partners, seems to have played an important role in the development of students' communication skills.

## Keywords

Global Communication skills, Cross-Cultural Distance learning, Social and Emotional learning (SEL)

## 1 Introduction

### 1.1 Cross-Cultural Distance Learning

Cross-Cultural Distance Learning (CCDL) is "a joint seminar between two universities in Asia using an Internet video-conferencing system" (Nakano, Yoshida, and Owada, 2008: 191). At present there is a series of three CCDL text books, each with a different theme: Social and Global Issues, International Career Path, and Media. Some CCDL communication skills, such as facilitation skills, are

common across more than textbook. Others, such as nonverbal communication skills, and high or low context communication styles are specific mainly to one course.

While not taught explicitly to students, Goleman's (2006) concept of 'social competence' appears in some detail in teacher's notes accompanying the core textbook (Nakano et al, 2008: iii). The teacher uses this broadbased understanding herself, to inform her own instructions with the students when playing the role of 'teacher-facilitator' during activities from the student textbook. All ten facilitation skills listed in student text represent expressive communication skills (Table 2). Facilitation skills were originally formalized as a part of the CCDL course in order to help students cope with the low context (Hall, 1959), virtual environment of video-conferencing and to mitigate against 'white-noise' and 'time lag' endemic to this technology in such early stages of its development. Effective use of technology to communicate across geographies and to manage relationships despite distance is critically imported for citizens that wish to be active transnationally. Facilitation is most simply explained in terms of 'helping a conversation go smoothly' between people with different personalities and communication styles.

The course emphasises the development of social competence and communication skills as "transferable skills" that can be "useful adapted to a range of situations beyond the end of the course" (Nakano et al, 2008: vii). The students who take CCDL courses not only have the opportunity to develop these skills with others from their own country, but also with students from different cultural backgrounds across East Asia. CCDL aims to help students notice how different communication skills, such as expressive and receptive skills, are used to different extents by people with different cultural backgrounds and personalities. In this sense, the communication skills developed through CCDL are considered useful for participation in global, regional, and local society.

As shown in Table 1, CCDL consists of three

types of classes: 'Preparation', 'Joint', 'Post Joint' classes. All class-types give students discussion opportunities.

Table 1: Media Course

Media Schedule Overview 2008				
CLASS	Preparation/Joint	TOPIC	SKILLS	Homework
WEEK 1	Preparation	Introduction	Facilitation	None
WEEK 2				HW1: Reading facilitation skills & B self-introduction
WEEK 3				HW2: Reading & video conversation management & Dr. Nakano's lecture
WEEK 4	Preparation	Advertising	Nonverbal	HW3: Research Car Advert worksh
WEEK 5	Joint Class 2		Communication	HW4: Reading
WEEK 6	Preparation	Front Page	Context	HW5: Reflection Paper Topic 1&2
WEEK 7	Joint Class 3			HW6: Research Front page broashe news
WEEK 8	Preparation	Country Image	D.I.E. <sup>1</sup>	HW7: Reading D.I.E. & O.D.I.S.
WEEK 9	Joint Class 4			HW8: Research
WEEK 10	Preparation	Popular Culture	O.D.I.S. <sup>2</sup>	HW9: Reflection Paper Topic 3&4
WEEK 11	Joint Class 5			HW10: Reading
WEEK 12	Post Joint Class	Presentations	Multiple Intelligence	HW11: Power Point
WEEK 13				HW12: Rehearsal
WEEK 14				
WEEK 15				

Source from CCDL Teacher's Notes (Nakano et al, 2008:1).

1. D.I.E. stands for description, interpretations, and evaluation.
2. O.D.I.S. stands for observation, description, interpretation, suspending ethnocentric evaluations.

Preparation Classes, conducted in a face-to-face discussion environment, are structured around one-to-one, small group, and whole class activity depending on the communication skills being taught. This enables learning to take place on three levels. Students learn about facilitator's role by facilitating a discussion themselves, by experiencing being a participant in a facilitated discussion, and by observing others facilitate (ibid: 20). First, students have the chance to learn experientially through playing an active, expressive role, such as being a facilitator. For example, at the start of each small group discussion, students choose one person to be a facilitator for that activity. A different student takes this role each time to gain practical experience using one or more of ten facilitation skills as follows (Nakano et al, 2008: 20).

Table 2: Facilitation Skills

	Facilitation Skills
1	Helping things go smoothly; stopping things from getting stuck, keeping the conversation going.
2	Ensuring everyone is participating and can understand the discussion
3	Writing main discussion questions and keywords in the TextBox to make sure everyone basically follows the discussion.
4	Inviting a student to speak up or ask questions if they have been quiet for some time.

5	Preventing one person from dominating the discussion.
6	Deciding who should go first if 2 or more people want to speak at the same time.
7	Directing discussion by calling on individual people by name.
8	Preparing extra questions and/or topic-related material before discussion to deepen and/or extend the discussion.
9	Asking 'why' questions during the discussion to expand understanding of reasons and avoid making assumptions.
10	Enabling group members to make personal and topic-related connections at a deeper level.

(Source from Nakano et al, 2008: 20)

This is the first level in which students 'learn by doing'. Simultaneously, other students in the group, learn about facilitation through their experience of being a participant in a discussion facilitated by another student. They 'learn by participating'. Students experience first hand what it is like to be invited to share their opinion, or to be asked "Why?" in order to clarify their reasons or check assumptions. During this second experiential level of learning, students can gain both positive and negative examples as the novice facilitators climb the ladder from unconscious incompetence to competence. The third level of learning occurs during whole-class discussion. This 'debrief activity' is facilitated by the teacher. She elicits content-level learning from each small group in order to elaborate the definition of skill used in each topic, either explicitly with a definition or more tacitly through examples. As one member from each small group shares something from their discussion, the teacher is able to use many of the facilitation skills herself as a model for students to learn from.

During Preparation Classes, the length of group discussion time ranges from 1 to 15 minutes, depending on the topics, activities, and purposes. Students are given the chance to work in different groups and with different individuals in order to maximize their exposure to different personalities and communication styles (at the process level), and their exposure to different ideas and points of view (at the content level). Implicitly, this opportunity creates the need for students to engage with and adapt their communication style to a greater range of people than might occur if working consistently with students of their own choice.

In Joint Classes, by contrast, there is continuity amongst 'ChatRoom' members each week (except for student absences). Joint Classes consist of 50-minute discussion periods with overseas students, in small ChatRoom groups of 2 to 6 students at least one student from each country's university. The remaining 40 minutes of the regular 90-minute period is used for preparation and debrief. Avoiding the need for repeated introductions to new ChatRoom members each week and enabling rapport to develop with overseas

students is considered an important for the quality and depth of discussion. The two countries involved take turns to facilitate. For example, a student from Japan facilitates one week, followed by a facilitator from Taiwan the next.

## 1.2 CCDL Media Course

The present study focuses on learning from the Media course which includes both facilitation skills (expressive) and nonverbal communication skills (expressive and receptive). As the Media schedule shows (Table 1), Joint Classes are held on alternate weeks during week 3, 5, 7, 9, and 11. The partner country in the case of the course observed was Taiwan.

## 1.3 Research Objectives

What communication skill do students gain from CCDL?

# 2 Method

## 2.1 Applying Social Emotional Learning to University Environment

Items on one receptive communication skill and three expressive communication skills were adopted in the questionnaire. Items on these communication skills were excerpted from Social Emotional Learning (SEL) category (Appendix A) and translated into Japanese. SEL has been mainly used in United States, Denmark, Sweden, Israel, England and it has around eighty programs to solve current social issues. It aims for each student to be a successful person in a society. SEL was defined by the Collaborative for Academic, Social, and Emotional Learning (CASEL) as “the process of acquiring and effectively applying the knowledge, attitudes, and skills necessary to recognize and manage emotions; developing caring and concern for others, making responsible decisions; establishing positive relationships; and handling challenging situation capably” (Zins and Elias, 2004: 1). Such processes of acquiring and applying the knowledge, attitudes, and skills refer to the main learning point in CCDL.

Take one of the core skills sets in CCDL, facilitation skills, as example, students learn about what facilitation means by observing a facilitator in a group, experiencing being as a participant in the discussion facilitated by another student, as well as facilitating a discussion themselves. In the concept of SEL, such facilitation in CCDL can be explained as “making responsible decisions; establishing positive relationships; and handling challenging situation capably” (ibid: 1). Therefore it is possible to say that SEL concept covers some of CCDL skills.

SEL has been widely used in primary school or high school, however, Shimai et al (2007) argues for the value of applying SEL in a university environment. According to Shimai et al’s study, the university students group who had received employment offers from more than three companies showed highest scores in intrapersonal, interpersonal, and situational abilities of emotional intelligence (Shimai et al, 2007). This result shows SEL use in university students is effective in their future life. Such concept of using skills is also similar to the concept of CCDL skill use in a real life domain.

Two types of communication skills, ‘receptive communication’ and ‘expressive communication’ were excerpted from the SEL categories. These definitions were revised as question type for the questionnaire (Appendix A).

### Communication Skills

1. *Receptive communication*: The capacity to attend to others both verbally and nonverbally to receive messages accurately
2. *Expressive Communication*: The capacity to initiate and maintain conversation, express one’s thoughts and feelings clearly both verbally and nonverbally” (Patricia et al, 2000).

From *Receptive communication*, “The capacity to attend to others both verbally and nonverbally to receive messages accurately” was revised as Survey Question Two. From *Expressive communication*, “The capacity to initiate and maintain conversation” was revised as Survey Question One, “express one’s thoughts and feelings clearly both verbally and nonverbally” as Survey Question Three. Survey question Four was “Did you check if the listener understood you”. Each item on communication skills in the questionnaire is shown in Table 2.

Table 3: Receptive/Expressive Communication Skills

Receptive/Expressive Communication Skills	
1.	Did you initiate and maintain conversation?
2.	Did you attend to the other to receive messages accurately?
3.	Did you express your thoughts and feelings clearly?
4.	Did you check if the listener understood you?

\*Q1, Q3, and Q4 are Expressive Communication

\*Q2 is Receptive Communication

The questionnaire has a five-point scale (1. not at all, 2. not much, 3. somewhat, 4. a little, 5. a lot). (Appendix B).

## 2.2 Participants

Participants comprised ten university students, One Chinese and nine Japanese, who took CCDL Media course, fall semester in 2008. Although the majority were not EFL majors, most had an interest in either 'maintaining' or 'improving' their English skills. Other reasons given for taking CCDL include 'interest in other Asian cultures/countries' and 'to communicating with students in Asia'. As CCDL is not a compulsory course, participation in the course was entirely for voluntary. Each student had a complement of different study purposes. All students except one had experiences of study abroad from one month to a year. Departments of the students varied such as Education, Humanities and Social Sciences, International Liberal Studies, Law, Letters, Arts and Social Sciences, Political Science and Economics, Social Science, and Sport Sciences. The grade of the students varies from sophomore to senior.

## 2.3 Data processing

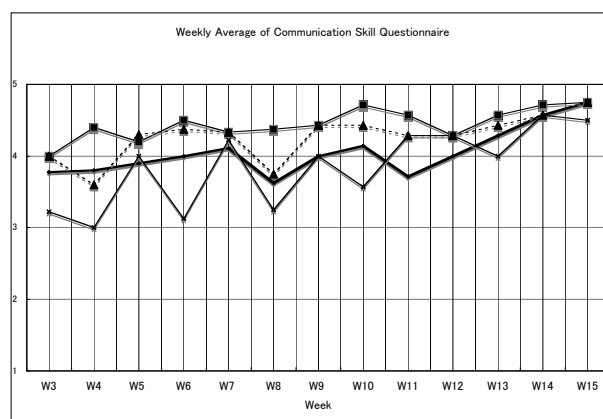
This questionnaire was conducted weekly after the class and collected following completion on the same day. The sum of the individual rate in each item was divided by the numbers of the participants in the class. Mean (M) and Standard Deviation (SD) of the questionnaire were calculated.

## 3 Results and Discussion

This report discusses two main findings of the research. Other results are discussed elsewhere (Watanabe, 2008). (1) Overall, there was both an increase and convergence in students' reported use of all skills towards the end of the course. (2) Context-sensitive use of skills was observed in respect to Joint and Preparation classes.

### 3.1 Increase and convergence in reported use of skills

As shown in Figure 1, the average rate of the students' self-evaluations increased over the semester. This is true for each of the 4 communication skills. Students reported using the skills less towards the beginning of the course and more towards the end.



From top to bottom line,

■ (Q2) Have you attended to other to receive messages accurately?

▲ (Q3) Have you expressed your thoughts and feelings clearly?

◆ (Q1) Have you initiated and maintained conversation?

× (Q4) Did you ask listeners if they understood you?

Figure1: Students' Weekly Self-Evaluation of Communication Processes

Along with increasing use of skills, students' perception of the skills also seems to have become more similar as the course progressed. There was a dramatic decrease in the SD for each item. This convergence is especially apparent for Q4 which asks whether students confirmed understanding on the part of the listener. SD for this item in week 3 was 1.48 (9 participants) and in week 14 was 0.53 (7 participants). Why did this alignment in ratings occur? One interpretation is that students' individual goals aligned as a result of discussion over the course. The range of study purposes students brought to this elective course became more focussed on acquiring communication skills *per se* as their understanding of different types of communication skills increased and as their ability to use them improved. Opportunities for structured practice on three levels may have played a part in this. Through observing others, experiencing being a participant in the discussion facilitated by another student, and facilitating a discussion themselves, the students could focus more on a similar goal or similar direction. Another contributing factor may have been the continuity of ChatRoom members in Joint Lessons. Increasing familiarity with group members may have provided students with the confidence to communicate in a more active way. This continuity may also have provided them with a baseline by which to compare their participation each week. However, as there was no control group in this research, the existence of experimenter effects cannot be ruled out. It is possible that the presence of the researcher and the simple act of completing the questionnaire each week reinforced the centrality of the skills in students' minds. Further research is needed to investigate the causes

of this convergence.

Table 4: Mean and Standard Deviation of the Self Evaluation Questionnaire

WEEK	Q 1		Q2		Q3		Q4		N
	M	SD	M	SD	M	SD	M	SD	
<b>W3</b>	3.78	(1.20)	4.00	(1.00)	4.00	(1.22)	3.22	(1.48)	9
W4	3.80	(0.84)	4.40	(0.89)	3.60	(1.14)	3.00	(1.22)	5
<b>W5</b>	3.90	(0.88)	4.20	(0.79)	4.30	(0.67)	4.00	(0.94)	10
W6	4.00	(0.53)	4.50	(0.53)	4.38	(0.92)	3.13	(1.13)	8
<b>W7</b>	4.11	(0.78)	4.33	(0.71)	4.33	(0.71)	4.22	(0.67)	9
W8	3.63	(0.92)	4.38	(1.06)	3.75	(1.04)	3.25	(1.39)	8
<b>W9</b>	4.00	(0.58)	4.43	(0.53)	4.43	(0.79)	4.00	(1.00)	7
W10	4.14	(0.38)	4.71	(0.49)	4.43	(0.79)	3.57	(1.51)	7
<b>W11</b>	3.71	(0.76)	4.57	(0.53)	4.29	(0.76)	4.29	(0.76)	7
W12	4.00	(0.82)	4.29	(0.76)	4.29	(0.76)	4.29	(0.76)	7
W13	4.29	(0.76)	4.57	(0.53)	4.43	(0.53)	4.00	(0.82)	7
W14	4.57	(0.53)	4.71	(0.49)	4.57	(0.53)	4.57	(0.53)	7
W15	4.75	(0.50)	4.75	(0.50)	4.75	(0.50)	4.50	(0.58)	4

\*Q1, Q3, and Q4 are Expressive Communication

\*Q2 is Receptive Communication

\* N: the number of the students attended in class

\*Joint Discussion was held in week 3, 5, 7, 9, 11.

(Q1) Have you initiated and maintained conversation?

(Q2) Have you attended to other to receive messages accurately?

(Q3) Have you expressed your thoughts and feelings clearly?

(Q4) Did you ask listeners if they understood you?

### 3.2 Context-sensitive use of skills

The second finding discussed here regards context-sensitive use of skills. The students used receptive and expressive communication skills differently in Joint Class and Preparation Class. When using English in Japan, these learners appeared use a more receptive communication style. The result of the questionnaire demonstrates item of Q2 (receptive communication item) is the highest average among all. This tendency could be seen clearly in Preparation Class on week 4, 6, and 10. This shows that students used receptive communication skill more when they communicated with people receptive communication skills more often when they communicated with people who shared the same cultural background. This also applies to some extent in the case of the Chinese student studying in Japan, as she had extended experience in the Japanese context and with its communicative norms. In traditional Japanese social context, more emphasis is placed on listening and interpreting the message rather than direct expression (Ting-Toomey and Chung, 2005). In particular, receptive skills, such as noticing people's body language reaction, or lack of it, as part of listening and understanding the other person's point of view and feeling, plays an important role. This high Japanese type of listening holds true also in the

academic environment. This difference was most striking in relation to Q4. The average rate of Q4 was explicitly high in Joint Class compared to the rate in Preparation Class. In Table 3, the average rate of their evaluation on Q4 was 3.22, 4.00, 4.22, 4.00, 4.29 in week 3, 5, 7, 9, and 11 respectively, while the one in Preparation Class was 3.00, 3.13, 3.25, 3.57 in week 4, 6, 8, and 10. Compared these mean difference between Joint Class and Preparation Class, it can be said that the way the students use communication skill will be different depending on class types. Furthermore, the result shows that there were high average rate in Q2 and low average rate in Q4 in Preparation Class. This demonstrates that the students were more receptive in Japanese social context.

## 4 Conclusion

The results of self-evaluation questionnaire demonstrate increase and convergence in reported use of skill and context-sensitive use of skills. Students' high average rate of listening to others can be said to reflect Japanese tendency to use a more receptive communication style according to cultural norms. This was particularly evident during Preparation Classes. However, the students' increasing reported using a more expressive communication style both in the practice situation of the Preparation Classes and during Joint Class discussion with students in Taiwan. From the results of decreasing SD in the average rate on each communication skill, it seemed that students' perceptions on skills became more similar as the course progress. An increase average rate of skills use with decreasing SD shows students' similar evaluation of themselves. This could be that the students shared a common goal with partners involved in a discussion or even with wider cross-cultural norms.

### Acknowledgement

I would like to express my sincerest thanks to Prof. Michiko, Nakano introducing me to CCDL courses. I also appreciate Karseras, Sumi Annette who allowed me to observe her class and to collect questionnaire from the students over the course after their class hours.

### References

- Goleman, D. (2006) Social and Intelligence. New York: Bantam Books.
- Graczyk, P. A. et al. (2000). Criteria for evaluating the quality of school-based social and emotional learning programs. In Bar-On, Reuven & James, D. A. Parker (eds.), *The handbook of emotional intelligence: theory, development, assessment, and application at*

- home, school, and in the workplace (pp. 391-410). San Francisco, CA: Jossey-Bass.
- Hall, E.T. (1959) the silent language. New York: Doubleday.
- Nakano, M., Karseras Sumi, A., and Yokota, M. (2008). *Cross cultural distance learning media: Teacher's notes*, Waseda University International Corporation.
- Nakano, M. et al. (2008). Developing transferable skills and social intelligence through theme-based cross cultural distance learning 2008 spring semester. In JACET select committee (eds.), *Information Communication Technology Practice & Research 2007* (pp.191-199). Tokyo: Waseda University International Corporation,.
- Nakano, M., Yoshida, S., and Owada, K. (2008, November). *Cross-cultural distance learning (CCDL) programs and program assessments*. Paper presented at annual meeting of the Association of Pacific Rim Universities 9<sup>th</sup> distance learning and the Internet conference 2008, Tokyo, Japan, 189-198.
- Shimai, S. et al. (2007). Jyoshidaigakusei ni okeru jyoudouchinou no yakuwari. *Japan journal of Administrative Science*. (20), no.3, 317-324.
- Ting-Toomey, S. and Chung, L. C. (2005) *Understanding Intercultural Communication*. CA: Roxbury.
- Watanabe, A. (2008). *CCDL Media classroom observation notes in 2008*.
- Zins, J. E. and Elias, M. J. (2006). Social and emotional learning. In G. G. Bear, & K. M. Minke (eds.), *Children's needs III: Development, prevention, and interaction* (pp.1-13). Bethesda, MD: National Association of School Psychologists.

## Appendices

### Appendix A

#### Items on communication skills in the Questionnaire excerpted from the SEL category.

Social and Emotional Learning (SEL): Key SEL Competencies by Category

##### Communication Skills

3. *Receptive communication*: The capacity to attend to others both verbally and nonverbally to receive messages accurately
4. *Expressive Communication*: The capacity to initiate and maintain conversation, express one's thoughts and feelings clearly both verbally and nonverbally, and demonstrate to other speakers that they have been understood (Patricia et al, 2000: 400)

Source from *The handbook of emotional intelligence*.

### Appendix B

#### Items on receptive and expressive communication skills excerpted from Questionnaire for this study.

Please give your frank opinion about today's lesson. Choose the number, which is suitable for your feeling about the lesson.

1=completely not 2=not quite 3=so-so 4=little 5=very much

(1)Have you initiated and maintained conversations?(自分から会話を始め続けることができましたか?)

1-----2-----3-----4-----5

(2)Have you attended to other (both verbally and nonverbally) to receive messages accurately?

相手の意味することを正しく受け取るために(言葉であるいは言葉を使わないで)注意して話を聞きましたか?

1-----2-----3-----4-----5

(3) Have you expressed your thoughts and feelings clearly (both verbally and nonverbally)?

自分の思ったことや感じたことを、言葉で、あるいは言葉を使わないで、はっきりと表現しましたか?

1-----2-----3-----4-----5

(4) Did you ask listeners if they understood you?

1-----2-----3-----4-----5

あなたは相手に対してあなたの話を理解したかどうか尋ねましたか?

\*Some of the items in the questionnaire were excluded as they were not used in the study.

# Vocabulary Frequency in Japanese Entrance Examinations from 2005 to 2008

Takane YAMAGUCHI

Waseda University

takane@aoni.waseda.jp

## Abstract

This paper describes on written English vocabulary frequency in Japanese University entrance examination corpus in 2005, 2006, 2007, and 2008 respectively, in terms of how different in word frequency ranks each word family in each word list is. The difference of the ranks seems to be necessary when attempts have been made to make educational word lists, because some educational word lists happen not to contain words seemingly necessary for learners. The research question is on how many tokens are necessary for making word lists for learners of English.

## Keywords

Entrance examinations, Vocabulary, Word frequency, Word list

## 1 Procedure for word lists

### 1.1 Whole the procedure for word lists

All English words in each examination corpus were picked up, decapitalized, lemmatized, and counted by frequency. The procedure is the same as that reported in Yamaguchi (2006).

#### 1.1.1 Decapitalization

Because of the decapitalization, the word lists in this study cannot tell originally capitalized words from originally non-capitalized words. For example, word family ‘i’ contains the word ‘I’ as well as the word ‘i’.

#### 1.1.2 Lemmatization

Lemmatization is the change of some word tokens into the same word family or lemma. The table for lemmatization was originally made by using data produced by Yasumasa Someya.

## 2 Results

Table 1 shows all the word lists in this study.

Table 1: Word lists in this study

	2005	2006	2007	2008
Token number	803096	177029	74422	83217
Lemma number	20733	9817	6116	5086

### 2.1 The size of tokens necessary

Table 2 is an example of ranks of the words randomly selected by the author.

Table 2: ranks of the words in each corpus

	2005	2006	2007	2008
Admire	1798	-	-	-
Pretend	2230	2746	-	708
explain	788	777	774	784
effective	1038	2048	3384	784
Suppose	11325	1101	1527	826
government	379	530	221	927
worker	514	1836	1889	991

While the word “suppose” appears at the 11325<sup>th</sup> in 2005 word list, it comes up at the 826<sup>th</sup> in 2008, which shows the possibility that corpora containing 803096 tokens is not ideally correct as well as the one that 2008 corpus, which has fewer tokens than 2005, cannot state ideal word orders.

The word “pretend” is not found in 2007 word list. Because the author believes this word is necessary for learners, 2007 corpus alone, which token size is 74422, seems to be inappropriate for making word lists.

Similarly, the word “admire” is not found except 2005 word list. Given that the reason for its absence from the other word lists is the size of tokens, it is assumed that more than 177029 tokens are needed for word lists appropriate for learners.

Thought along the same line, it can be doubt if 2005 corpus is appropriate for educational word lists, which is to be examined in next research.

<b>Reference</b>		48	2351	time
Yamaguchi, T. (2006). Vocabulary Frequency in University Entrance Examination in Japan. <i>Proceedings of the 11th Conference of Pan-Pacific Association of Applied Linguistics.</i>		49	2323	were
		50	2315	she
		51	2277	all
		52	2266	e
		53	2257	there
		54	2231	say
		55	2220	has
		56	2191	so
		57	2162	if
		58	2098	will
<b>Appendix A. 2005 Word list.</b>		59	2076	her
Rank	Frequency Word	60	2063	which
1	44203 the	61	2048	use
2	28209 a	62	1981	than
3	24451 to	63	1962	who
4	21691 of	64	1894	your
5	17460 and	65	1867	year
6	16011 in	66	1779	take
7	11097 that	67	1772	would
8	10287 is	68	1766	no
9	9429 i	69	1728	other
10	8504 it	70	1712	because
11	8324 b	71	1703	me
12	7558 c	72	1660	how
13	7170 for	73	1644	think
14	7089 be	74	1632	up
15	7082 you	75	1618	them
16	6242 d	76	1617	like
17	5888 was	77	1602	many
18	5706 are	78	1587	some
19	5676 s	79	1575	work
20	5398 as	80	1529	out
21	4797 on	81	1468	our
22	4762 have	82	1432	see
23	4713 with	83	1418	child
24	4504 this	84	1411	most
25	4371 they	85	1356	new
26	4361 not	86	1338	world
27	4061 he	87	1324	very
28	3725 at	88	1299	way
29	3621 by	89	1274	only
30	3557 we	90	1247	could
31	3447 but	91	1214	f
32	3428 from	92	1209	come
33	3363 or	93	1198	into
34	3208 t	94	1183	day
35	3207 people	95	1181	first
36	3202 do	96	1174	much
37	3111 their	97	1170	become
38	2904 can	98	1118	make
39	2808 what	99	1097	get
40	2754 one	99	1097	good
41	2747 more	101	1056	want
42	2707 had	102	1054	just
43	2655 an	103	1035	know
44	2604 when	104	1020	don
45	2585 about			
46	2452 his			
47	2393 my			

105	1018	even	162	626	need
106	1016	life	163	624	back
106	1016	well	164	617	part
108	1006	may	165	614	going
109	992	now	166	612	try
110	990	after	167	610	japan
111	987	its	168	609	however
112	980	should	169	605	always
113	971	also	170	602	often
114	958	thing	171	601	own
115	950	such	172	594	must
116	947	two	173	593	through
117	926	over	174	591	does
117	926	country	174	591	old
119	921	go	176	585	better
120	916	him	176	585	help
120	916	us	178	578	book
122	902	then	179	577	every
123	888	change	179	577	person
124	880	m	181	572	between
125	871	tell	181	572	little
126	860	too	183	568	city
127	859	woman	184	567	still
128	855	word	184	567	water
129	849	study	184	567	state
130	847	american	187	566	question
131	827	english	187	566	seem
132	824	where	189	564	great
133	821	any	190	563	speak
134	818	right	191	560	really
135	807	call	192	553	less
136	790	school	193	551	animal
137	784	japanese	194	549	parent
138	782	made	195	545	money
139	781	problem	196	543	man
140	777	language	197	540	around
141	773	write	198	539	down
142	771	why	199	538	human
143	761	before	200	535	number
144	735	student	201	532	play
145	721	read	202	528	system
146	718	same	203	525	both
147	709	ask	203	525	find
148	672	different	203	525	talk
149	671	long	206	524	ll
150	667	place	207	519	best
151	663	each	208	515	small
152	662	important	209	514	without
153	660	never	210	513	today
153	660	something	211	508	believe
155	657	home	212	507	three
156	655	food	213	502	g
157	652	begin	213	502	answer
158	647	show	215	499	men
159	637	start	216	498	cause
160	631	another	217	494	since
161	627	while	218	492	few

219	490	last	276	366	let
220	487	family	277	364	author
221	484	re	277	364	course
222	481	mother	277	364	spend
223	480	develop	280	361	far
224	479	look	280	361	learn
225	469	here	280	361	society
226	465	might	283	359	group
227	464	found	283	359	percent
227	464	off	285	358	house
229	463	car	285	358	kind
230	462	didn	287	357	mr
231	456	put	288	356	information
232	452	high	288	356	until
233	449	young	290	355	week
234	443	age	291	354	area
235	440	accord	292	353	almost
236	439	yes	292	353	social
237	435	got	292	353	run
238	431	example	295	346	close
239	426	buy	296	345	stop
240	423	during	297	344	possible
241	422	give	298	342	unite
242	417	reason	299	339	science
243	414	left	300	338	ago
243	414	turn	300	338	america
245	413	next	300	338	father
246	409	university	300	338	big
247	406	computer	304	336	h
248	404	feel	305	335	means
249	403	away	306	332	together
249	403	century	306	332	point
249	403	understand	308	330	idea
249	403	class	308	330	order
253	401	story	310	329	friends
254	400	having	311	328	ever
255	398	lot	311	328	hard
256	397	large	311	328	produce
257	394	live	314	327	early
257	394	business	314	327	sure
259	391	am	316	325	set
260	390	others	316	325	watch
261	389	bring	318	323	cannot
262	386	culture	318	323	ve
263	385	enough	318	323	yet
264	382	scientist	318	323	teacher
264	382	body	322	322	control
264	382	result	323	318	power
267	381	following	324	317	history
268	380	able	324	317	night
269	378	fact	326	315	choose
270	376	once	327	313	later
271	374	went	327	313	someone
272	371	living	327	313	difference
272	371	company	327	313	create
274	369	though	331	312	keep
275	368	train	331	312	making

333	311	soon	389	271	die
333	311	technology	391	270	doing
335	310	second	391	270	end
335	310	test	393	269	sense
337	308	eat	393	269	rule
337	308	job	395	268	matter
337	308	consider	396	267	everyone
340	307	nothing	396	267	plan
341	306	hand	398	266	tree
342	304	again	399	265	morning
343	303	health	399	265	process
344	301	mind	399	265	program
345	298	product	399	265	wait
346	297	difficult	403	263	among
346	297	earth	403	263	view
346	297	room	405	262	four
346	297	walk	405	262	pay
350	296	usually	405	262	plant
351	295	although	408	261	common
351	295	passage	408	261	special
351	295	rather	408	261	cost
351	295	win	411	260	brain
351	295	provide	412	259	education
356	294	case	412	259	several
357	293	build	412	259	report
358	292	against	415	258	makes
358	292	war	415	258	species
360	291	five	417	257	population
360	291	late	417	257	situation
360	291	lives	417	257	visit
363	290	public	420	254	break
363	290	phone	421	252	done
363	290	teach	422	250	bad
366	289	stay	422	250	longer
367	288	research	422	250	sometimes
367	288	whether	422	250	college
369	286	decide	426	249	form
369	286	thank	426	249	mary
371	285	TRUE	428	248	natural
372	283	air	428	248	ability
373	282	known	428	248	sound
374	281	themselves	428	248	reach
375	280	already	432	246	least
375	280	hours	432	246	concern
375	280	care	434	245	under
375	280	allow	434	245	space
379	278	government	434	245	appear
379	278	john	437	244	given
379	278	value	438	242	name
382	276	modern	438	242	past
383	275	anything	438	242	strong
383	275	sleep	441	241	future
385	274	u	441	241	half
385	274	continue	443	240	paragraph
387	273	please	443	240	probably
388	272	likely	445	239	experience
389	271	television	446	238	remember

447	237	british	503	208	save
447	237	writer	505	207	music
447	237	support	505	207	remain
450	236	development	507	206	major
450	236	receive	507	206	similar
452	234	getting	507	206	town
452	234	land	510	205	lost
454	233	quite	510	205	realize
455	232	certain	512	204	skill
455	232	europe	512	204	store
455	232	looking	514	203	actually
458	231	coffee	514	203	travel
458	231	real	514	203	worker
460	230	doesn	514	203	activity
460	230	international	518	202	free
460	230	present	518	202	team
460	230	sit	520	200	face
464	229	suggest	520	200	knew
465	228	humans	520	200	light
465	228	million	520	200	whole
467	226	love	520	200	stress
467	226	pass	525	199	especially
469	225	popular	525	199	game
470	224	foreign	525	199	hear
470	224	instead	525	199	paper
470	224	looked	529	198	energy
470	224	side	530	196	else
474	223	languages	530	196	ten
474	223	boy	532	195	nature
474	223	require	532	195	office
474	223	carry	532	195	tend
478	222	friend	532	195	picture
478	222	learning	532	195	sell
480	221	news	537	194	heard
480	221	stand	537	194	open
482	220	adult	539	192	disease
483	219	easy	539	192	eyes
483	219	happy	539	192	move
483	219	death	539	192	sport
483	219	rise	539	192	condition
483	219	send	544	191	cultural
488	218	knowledge	544	191	within
488	218	mean	544	191	researcher
490	217	everything	547	190	economic
490	217	term	548	189	along
492	216	oh	548	189	lived
493	215	felt	548	189	rate
493	215	leave	551	188	french
495	214	base	551	188	hope
495	214	arrive	551	188	surprise
497	211	environment	554	187	gave
497	211	worry	554	187	sign
499	210	main	554	187	accept
499	210	share	554	187	thousand
501	209	amount	558	186	short
501	209	act			
503	208	britain			

# Syntactic Categorization of Ditransitive and Dative Verbs by Japanese EFL Learners

Naoki Sugino<sup>1</sup>, Simon Fraser<sup>2</sup>, Chika Ikeda<sup>3</sup>, Hirokatsu Kawashima<sup>4</sup>, Yuya Koga<sup>5</sup> and Naoki Sugimori<sup>1</sup>

<sup>1</sup>Ritsumeikan University, <sup>2</sup>Hiroshima University, <sup>3</sup>Aichi Prefectural University,

<sup>4</sup>Nagasaki University of Foreign Studies, <sup>5</sup>Shukugawa Gakuin College

gwisno@is.ritsumeikan.ac.jp, fraser@hiroshima-u.ac.jp,

iked@for.aichi-pu.ac.jp, kawashima@tc.nagasaki-gaigo.ac.jp,

koga@shukugawa-c.ac.jp, naosugi@is.ritsumeikan.ac.jp

## Abstract

The present study attempts to explicate how different classes of verbs are represented by Japanese EFL learners with different proficiency levels. Sugino et al. (2008), focusing on the English psych verbs, reported that (a) the categorization of the verbs by the participants defied the theoretically-predicted two categories, i.e., the experiencer-subject (ES) verbs and the experiencer-object (EO) verbs, but (b) each verb was represented with different weighting patterns. Also observed were (c) the high proficiency group displayed three clusters, one of which appeared to reflect characteristics of the EO verbs, although the EO verbs were not always mapped onto this cluster, and (d) the low proficiency group displayed four weighting patterns, none of which corresponded to the EO-ES distinction, and (e) the ES verbs appeared to be represented in two groups.

Using the same research design, this study focuses on verbs that can be used both in the ditransitive and the prepositional dative constructions and those that can be used only in the prepositional dative construction. Grammaticality judgment scores are used to give differentiated weightings to each verb, and the data are analyzed using self-organizing maps (Kohonen, 1997). Major findings of the present study will be reported in the presentation.

## Keywords

English verbs; ditransitive; dative; self-organizing maps; mental lexicon; representation

## Introduction

This study is carried out as part of our research project that aims at developing a computer-mediated vocabulary learning tool that displays a learner's state of the mental lexicon three-

dimensionally. In order to develop such a tool, it is necessary to grasp how a word is represented in relation to other words. In our review of the previous studies in second language mental lexicon (Fraser et al., 2008), it has been pointed out that the network structure of the mental lexicon consists of several subnetworks, each of which represents the subclasses of related words. Some examples of such related words are different classes of English verbs that share common lexico-syntactic features (Goldberg, 1995; Levin, 1993; Pinker, 1989).

In the context of SLA, such clusters can function as facilitative factors in learning new vocabulary items. For instance, if a learner has established a cluster with sufficient members and is aware of the shared features, s/he can use that knowledge to infer the meaning and syntactic behavior of a new word that also shares the same features. At the same time, it has been also pointed out that some of the clusters pose difficulties for learners because different clusters share similar semantic or syntactic features. One such difficulty is concerned with dative alternation. While some of the verbs can be used both in the ditransitive construction and in the prepositional dative construction, there is a cluster of verbs that can be used only in the prepositional dative construction. The present study focuses on this aspect and attempts to elucidate how Japanese EFL learners represent these two classes of English verbs.

## 1 The present study

### 1.1 Target verbs and structures

Two classes of English verbs, four verbs that can be used in both constructions and for that can be used only in the prepositional dative construction are used in the present study. All of them are in 900 words that are familiar to the Japanese EFL learners

(Yokokawa, 2006), and therefore, it would be safe to assume that the participants in the present study possessed some knowledge of these verbs.

Table 1: Verbs used in the present study

Ditransitives		Datives	
V	familiarity rank	V	familiarity rank
<i>find</i>	640	<i>carry</i>	640
<i>give</i>	61	<i>discover</i>	577
<i>make</i>	87	<i>report</i>	447
<i>send</i>	877	<i>select</i>	796

Note: Based on Yokokawa (2006)

These verbs were used in all of the following 10 sentence patterns, yielding 80 target sentences.

Table 2: Target sentence structures

		Dit.	Dat.
F1	Human NP1 + V + NP2 + NP3	5	1
F2	Inanimate NP1 + V + NP2 + NP3	5	1
F3	Human NP <sub>i</sub> + V + NP <sub>ii</sub> + PP	5	5
F4	Inanimate NP <sub>i</sub> + V + NP <sub>ii</sub> - PP	5	5
F5	NP2 + <i>be</i> Vp.p. + NP3 + <i>by</i> Human NP1	5	1
F6	NP2 + <i>be</i> Vp.p. + NP3 + <i>by</i> Inanimate NP1	5	1
F7	NP3 + <i>be</i> Vp.p. + NP2 + <i>by</i> Human NP1	1	1
F8	NP3 + <i>be</i> Vp.p. + NP2 + <i>by</i> Inanimate NP1	1	1
F9	NP <sub>ii</sub> + <i>be</i> Vp.p. + PP + <i>by</i> Human NP <sub>i</sub>	5	5
F10	NP <sub>ii</sub> + <i>be</i> Vp.p. + PP + <i>by</i> Inanimate NP <sub>i</sub>	5	5

Note: "Dit." designates "ditransitive", and "Dat." "dative". "5" indicates that the sentence is well-formed, while "1" grammatically incorrect.

## 1.2 Participants and procedure

The participants were Japanese EFL learners studying at various universities and colleges in Japan. The participants were first instructed to take a general English language proficiency test, *Measure of English Grammar* (Shimizu et al., 2006). The 80 target sentences, with 20 filler sentences, were randomly ordered and divided into two test sets, each consists of 50 grammaticality judgment items. These tests were administered either on-line or as paper-based tests, as appropriate to the testing situation. The participants were given approximately 20 minutes to complete each.

## 2 Analysis

Since this study focuses on how the two classes of verbs are represented in the learner's mental lexicon, the analysis focused on the response patterns each participant attributed to each of the target verbs. In order to understand the logic behind the procedure adapted in this study, suppose a hypothetical learner A, who has acquired full competence in the two classes of verbs. A's responses to the target sentences (1a) to (1j) containing *give* will be as are shown in square brackets:

- (1) a. Mr. Jones gave me some money. [5]
- b. The company gave him a new job. [5]
- c. Mr. Jones gave some money to me. [5]
- d. The company gave a new job to him. [5]
- e. I was given some money by Mr. Jones. [5]
- f. He was given a new job by the company. [5]
- g. Some money was given me by Mr. Jones. [1]
- h. The new job was given him by the company. [1]
- i. Some money was given to me by Mr. Jones. [5]
- j. A new job was given to him by the company. [5]

Thus, A's knowledge about *give* can be represented as (2):

$$(2) \text{ give}_A(5, 5, 5, 5, 5, 5, 1, 1, 5, 5)$$

Similarly, A's responses to the other ditransitive verbs will be the same as (2). Therefore, we can assume that A has formed a cluster of verbs, all sharing the same response patterns as (2').

$$(2') \text{ give}_A(5, 5, 5, 5, 5, 5, 1, 1, 5, 5), \\ \text{find}_A(5, 5, 5, 5, 5, 5, 1, 1, 5, 5), \\ \text{make}_A(5, 5, 5, 5, 5, 5, 1, 1, 5, 5), \\ \text{send}_A(5, 5, 5, 5, 5, 5, 1, 1, 5, 5)$$

On the other hand, A's knowledge about the other cluster of verbs can be represented as in (3):

$$(3) \text{ carry}_A(1, 1, 5, 5, 1, 1, 1, 1, 5, 5), \\ \text{discover}_A(1, 1, 5, 5, 1, 1, 1, 1, 5, 5), \\ \text{report}_A(1, 1, 5, 5, 1, 1, 1, 1, 5, 5), \\ \text{select}_A(1, 1, 5, 5, 1, 1, 1, 1, 5, 5)$$

In this way, it would be possible to depict a learner's representation of a verb as a set of responses; how the words are clustered, and how they are different, can be understood in terms of similarities in the response patterns.

Let us further suppose that other hypothetical learners, with the same level of mastery in the two classes of verbs, participated in the present study.

Their response patterns would be identical with those of A's, and we can observe that they share the same categorization of the verbs.

As a matter of course, the real learners do not judge the grammaticality as clearly and correctly as these hypothetical learners. There will be some gradual differences and similarities among the participants' response patterns.

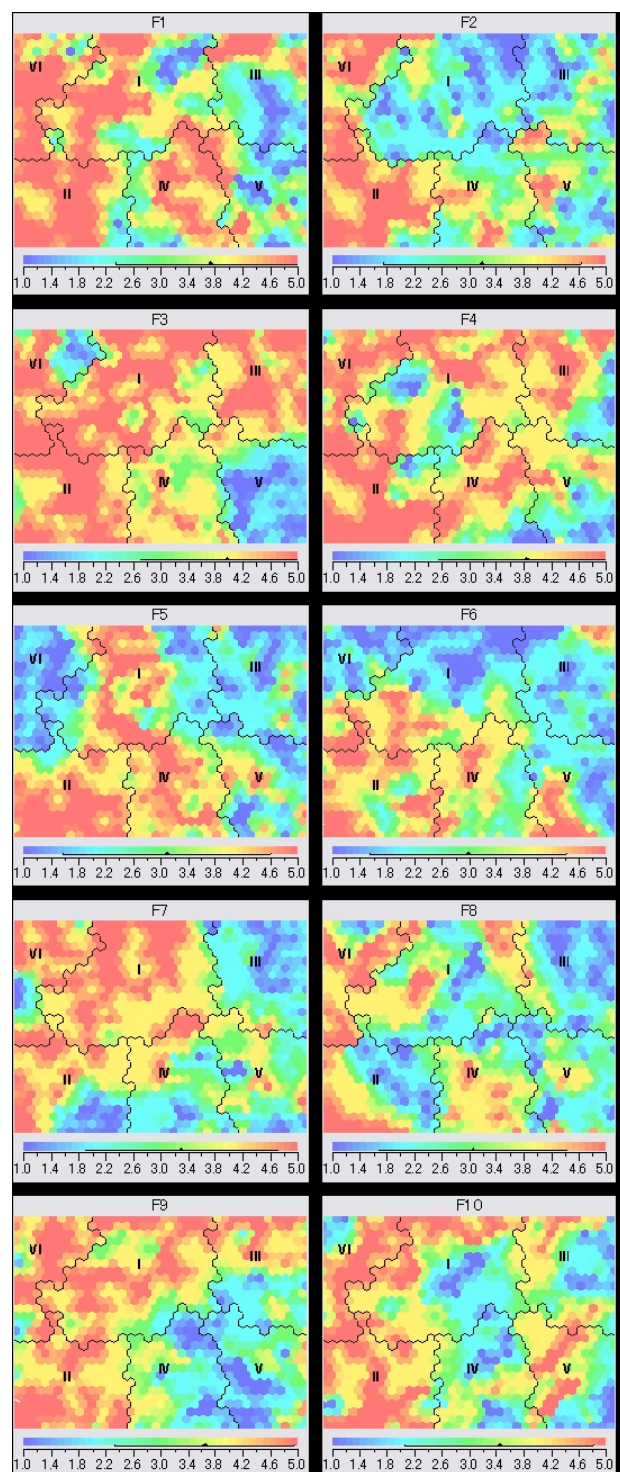


Figure 1: Responses to each sentence structure

Kohonen (1997)'s self-organizing map (SOM) is employed to capture how each of the verbs are represented by individual learners and to what

degree the representations of the verbs are shared among the participants. SOM is a kind of cluster analysis, in which a "node" with a weight vector is positioned onto a two-dimensional map.

New data, represented as new nodes with similar vectors, are placed closer to the original node, while those with different vectors are placed further away from the original. Thus, the response patterns each individual attributes to each of the verbs are regarded as these weight vectors, yielding topological mappings of individual learner's representation. Consequently, if the nodes concentrate in one cluster on the map, it would mean that the response patterns are shared by the participants. By analyzing the characteristics of the cluster, the nature of the shared representation can be obtained.

### 3 Results and observation

The data from 50 participants from a university were examined as a preliminary analysis. Figure 1 shows participants' responses in each of the 10 sentence structures (Table 2 above).

Combining these 10 figures, the software supported 6 clusters that share similar response patterns, as is displayed in Figure 2.

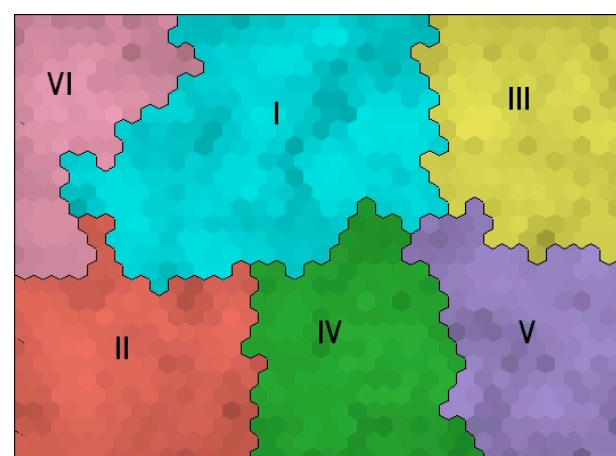


Figure 2: 6 representational clusters

Cluster I can be characterized by relatively lower responses to F2 and F5. Cluster II shows relatively higher responses to all sentence structures but less in F7 and F8, and therefore, displays the characteristics of the ditransitive verbs. On the other hand, lower responses in F5, F6, F7 and F8 characterize Cluster III; taken together with the relatively lower response in F1 and F2 in this area, Cluster III displays the characteristics of the dative verbs. In between Clusters I and II is Cluster IV, characterized by relatively lower, or sometimes mixed, responses to F7, F9 and F10. Low responses in F3 and mixed responses in other sentence structures are the characteristics of Cluster V, and

those lower responses in F5 and F6 characterize Cluster VI.

Due to the limitation of space, only a few examples can be reported here. Figure 3 shows the responses to *give*, and clearly, we can observe strong concentration in Cluster II.

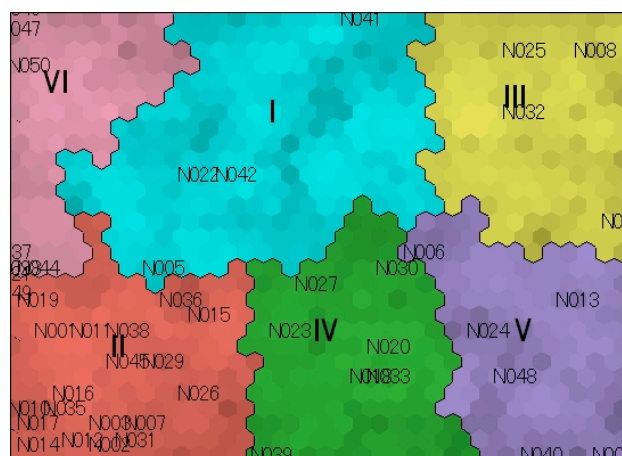


Figure 3: Mapping of responses to *give*

However, there were no dative verbs that displayed clear concentration in Cluster II, except for weak concentration of *discover* as shown in Figure 4.

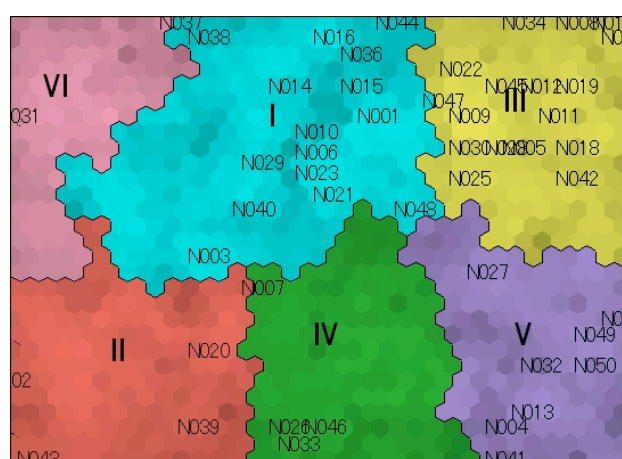


Figure 4: Mapping of responses to *discover*

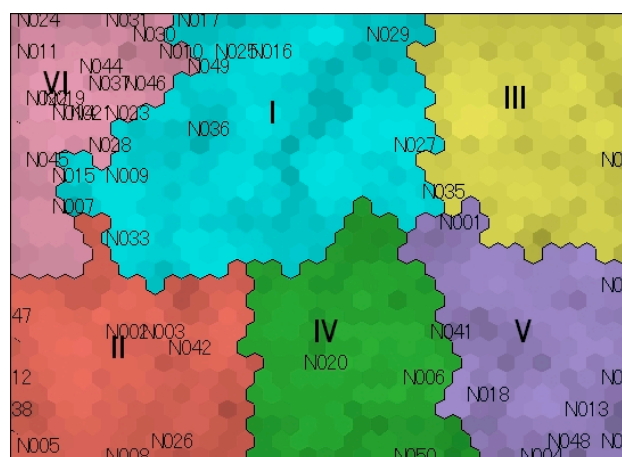


Figure 5: Mapping of responses to *send*

Some verbs concentrated in clusters other than those that shares target characteristics. One example is *send*, as shown in Figure 5: The participants' responses gathered in Cluster VI, characterized by the negative responses in F5 and F6. In order for the response pattern to be ideal, learners would need to learn that *send* cannot be used in the sentence structures F1, F2, F7 and F8. This implies that the SOM can be used diagnostically in understanding learners' difficulty in vocabulary acquisition.

In the presentation, results from larger samples will be discussed and implications and points for further studies will be sought for.

## Acknowledgement

The present study is supported by Grant-in-Aid for Scientific Research (B) (19320088) from Japan Society for the Promotion of Science.

## References

- Fraser, S., Sugino, N., Koga, Y., Kawashima, H., Ikeda, C. and Sugimori, N. (2003). Methodological considerations of mental lexicon research: Focusing on representational links among lexical entries. *Hiroshima Studies in Language and Language Education*, 11, 133-151.
- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago, IL.: University of Chicago Press.
- Kohonen, T. (1997). *Self-organizing maps*, 2nd edition. Heidelberg: Springer-Verlag.
- Levin, B. (1993). *English verb classes and alternations*. Chicago, IL.: University of Chicago Press.
- Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, Mass.: MIT Press.
- Shimizu, Y., Yamakawa, K., Sugino, N., Ohba, H., and Nakano, M. (2006). Developing a sample-free grammatical proficiency test for SLA research. *Proceedings of the 10th Conference of Pan-Pacific Association of Applied Linguistics*, 227-236.
- Sugino, N., Koga, Y., Ikeda, C., Kawashima, H., Fraser, S. and Sugimori, N. (2008). Investigating lexical relatedness in EFL learners' mental lexicon: Focusing on the syntactic features of English verbs. Poster presentation at CLaSIC 2008, National University of Singapore, December 5, 2008.
- Yokokawa, H. (Ed.) (2006). *Nihonjin eigo gakushusha no eitango shinmitsudo: Moji-hen*. Tokyo: Kuroshio.

# Quantitative Analysis of Speaking in Group Oral Interaction – Range of Speaking on the Basis of CEF Assessment –

Junko Negishi

Open Education Center, Waseda University

jnegishi@dream.com

## Abstract

Speaking assessments in Japan are meant to facilitate the use of picture description tasks or interviews by individual examinees, while group oral discussions are rarely carried out. Given this factor, this study aimed first to find some features and developmental phenomena of the participants as second language learners of English at junior high school, senior high school and university level when they underwent group oral discussions in a group of three. Second, the relationship between the results obtained from the assessment given by raters utilizing CEF rating scales and the variables obtained by the analysis of subcategories was investigated. This paper solely reports the results of investigation about Range among the five rating subcategories because of the space constraint.

For the first question, speaking ability of the participants given by multi-faceted analysis was converged to Basic Users (A1, A2) and Independent Users (B1, B2) and no students were allocated to Proficient Users (C1, C2). In terms of the second question, one of the subcategories, Range, was explored by analyzing six items. The items that explained Range most was Guiraud index followed by the number of complex sentences.

## Keywords

Speaking, Group discussion, CEF, Rating

## 1 Introduction

### 1.1 Group oral discussion

The emphasis of English Test as a second language in Japan has been put on writing for a long time. It was not until a few years ago that listening test was introduced to National Center Test for Universities. The idea of giving an oral performance test to students would be the last choice as it is costly and time consuming.

Few oral performance tests available in Japan at present are predominantly for a single candidate. Nonetheless, individual oral performance tests typified by monologue cannot be regarded as an actual language use (Long, 2000). Interviews are

also considered to be problematic for the reason that they are inauthentic and asymmetric (Johnson & Tyler, 1998).

English ability that society calls for in the present circumstances is speaking ability in a discussion or communication to some extent where speakers negotiate meanings, which suggests that there is a discrepancy between what schools focus on and what society requires.

Group discussions have been introduced, for example, to Cambridge First Certificate in English and to an oral examination practiced by Council of Europe (2001) within European Union which has multi-languages and multi-cultures. It is impractical to measure overall speaking ability by employing a single group oral discussion; however, it has some advantages. For example, speaking activities in classrooms can be utilized and it is practical because it is what companies and society call for. Swain (2001) reports that participants can co-construct a dynamic representation of interaction in a given situation. In addition, teachers and raters can concentrate on ratings with less time to more examinees, which could be utilized in schools.

### 1.2 Ratings of group oral discussion using CEF rating scales

The rating procedure for oral performance is classified into two: a holistic rating and an analytic rating. Council of Europe (2001) mentioned above has released its Common European Framework of Reference for Languages: learning, teaching, assessment, (CEF hereafter) in which it gives two kinds of rating scales: CEF Global Oral Assessment Scale for a holistic assessment and CEF Oral Assessment Criteria Grid for an analytical assessment. What they have in common is that there are 6-level 'can-do' descriptors, that is, A1, A2 for Basic Users, B1, B2 for Independent Users, and C1, C2 for Proficient Users. The difference between the two rating scales is that the former sees an examinee's performance from a holistic point of view and the latter has 5 subcategorized analytic grids: Range, Accuracy, Fluency, Interaction, and

Coherence. Table 1 shows the former, a part of CEF Global Oral Assessment Scale. The latter is on Table 3 below.

Table 1: CEF Global Oral Assessment Scale

C2	Conveys finer shades of meaning precisely and naturally.
C1	Shows fluent, spontaneous expression in clear, well-structured speech.
B2	Expresses points of view without noticeable strain.
B1	Relates comprehensibly the main points he/she wants to make.
A2	Relates basic information on, e.g., work, family, free time, etc.
A1	Makes simple statements on personal details and very familiar topics.
Below A1	Does not reach the standard for A1.

### 1.3 Purpose of the study

As mentioned above, speaking assessments in Japan are meant to facilitate the use of picture description tasks or interviews by individual examinees, while group oral discussions are rarely carried out. Given this factor, this study aimed first to find some features and developmental phenomena of the participants as second language learners of English at junior high school, senior high school and university level when they underwent group oral discussions in a group of three. Second, the relationship between the results obtained from the assessment given by raters utilizing CEF rating scales and the valuables obtained by the analysis of subcategories was investigated. This paper solely reports the results of investigation about Range among the five rating subcategories.

## 2 Method

### 2.1 Participants

The participants of the study were 135 students who had been studying English at junior high schools, senior high schools, and universities in and around Tokyo area. Forty-five groups were composed by the fifteen groups from three levels of educational institutions: junior high, senior high and university respectively. Each group comprised three students participating in a discussion according to a given topic. Table 2 shows the number of participants in each educational institution, public or private, and the ratio of males and females.

Table 2: Number of Participants in Each Educational Institution

	Public School		Private School		Male	Female	Total
	Participants	Groups	Participants	Groups			
Junior High School	27	(9)	18	(6)	8	19	27
					18	0	18
Senior High School	24	(8)	21	(7)	19	5	24
					21	0	21
University	15	(5)			8	7	15
			15	(5)	6	9	15
			15	(5)	4	11	15
Total	66	(22)	69	(23)	84	51	135

### 2.2 Data collection

Spoken data of the participants were collected in the following way: 1) three students were allocated to one group at random, 2) each group drew one card from five on which different topics were

written as School, Family, Friends, Hobby and Culture, 3) the students were informed to have an oral interaction about the topic for five minutes after having a five-minute planning time individually, 4) before the interaction, each member of the group introduced themselves for about thirty seconds as a warm-up activity, 5) they started an oral interaction based on the topic they drew looking at each other's face. The oral interaction was videotaped and later burned to a DVD for rating. The participants' conversation was subsequently transcribed for analysis.

### 2.3 Raters

Raters were ten Japanese teachers of English who had Master's degrees in the field of English education or linguistics. Before rating, they received training using a DVD produced by the Council of Europe. They rated the students by applying both a holistic rating scale and analytic rating criteria. The latter consists of five subcategories; Range, Accuracy, Fluency, Interaction, and Coherence. The raters assessed the students while watching a DVD by 7 scales: Below A1, A1, A2, B1, B2, C1, C2.

### 2.4 Multi-faceted Rasch analysis

In most of the oral performance tests as a second language, raters assess participants' speaking ability by means of rating scales. No matter how precise the rating scales are, the assessment might be subjective or inconsistent because of the fact that ratings are administered by humans. Training is mandatory which can reduce irregular errors and can improve inter-rater reliability and intra-rater consistency; however, it is impossible to remove all the severity or leniency of the ratings (McNamara, 1996; Weigle, 1998).

Multi-faceted analysis compensates diverse influences as much as possible, which are caused by many facets such as raters' differences, task difficulties, examinees' abilities, and their interactions. It estimates and transforms original 'raw scores' given by raters to 'measures,' which are on an interval scale. This study utilized multi-faceted analysis software, FACETS Version No. 3.6.4.0 (Linacre, 2008) to eliminate subjective judgement as much as possible.

### 2.5 Items that may represent Range

What items should represent a subcategory, Range? Table 3 shows CEF Oral Assessment Criteria for Range. Seeing that the participants were proved to be within the range of either A or B on the scale, several indicative words were extracted from the criteria: words and simple phrases (A1), groups of a few words and formulae (A2), sufficient vocabulary

(B1), complex sentence forms (B2). According to them, the following items, 1) to 6), were selected for the analysis in the study.

Table 3: CEFR Oral Assessment Criteria (Range)	
C2	Shows great flexibility reformulating ideas in different linguistic forms to convey finer shades of meaning precisely, to give emphasis, to differentiate and to eliminate ambiguity. Also has a good command of idiomatic expressions and colloquialisms.
C1	Has a good command of a broad range of language allowing him/her to select a formulation to express him/herself clearly in an appropriate style on a wide range of general academic, professional or leisure topics without having to restrict what he/she wants to say.
B2	Has a sufficient range of language to be able to give clear descriptions, express viewpoints on most general topics, without much conspicuous searching for words, using some complex sentence forms to do so.
B1	Has enough language to get by, with sufficient vocabulary to express him/herself with some hesitation and circumscriptions on topics such as family, hobbies and interests, work, travel and current events.
A2	Uses basic sentence patterns with memorised phrases, groups of a few words and formulae in order to communicate limited information in simple everyday situations.
A1	Has a very basic repertoire of words and simple phrases related to personal details and particular concrete situations.
Below A1	

- 1) Number of tokens (Tokens) which indicates the total number of words that a participant spoke in a five-minute oral interaction in a group of three.
- 2) Number of different Types of words (Types), or index words, that is, how many different words were used in the same text.
- 3) Type-token ratio (TTR): the ratio of different words to total words used, which has been used to indicate lexical diversity.
- 4) Guiraud index (Guiraud): TTR has been reported as problematic because large number of tokens produces lower TTRs than smaller ones. As for the reason that the participants produced different sizes of spoken samples in a five-minute interaction, mathematical transformations of TTR were carried out according to Guiraud (1960) to make it independent of the sample size which were calculated as  $\text{Types}/\sqrt{(\text{Tokens})}$
- 5) Chunks: groups of a few words and formulae were labeled as chunks in the study.
- 6) Complex sentences (Cmplx S.) including not only literal complex sentences but also other syntactically complex sentences such as embedded sentences, sentences with relative clauses, and also compound sentences because most of the sentences the participants uttered were simple sentences and there were not many real 'complex sentences.'

### 3. Results and Discussion

#### 3.1 Multi-faceted analysis and Measures

The results produced by the multi-faceted software, FACETS, demonstrated that difficulties of rating items did not indicate considerable difference and that rating items were proved to fit Rasch model. Reliability of rating scales was confirmed when the ten raters assessed the participants by means of the CEF rating criteria. In terms of the raters' severity, they showed different severities and their ratings turned out to have fairly large differences while the raters were proved to be self consistent.

Although FACETS analysis provided varied

information such as task difficulty, participants' ability, rater severity and so forth, this study solely used the produced interval scores or 'measures' for Global and for one of the subcategories, Range.

#### 3.2 Global speaking ability

Figure 1 shows a histogram of the measures supplied by multi-faceted analysis. There were more lower level students than average = '0' since one-thirds of the participants were junior high school students who were novice learners of English.

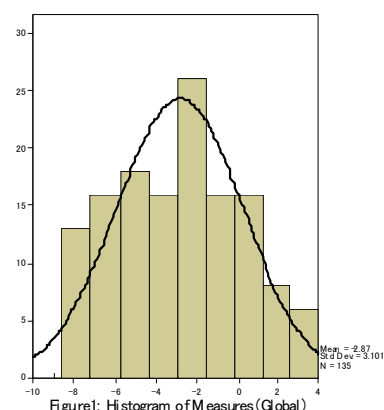
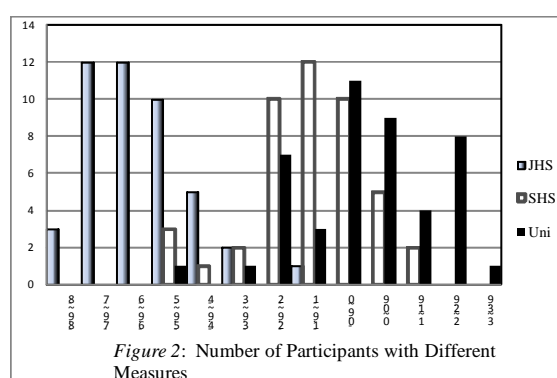


Figure 2 demonstrates the number of participants divided by the three educational institutions they belonged to (junior/senior high school and university) with interval scales (unit: logit) produced by FACETS. The left side of the figure or the smaller number indicates that the students' speaking ability was determined as low and the right side or the bigger number as high.



An average measure of junior high participants was -6.27 logits (A1 for CEF assessment), senior high was -2.40 logits (A2), and university was 0.06 logits (B1), all of which showed a clear difference of the participants which educational institution they belonged to. While junior high participants tended to remain on the left side and did not overlap, the other two, senior high and university students, tended to overlap; however, no senior high students

obtained measures over 2 logits and their peaks for senior high and university had a 2-logit difference. Senior high and university students showed bigger variances compared to junior high students.

Raters assessed not only Global speaking ability but also five subcategories, Range, Accuracy, Fluency, Interaction, and Coherence. However, the following analyses were conducted on the measures of Range. The Pearson correlation coefficient between Global and Range measure was very high, showing .995.

### 3.3 Range Variables

#### 3.3.1 1) Number of tokens and 2) different types of words

To analyze the number of tokens and different types of words, a software 'v8an' included in JACET list of 8000 basic words (JACET 2003; JACET 8000 hereafter) was utilized. JACET 8000 was originally built up from British National Corpus adding some Japanese-specific words. It lemmatized the transcribed text and then ranked them from level 1 (basic 1000 words) to level 8 (7001 to 8000 words approximately). It also produced the number and percentage of the total tokens or words, and different types of words or index words, from the text. The words that were not on the list were grouped as 'over 8' and contracted forms as 'cont.', proper nouns as 'prop.' and non-words such as numbers as 'non.'

Table 4: Number of Tokens and Different Types of Words

	level 1	level 2	level 3	level 4	level 5	level 6	level 7	level 8	over 8	cont. forms	proper nouns	non-words	total
<b>Tokens</b>													
Total	8210	337	172	59	48	119	7	44	172	141	8	225	9542
%	86.0%	3.5%	1.8%	0.6%	0.5%	1.2%	0.1%	0.5%	1.8%	1.5%	0.1%	2.4%	100.0%
Average	60.81	2.50	1.27	0.44	0.36	0.88	0.05	0.33	1.27	1.04	0.06	1.67	70.68
<b>Types</b>													
Total	4338	259	124	48	30	76	7	33	130	89	5	176	5315
%	81.6%	4.9%	2.3%	0.9%	0.6%	1.4%	0.1%	0.6%	2.4%	1.7%	0.1%	3.3%	100.0%
Average	32.13	1.92	0.92	0.36	0.22	0.56	0.05	0.24	0.96	0.66	0.04	1.30	39.37

Table 4 shows the total number and percentage of tokens (Tokens) and different types of words (Types) of 135 participants produced by v8an (JACET 8000). On the condition that the participants had an oral interaction for five minutes in a group of three, the average Tokens were 70.7 words, out of which, words uttered from level 1 were 60.8 words showing 86% of the total tokens. On the other hand, the average Types were 39.4 words, out of which, words used from level 1 were 32.1 words indicating 81.6% of the total Types.

Figure 3 displays the average Tokens the participants used and Figure 4 shows Types, both of which were separated by the three educational institutions. Both figures showed that the higher educational institution they belonged to, the more Tokens and Types they uttered, that is, their vocabulary size grew. With regard to the use of level 1 words, both the Tokens and Types increased

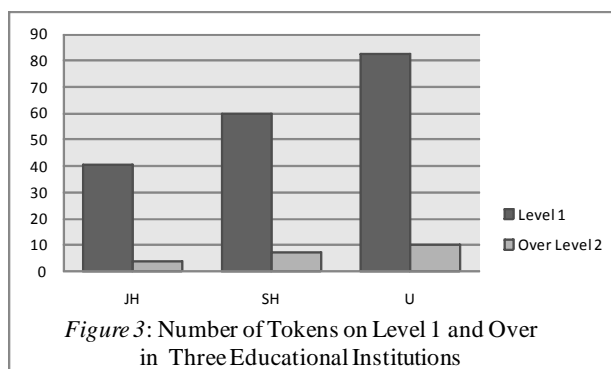


Figure 3: Number of Tokens on Level 1 and Over in Three Educational Institutions

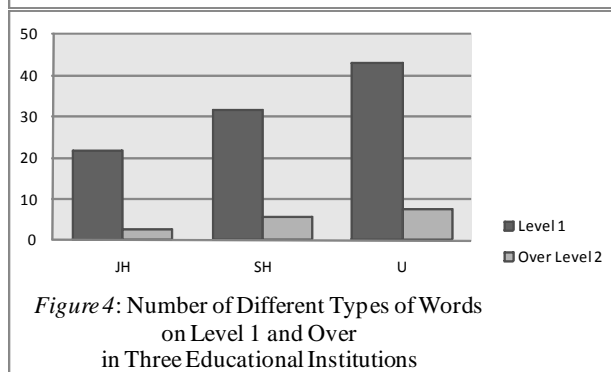


Figure 4: Number of Different Types of Words on Level 1 and Over in Three Educational Institutions

nearly 1.5 times from junior high to senior high and around 2 times from junior high to university. In terms of the use of level 2 to over 8 words, both increased two times from junior high to senior high and around 2.8 times from junior high to university, which means that the participants demonstrated that they acquired more difficult words when they proceeded to the higher institutions.

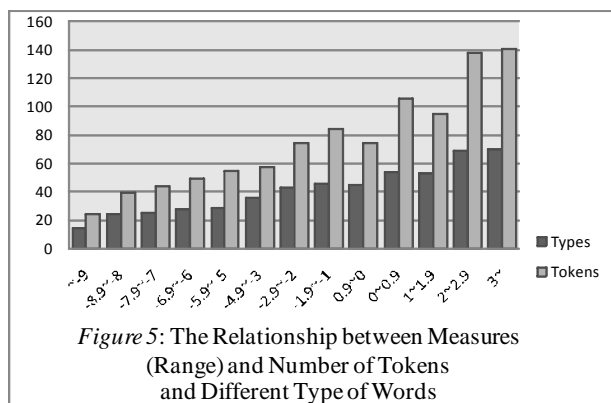


Figure 5: The Relationship between Measures (Range) and Number of Tokens and Different Type of Words

Figure 5 displays Tokens and Types the speakers uttered in a five-minute conversation according to the measures they received regardless of their institutions. As shown in the figure, the more measures the speakers gained, or the more proficient they became, the more Tokens and Types they tended to utter. Although Pearson correlation coefficient between the measures for Range and Tokens was .693\*\* and Types was .786 \*\*, showing significant correlation, there were a few reversed part. Qualitative analysis should be conducted to explore the reason for the inverse.

### 3.3.2 3) TTR and 4) Guiraud index

Type-token ratio (TTR), the ratio of different words to total words used, is often used to measure lexical diversity as it has been believed to imply language learners' active vocabulary size. Nonetheless TTR has equally been reported as problematic because large number of tokens produces lower TTRs than smaller ones. As the participants of the study produced different sizes of spoken samples in a five-minute interaction, mathematical transformations of TTR was carried out according to Guiraud (1960) to make it independent of the sample size. Guiraud index (Guiraud) was calculated as  $\text{Types}/\sqrt{\text{Tokens}}$ .

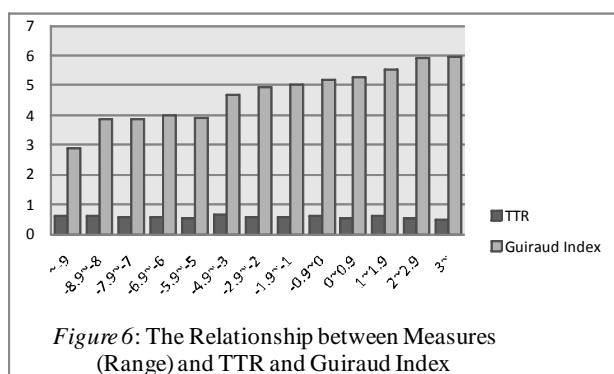


Figure 6: The Relationship between Measures (Range) and TTR and Guiraud Index

Figure 6 shows the number of TTR and Guiraud index with different levels of the participants according to the FACETS measures of Range. As shown in the figure, TTR stayed stable regardless of the speakers' proficiency. On the other hand, Guiraud index can be said that it had a proportional relation between the proficiency levels. The Pearson's correlation coefficient between the measures for Range and Guiraud index was significant .791,  $p < .001$ ; however, for Range and TTR was -.142, indicating very weak inverse correlation. This result suggested that Guiraud index could be an indicator for Range, but not TTR.

### 3.3.3 5) Chunks and 6) complex sentences

Chunks – groups of a few words and formulae – and complex sentences were counted.

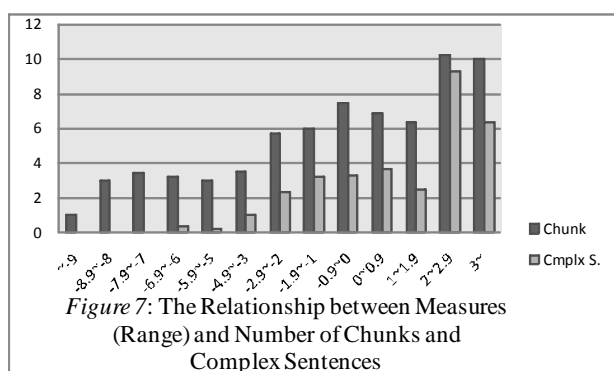


Figure 7: The Relationship between Measures (Range) and Number of Chunks and Complex Sentences

As shown in Figure 7, in terms of the number of chunks (Chunks) the participants used, the more they became proficient, the more Chunks they used; however, it was not always the case. The same phenomenon was seen in the number of complex sentences (Cmplx S.). The participants who received less than -7.0 logits employed no Cmplx S. at all. It is likely that the raters regarded the speakers as less proficient when they spoke only simple sentences.

## 3.4 Correlations

The relationship between the Range measures and variables of six items were explored by means of Pearson correlation coefficient.

Variables	Measure	Tokens	Type	TTR	Guiraud	Chunks
1 Measure (Range)	-					
2 No. of tokens (Tokens)	0.693 **	-				
3 No. of index words (Types)	0.786 **	0.938 **	-			
4 Type-token ratio (TTR)	-0.17 *	0.004	-0.386	-		
5 Guiraud index (Guiraud)	0.788 **	0.694 **	0.894 **	0.004	-	
6 No. of chunks (Chunks)	0.593 **	0.754 **	0.702 **	-0.447 **	0.543 **	-
7 No. of complex sentences (Cmplx S.)	0.648 **	0.705 **	0.691 **	-0.318 **	0.542 **	0.621 **

Note: \*\* $p < .01$ , \* $p < .05$  (two-tailed tests)

As Table 5 shows the correlation between the measure and Types was the most significant, followed by Guiraud, Cmplx S., Tokens, and Chunks. Only TTR indicated that there was no correlation toward the measure. In summary, the three variables –Types, Guiraud and Cmplx S. – might be good indicators to explain whether or not the speaker had wide range of speech.

## 3.5 Multiple regression

Model	Unstandardized Coefficients		Standardized Coefficients	<i>t</i>	Sig.
	<i>B</i>	<i>Std. Error</i>	<i>Beta</i>		
Step 1					
Constant	-17.062	0.964		-17.696 **	
Guiraud index	2.983	0.202	0.788	14.741 **	
Step 2					
Constant	-14.176	1.203		-11.787 **	
Guiraud index	1.994	0.422	0.526	4.729 **	
No of Cmplx S.	0.328	0.082	0.277	4.011 **	
No of Types	0.026	0.027	0.124	0.96	

Note:  $R^2 = .620$  for Step 1, \*\* $p < .001$ ;  $R^2 = .692$  for Step 2, \*\* $p < .001$

Multiple regression analysis was conducted to predict the Range measures from several predictor variables. A value of  $R^2$  which is a measure of how much of the variability in the outcome was accounted for by the predictors. For the first model its value was .620, which meant that Guiraud accounted for 62.0% of the variation in Range measures. To obtain the other two predictors, the multiple regression analyses were carried out several times and some variables were excluded to eliminate the variables that showed collinearity. When the other two predictors, Cmplx S. and Types were included as well, this value increased to

69.2%.

The b-values (*B*) in Table 6 indicate the relationship between Range and each predictor. As the value was positive, there was a positive coefficient which subsequently represented a positive relationship. The b-value of Guiraud, 2.983, indicated that as Guiraud increased by one, Range measure increased by 2.983. This interpretation was true only if the effects of the other two valuables were held constant. The multiple regression model would then be:

$$\text{Range}_i = b_0 + b_1 \text{Guiraud}_i + b_2 \text{Cmplx S.}_i + b_3 \text{Types}_i \\ = -14.176 + (1.994 \text{Guiraud}) + (0.328 \text{Cmplx S.}) + (0.026 \text{Types})$$

As the *t*-test associated with a b-value of Guiraud was significant and the smaller the value of Sig., the greater the contribution of that predictor. For this model, Guiraud (*t* (131) = 4.729, *p* < .001) and Cmplx S. (*t* (131) = 4.011, *p* < .001), were significant predictors of Range, but not Types (*t* (131) = .960, *p* = .339).

#### 4 Conclusion

The study mainly investigated two phenomena. The first was to find some features and developmental phenomena of the participants and the second was to search for the relationship between the measures and various variables.

For the first question, speaking ability of the participants given by multi-faceted analysis was converged to Basic Users (A1, A2) and Independent Users (B1) and no students were allocated to Proficient Users (C1, C2). Among them, junior high school students tended to be grouped as A1, who were isolated from the other students. Albeit senior high school students were mostly labeled as A2, some of them showed higher ability than university participants who were labeled as B1 as a whole. Only university students obtained the highest ratings and the difference between the senior high students and university students was 2 logits.

In terms of the second question, one of the subcategories, Range, was investigated by analyzing six items: Tokens, Types, TTR, Guiraud, Chunks, and Cmplx S. All the items except TTR showed good relationships with the measures for Range. The items that proved the most significant correlation coefficient were Types, Guiraud, and Cmplx S.; however, multiple regression suggested that Guiraud and Cmplx S. which should explain Range well. More detailed investigation is essential by means of qualitative analysis.

What could be said to bridge the results with pedagogical purposes? To widen the range of learners' speaking ability, improving his or her

vocabulary would be imperative which includes phrases and chunks. Teaching how to use complex sentences would be indispensable as well. For example, a complex sentence using a conjunction 'when' is taught to second year of junior high school students and a sentence with relative clause to third year, which means it is not unachievable for at least senior high school students to make use of complex sentences in speaking. What teachers need to do in a classroom is to provide opportunities to express themselves using English.

#### References

- Council of Europe (2001). *European language portfolio*. Retrieved March 7, 2008, from [www.coe.int/portfolio](http://www.coe.int/portfolio).
- Guiraud, P. (1960). *Problemes et methods de la statistique linguistique*. Dordrecht: Reidel.
- JACET (2003). *JACET list of 8000 basic words*. Tokyo: Japan Association of College English Teachers.
- Johnson, M., and Tyler, A. (1998). Re-analyzing the OPI: How much does it look like natural conversation? pp. 27-51. In R. Young & A. W. He (Eds.), *Talking and testing: Discourse approaches to the assessment of oral proficiency*. Amsterdam: John Benjamins.
- Linacre, J. M. (2008). *Facets Rasch measurement computer program*. Chicago: Winsteps.com.
- McNamara, T.F. (1996). *Measuring second language performance*. New York: Longman.
- Swain, M. (2001). Examining dialogue: Another approach to content specification and to validating inferences drawn from test scores. *Language Testing*, 18, 275-302.
- Young, R. F. (2000). *Interactional competence: Challenges for validity*. Paper presented at the Language Testing Research Colloquium, Vancouver, Canada, March. ERIC.
- Weigle, S.C. (1998) Using FACETS to model rater training effects. *Language Testing* 15, 263-87.

# Towards the Assessment of the Use of Prosodic Cues in Speech Recognition

Remi Murao<sup>1</sup> and Michiko Nakano<sup>2</sup>

<sup>1</sup>Open Education Center, Waseda University, <sup>2</sup> School of Education, Waseda University

[murao@aoni.waseda.jp](mailto:murao@aoni.waseda.jp), [nakanom@waseda.jp](mailto:nakanom@waseda.jp)

## Abstract

In this study, Japanese learners of English's ability to use prosodic cue for sentence segmentation is examined. Five types of sentence pairs were used to explore whether the learners could identify the sentence structure by their prosodic differences; 1) late/early closure, 2) SVOC/SVO, 3) phrasal verbs, 4) compounds, and 5) content/function words.

## Keywords

Sentence segmentation, Prosodic cues, Speech Recognition

## 1 Previous Studies

Previous study has shown that native speakers of English and advanced level English learners utilize prosodic cues to recognize the utterances when segmental information is distorted (Murao, 2006). However, it is reported that the ability to use prosodic cues for sentence segmentation is low for English learners whose first language is a tone language. The study by Pennington & Ellis (2000) showed that Cantonese speakers' memory for English sentences with prosodic cues for (1) phrase structure and (2) linguistic boundary were especially poor.

### (1) Phrase structure

- 1a. *The fight is over Fred.* [The fight is about Fred.]
- 1b. *The fight is over, Fred.* [I am telling you Fred that the fight is finished.]

### (2) Linguistic boundary

- 2a. *She's a lighthouse keeper.* [Her job is to look after a lighthouse.]
- 2b. *She's a light housekeeper.* [She is a housekeeper who does light housework.]

This result implies that there is a tendency for speakers of tone language to store pitch patterns as lexical representation, but they are not used to recognizing and analyzing sentence prosody.

Native speakers of English, on the other hand, are sensitive to sentence prosody in that they use prosodic cues to locate linguistic boundaries from a speech stream when acquiring their first language. Soderstrom, Seidl, Nelson & Jusczyk (2003) have reported in their study that even six-month old infants were sensitive to phrase-level prosodic cues and could differentiate between linguistic units ("Today, *people by the hole* seem scary.") and non-linguistic units ("In fact, some *people # buy the whole* supply of them").

Although it is widely acknowledged that prosodic cue plays a role in "sentence segmentation (deciding whether two adjacent section of speech belong to the same or different linguistic units)", it is questioned that it plays a role in "structure bracketing (discerning structural relations among the linguistic units)" (Gerken, 1996. Page 345). Whether speakers place prosodic breaks and pitch accent to designate hierarchical structure are optional for the speakers. For example, speakers may or may not offer prosodic cues before the subordinate clause in sentences like "*Ramona left [before the party started]*" (Frazier, Carlson & Clifton, 2006. Page 245). According to Frazier et al. (2006) there are some aspects of prosody that are obligatory. For example, speakers use prosody to disambiguate sentences with syntactic ambiguity such as "*When Roger leaves # the house is dark*".

## 2 Present Study

The present study examines whether Japanese learners of English are sensitive to prosodic cues for sentence segmentation. As with the Cantonese speakers in Pennington & Ellis's (2000) study, it is assumed that Japanese speakers have a tendency to attend to lexical prosody over sentence prosody due to their first language background. While Pennington & Ellis (2000) examined the recognition memory for English sentences based on prosodic information, the present study uses a different method to reveal the use of prosodic cues in order to compare the results.

## 2.1 Materials

Five pairs of sentences each in the following five sentence types were developed as the test items. The number of syllables is the same in each pair, and the words used are mostly the same with similar segmental phonemes. The sentence pairs differ in prosodic information such as pause and pitch accent. Words in capital letters show where pitch accent is placed. Twenty-five sentences in total were recorded by a female British speaker of English. The speaker was instructed to pronounce the sentence pairs in a way to differentiate from each other.

- (3) Late/Early Closure
  - 3a. *Since I always DRINK, wine is gone.*
  - 3b. *Since I always drink WINE, it's gone.*
- (4) SVOC/SVO
  - 4a. *He pushed the door CLOSED.*
  - 4b. *He pushed the CLOSED door.*
- (5) Phrasal Verbs
  - 5a. *Put on your SHOES.*
  - 5b. *Put your SHOES on.*
- (6) Compounds
  - 6a. *He's at the GREENhouse.*
  - 6b. *He's at the green HOUSE.*
- (7) Content/function words
  - 7a. *Mary bought HALF sandwich.*
  - 7b. *Mary have bought sandwich.*

## 2.2 Participants

Participants of the present study are thirty Japanese university students studying English as a foreign language. They were asked to report their scores of English proficiency test such as TOEIC, TOEFL, and WeTEC. Five native speakers of English also participated in the study as a control group.

## 2.3 Preliminary Research

The present study aims to examine Japanese learners of English's sensitivity to prosodic cues by using low-path filtered sentences. Test sentences are low-path filtered so as to leave prosodic information such as pauses and pitch accents and distort segmental information. Participants should be able to use prosodic cues to differentiate the sentence structure between pairs. In order to be fully convinced that participants used prosodic cues rather than segmental information, preliminary research was conducted to verify the effectiveness of low-path filtering technique. One native speaker of English and two near-native speakers of English

who are Japanese were asked to listen to sentences of which pitch movement was eliminated and was low-path filtered with cut-off point set to either 400Hz, 325Hz, or 200Hz. The result show that three participants could not differentiate between the two possible sentences ("*Mark loves the cat*" and "*Mike hates the dog*") when the sentence they heard lacked prosodic cue and was low-path filtered with cut-off point at 325Hz. This shows that low-path filtering will rule out the possibility of being able to use segmental information to identify the correct sentence.

## 2.4 Procedures

Participants will hear twenty-five sentences in a random order. All the test sentences are low-path filtered with cut-off point at 325Hz so as to distort the segmental information. Participants are asked to choose the correct sentence from the written sentence pairs. In order to avoid chance-level performance in selecting from the two alternative sentences, (a) and (b), participants were given the choice of (c); "I don't know". The sentences can be listened to for no more than three times.

After selecting from three choices, participants are asked to read aloud the sentence pairs. This will be recorded for the purpose of speech analysis

## 3 Results and Discussion

The result of a pilot study showed that advanced level Japanese learners of English performed exceptionally well on identifying the construction of sentences including phrasal verbs, implying that high frequency collocations such as *put-on*, *shoes-on*, *turn-off*, *light-off*, are stored in their mental lexicon with prosodic representation. Test scores of Japanese learners will be compared with that of native speakers, and difficulties among the sentence types will be discussed.

## References

- Frazier, L., Carlson, K. & Clifton, C. Jr. (2006). Prosodic phrasing is central to language comprehension. *TRENDS in Cognitive Sciences*, 10(6), 244-249.
- Gerken, L. (2009). Prosody's role in language acquisition and adult parsing. *Journal of Psycholinguistic Research*, 25 (2), 345-356.
- Pennington, M. C. & Ellis, N. C. (2000). Cantonese speakers' memory for English sentences with prosodic cues. *The Modern Language Journal*, 84(3), 372-389.
- Soderstrom, M., Seidl, A., Nelson, D. G. K. & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language*, 49, 249-267.

# The effect of two types of feedback on Iranian EFL students' tense-marker production

Somayeh Kazemi Tari<sup>1</sup>

<sup>1</sup>Department of Language and Literature, Science and Research University

somayehkazemitari@gmail.com

## Abstract

Although many researchers have worked on the role of feedback in immersion programs on second language learners' production, still little is known about the role of different types of immediate feedback on the grammatical accuracy of Iranian learners in an EFL setting. When one reviews the literature on feedback, s/he encounters such types of feedback as prompts and recasts and their different subcategories. When, where and with whom, any of these types of feedback might be appropriate? There are some doubts about the effectiveness of immediate types of feedback. Therefore, the researcher wanted to know which types of immediate oral feedback could better help Iranian EFL learners to produce the correct form of the tense markers of their target language—English. In other words, one of the primary objectives of the present study was to examine the effects of two types of oral feedback—namely explicit correction and implicit or-choice feedback — on Iranian EFL learners' production of different tense markers in English. 30 students participated in the study. Analyses of the results of the fill-in-the-blank pretest and posttest showed that explicit correction was more effective than implicit or-choice feedback in producing correct grammatical tense markers by Iranian EFL learners.

## Keywords

Feedback; explicit correction; or-choice feedback; tense-marker production

## Introduction

### 1.

#### ntroduction

The use of feedback in language teaching is nothing new. Traditional language teachers immediately corrected their students' errors when they appeared in the students' utterances. There are different views regarding the use of different types of

feedback. Interlanguage researchers argue that errors should be tolerated as much as they are the result of the learner's communicative strategies. Another view is that errors should be ignored on the whole. Still another view is that some types of errors should be corrected. Ellis (1990) provides taxonomy of the type of errors that should be corrected by the teacher. According to Ellis (1990, p. 54-55) global errors, errors that affect the overall comprehensibility of an utterance, stigmatized errors, and errors relating to the learner's next stage of development should be corrected.

There are a number of researches that have elicited teachers' rationales for correcting errors. Chaudron's (1986, in Chaudron, 1998, p. 136) research is among those. He compared the corrections made by teachers in French immersion classes in Canada and found that teachers corrected more "grammatical" (morphological and syntactic) errors in French classes (average 77%) than in other classes (average 37%). Also, Chaudron (1986, in Ellis, 1990, p. 73) found that fewer discourse errors were corrected compared to morphological errors.

Although the terms "feedback" and "correction" are usually used interchangeably, but Long (1977, cited in Ellis, 1990, p. 71) suggests the term "feedback" be used for the teacher's information about the correctness of the learner's utterances, while the term "correction" be used for the effect of feedback on learning. Ellis (1990, p. 71) uses error treatment and corrective feedback synonymously and states that the broadest definition of treatment is "any reaction by the teacher that clearly transforms, disapprovingly refers to or demands

improvement".

Correction may not always be effective; in order for the correction to be effective some prerequisite must be met. Johnston (1988) states that correction is effective when:

1. The learner desires or needs to eradicate the mistake.
2. She is able to form an internal representation of what the correct behavior looks like.
3. She realizes that her performance is flawed.
4. She has the opportunity to perform in real condition. (cited in Ellis, 1990, p. 181)

Feedback to learners can be either explicit through overt correction or implicit through negotiation work. Negotiation and interaction are important in L2 learning because as Mackey and Philp (1998, in Kaplan, 2002, pp. 172-76) have shown interaction can provide modified speech, confirmation check, comprehension check, clarification request, or-choice, topic-focused, elaboration, on the spot learning, delayed learning, and testing a hypothesis. And according to Mitchell and Myles (2004, p. 195) language learning is highly 'dependent on face to face interaction and shared processes, such as joint problem solving and discussion'.

There are different types of feedback compiled by Lightbown and Spada (2006), Ellis (1994), Ellis (2003), and Doughty and Long (2003). Nassaji provides three types of focus of form feedback: non-negotiated feedback, limited negotiation feedback, and extended negotiation feedback (cited in Fotos and Nassaji, 2007, pp. 123-24). Brown (2007, pp. 277-78) has provided taxonomy of six types of feedback that includes recast, clarification request, metalinguistic feedback, elicitation, explicit correction, and repetition. Among these six types of feedback, four types (i.e. clarification request, repetition, metalinguistic clues, and elicitation) are related to prompts. These four types of prompts have one crucial feature in common: "They withhold correct forms (and other signs of approval) and instead offer learners an opportunity to self-repair by generating their own modified response" (Lyster, 2004, p. 405).

Among different types of feedback one type that has had too many contradictory results on the learners' learning behavior is

"recast". Recast is "an implicit type of corrective feedback that reformulates or responds an ill-formed or incomplete utterance in an unobtrusive way" (Brown, 2007, p. 277). Lyster (1998, in Kaplan, 2002, p. 177) found that learners could not decipher whether recasts were signs of correction or approval. Lyster and Ranata (1997, in Kaplan, 2002, p. 177) also found that recasts did not have immediate effect on subsequent production. But Mackey and Philp (1998, cited in Kaplan, 2002, p. 177) considering delayed effects of feedback showed that recasts had positive effect on learners' production.

Comparing the effects of recasts and prompts, Ammar (2008) found superior effects for prompts over recasts in the acquisition of possessive determiners by Francophone learners of English as a second language. She also found that prompts were particularly effective for lower proficiency learners, whereas higher proficiency learners appeared to benefit similarly from both recasts and prompts. Lyster (2004) explains the four types of prompt used in French immersion classes mentioned before as follows:

1. *Clarification requests* are phrases such as "Pardon me" and "I don't understand" used to indicate that the student's message has either been misunderstood or ill formed. For example:

Student: *Et le coccinelle . . .* "And the (M) ladybug."

Teacher: *Pardon?* "Sorry?"

Student: *La coccinelle . . .* "The (F) ladybug."

2. *Repetitions* replicate the student's error verbatim, usually with rising intonation and stress to highlight the error. For example:

Student: *La chocolat . . .* "(F) Chocolate."

Teacher: *La chocolat?* "(F) Chocolate?"

Student: *Le chocolat.* "(M) Chocolate."

3. *Metalinguistic clues* provide comments, information, or questions related to the well-formedness of the student's utterance. For example:

Student: *Parce qu'elle cherche, euh, son, son carte.*

"Because she's looking for, um, her, her (M) card."

Teacher: *Pas son carte.* "Not her (M) card."

Student: *Euh, sa carte?* "Um, her (F) card?"

4. *Elicitation* entails direct questions such as "How do we say that in French?" or pauses that allow students to complete the teacher's utterance. For example:

Teacher: *Il vit où un animal domestique? Où est-ce que ça vit?*

“Where does a pet live? Where does it live?”  
 Student: *Dans un maison*. “In a (M) house.”  
 Teacher: *Dans . . . ? Attention*. “In . . . ? Careful.”  
 Student: *Dans une maison*. “In a (F) house.” (p. 405)  
 Brown (2007, p. 278) classifies responses to feedback into 3 types: (1) uptake, (2) repair, and (3) repetition.

### Research Questions

The study aims to test the effect of two types of feedback (i.e. explicit correction and elicitation), which are hypothesized to trigger different degrees of cognitive processing and awareness. The study’s research questions are thus formulated as follows:

1. Will the use of explicit and implicit feedback in class improve Iranian EFL students’ ability to accurately apply English tense markers?
2. Which type of feedback is more effective—explicit correction or implicit or-choice feedback?

### Research Hypotheses

To find answers to the above mentioned questions the following null hypotheses were formulated:

1. The use of explicit and implicit feedback in class will not improve Iranian EFL students’ ability to accurately apply English tense markers?
2. Explicit correction feedback is not more effective than or-choice feedback on Iranian EFL learners’ production of tense markers.

### Method

#### Participants

The participants in this study were thirty EFL students (all females) divided into two classes of fifteen. The students were all beginners. They received feedback on tense markers for 10 sessions. Each of the two groups received a particular type of feedback. One group, designated as *the explicit correction group* (Group A), received explicit correction following their errors in tense markers, whereas the other group designated as *the implicit or-choice group* (Group B), received or-choice feedback following errors in tense markers.

#### Procedures

Before instructional period, students received a fill-in-the-blank pretest (see Appendix) and after the instruction they received the same test as posttest (see Appendix). During the ten session instructional period, students in the two groups

received different types of feedback for 10 sessions.

Each session some sentences that included one of the simple present, simple past and present continuous tenses were extracted from one unit of the *Intro* book and their structures were practiced by the students. For examples of these structures see Table 1 below.

**Table 1.** Structures with simple present, simple past and present continuous tenses

Tense markers	Structures
Simple present	Wh-questions and statements with be
	Simple present Wh-questions and statements
	Yes/No and Wh-questions with do
	Questions with how
Simple past	Wh-questions and statements with be
	Simple past Wh-question and statements
	Past tense of be
Present continuous	Yes/No and Wh-questions and statements

Then, the researcher asked the students some oral questions to elicit the tense markers that were practiced on the same session or the sessions before. At the end of each session, the students were asked to make dialogues about their own experiences using the structures worked out on the same day and role play it in the class. During the whole class time students’ tense marker errors were immediately corrected in both groups. *The explicit correction group* (Group A) received *Type A* feedback (i.e. explicit error correction) and *the implicit or-choice group* (Group B) received *Type B* feedback (i.e. or-choice feedback). Examples of *Type A* and *Type B* feedbacks for some errors extracted from teacher-student interaction in class are presented in Table 2.

**Table 2.** *Type A* and *Type B* feedback with some examples extracted from teacher-student interaction.

Feedback Type	Example	
Type A (explicit correction)	1	T: What did you eat for breakfast? S: I eated tea, bread and butter for breakfast T: 'I eated tea', no. I drank tea & I ate bread.
	2	T: Where does your father work? S: My father work in bank. T: Uh uh! 'My father works' not 'my father work'.
	3	T: What are you doing? S: I stand up. T: 'I stand up', no. 'I'm standing'.
Type B (or-choice feedback)	1	T: What is Sara doing now? S: She listening you now. T: 'She listening to me' or 'she is listening to me'? S: She is listening.
	2	T: What did you do yesterday? S: I /ri:d/ my book. T: Yesterday, 'you /ri:d/ your book' or 'you /red/ your book'? S: I /red/.
	3	T: Did you go to bed late last night? S: No, I go to bed early. T: 'You go to be early' or 'you went to bed early'? S: I <i>went</i> to bed early.

### Target Features

To select the 3 target tense markers used in this study, the researcher analyzed the students' course book and selected the most frequent tense markers appearing in these materials. These tense

markers are presented in Table 3 with a few examples for each.

**Table 3.** Tense markers along with a few examples

Tense Markers	Examples
Simple present	They play football everyday. She <u>plays</u> the piano.
Simple past	I <u>walked</u> home yesterday. She went to the office. I read (pronounced /red/) an English book this morning.
Present continuous	I <u>am writing</u> a story. She <u>is smiling</u> . They <u>are listening</u> to the teacher.

### Treatment Materials

Ten topics were selected from among the students' course book material. The material, which included different topics such as greeting, daily schedule, clothes, families, and past experiences, provided the communicative context for the present study. For the purposes of this study, the topics for discussing and practicing different grammatical forms in the class were selected from *Intro*, an international communication book written by Richards (1997). The conversations and grammar exercises included in the book provided the context for drawing students' attention to the three tense markers which were the focus of the present study.

### Measures

The students' ability to use accurate tense markers was assessed by using a written 20-item fill-in-the-blank test for which the students were given 40 minutes to complete. The same test was used both as pretest and posttest. By comparing the pretest-posttest results with two t-tests the effectiveness of each type of feedback was measured. Then another t-test was used to compare the results of the two posttests to see which feedback was more effective.

Before the introduction of the study, in a pilot study conducted by the researcher, the 20-item teacher-made test was administered to a group of 15 students at the same level with the students in the researcher's classes to measure the reliability of the teacher-made test. By using the Kuder-Richardson formula, the researcher

measured the reliability of the test to be about 0.70. The measured reliability was within an acceptable range.

### Results and Discussion

The means and standard deviations of the pretests and posttests of the two groups are shown in Table 4 and Table 5. Also, figure 1 shows the means of the pretests and posttests for Group A (the explicit correction group) and Group B (the or-choice group).

**Table 4.** Pretest means (M) and standard deviations (SD) for the two groups

Group A <sup>1</sup> (n=15)		Group B <sup>2</sup> (n=15)	
M	SD	M	SD
6.8	2.8	6.4	2.9

<sup>1</sup> Group A: explicit correction group

<sup>2</sup> Group B: or-choice group

Table 4 shows that the means for the two groups on the pretest have no significant difference:  $t = 0.39$ ,  $p < 0.05$ . This confirms that the two groups were essentially equivalent before the instruction started. From Figure 1, it is quite evident that the posttest means are significantly higher than the pretest means:  $t = 6.92$ ,  $p < 0.05$  for Group A, and  $t = 4.30$ ,  $p < 0.05$  for Group B. This shows that both types of feedback (i.e. Group A feedback and Group B feedback) improved the students' accurate use of tense markers in English.

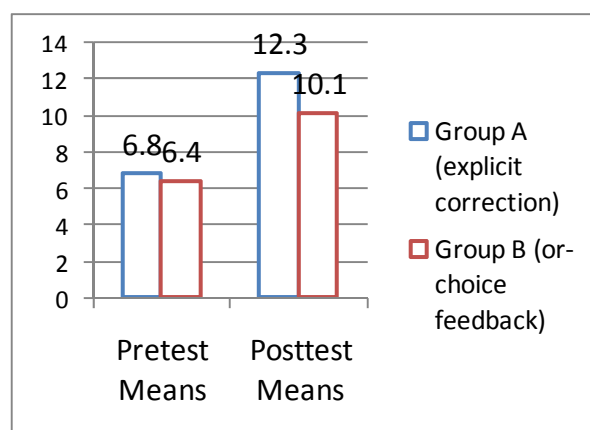
The means of the two groups on posttest, shown in Table 5 and also evident in Figure 1 shows that the *explicit correction* group outperformed the *or-choice* group on the posttest:  $t = 2.12$ ,  $p < 0.05$ .

**Table 5.** Posttest means (M) and standard deviations (SD) for the two groups

Group A <sup>1</sup> (n=15)		Group B <sup>2</sup> (n=15)	
M	SD	M	SD
12.3	1.4	10.1	1.8

<sup>1</sup> Group A: explicit correction group

<sup>2</sup> Group B: or-choice group



**Figure 1.** The pretest and posttest means for group A and group B.

The results of the t-tests rejected the first and second null hypotheses, thus confirming the effectiveness of the two types of feedback, and approving the superiority of the *Type A* feedback over *Type B* feedback on Iranian EFL learners' production.

The present study thus contributes to the arguments about the benefit of different types of corrective feedback by confirming that explicit correction is more beneficial compared to or-choice feedback to the beginner Iranian EFL learners in communicatively oriented classrooms. In terms of pedagogical implications, these findings may not be generalizable to other instructional contexts.

### 1.1 Appendix. Pretests and Posttests.

**Compete these conversations using the correct form of the following verb forms.**

See, spend, do/does, am/is/are, go, teach, work, look, love, live, stay

- A: What \_\_\_\_\_ you do?  
 B: I'm a student. I study English.  
 A: What \_\_\_\_\_ your mother do?  
 B: She is a teacher. She \_\_\_\_\_ English at a school in Iran.  
 A: Is your father a teacher, too?  
 B: No, he is a doctor. He \_\_\_\_\_ in a hospital.
- : Is your sister looking for a job?

- B: Yes, she is. She \_\_\_\_\_ in a restaurant now, but she \_\_\_\_\_ for a job in a theater. She \_\_\_\_\_ acting.
3. \_\_\_\_\_ A  
: \_\_\_\_\_ you \_\_\_\_\_ anything special over the weekend?  
B: Yes, I \_\_\_\_\_. I \_\_\_\_\_ shopping. I \_\_\_\_\_ all my money. Now I don't have any money.
4. \_\_\_\_\_ A  
: What \_\_\_\_\_ you \_\_\_\_\_ last night?  
B: I \_\_\_\_\_ the new Tom Cruise film. I \_\_\_\_\_ it!
5. \_\_\_\_\_ A  
: \_\_\_\_\_ you in France last summer?  
B: Yes, I \_\_\_\_\_.  
A: How long \_\_\_\_\_ you there?  
B: I \_\_\_\_\_ there for a month.

## References

- Ammar, A. (2008). Prompts and recasts: Differential effects on second language morphosyntax. *Language Teaching Research*, 12 (2), 183-210.
- Brown, H. D. (2007). *Principles of language learning and teaching* (5<sup>th</sup> ed.). New York: Pearson Education.
- Chaudron, C. (1998). *Second language classrooms: Research on teaching and learning*. Cambridge: Cambridge University Press.
- Doughty, C. J., & Long, M. H. (Eds.). (2003). *The handbook of second language acquisition*. New York: Blackwell.
- Ellis, R. (1990). *Instructed second language acquisition*. Oxford: Basil Blackwell.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Fotos, S., & Nassaji, H. (Eds.). (2007). *Form-focused instruction and teacher education: Studies in honor of Rod Ellis*. Oxford: Oxford University Press.
- Kaplan, R. B. (Ed.). (2002). *The Oxford handbook of applied linguistics*. Oxford: Oxford University Press.
- Lightbown, P.M., & Spada, N. (2006). *How languages are learned* (3<sup>rd</sup> ed.). Oxford: Oxford University Press.
- Lyster, R. (2004). Differential effects of prompts and recasts in Form-focused instruction. *SSLA*, 26, 399-432.
- Mitchell, R., & Myles, F. (2004). *Second language learning theories*. London: Arnold.
- Richards, J. C. (1997). *Intro*. Cambridge: Cambridge University Press.

## Information Sharing in a Pre-hospital Care Setting – A time related corpus based approach

Keiko Tsuchiya<sup>1</sup>, Miche Kawashima<sup>1</sup>, Tetsuo Yukioka<sup>1</sup>, Shoichi Ota<sup>1</sup>, Shiro Mishima<sup>1</sup>, Jun Oda<sup>1</sup>,  
Chikako Kawahara<sup>1</sup>, Masaki Onishi<sup>2</sup>, Ikushi Yoda<sup>2</sup>

<sup>1</sup>Department of Emergency and Critical Care Medicine, Tokyo Medical University,

<sup>2</sup>National Institute of Advanced Industrial Science and Technology

tsuchiya.keiko@gmail.com

### Abstract

A conversation analysis study in an emergency medical care setting in Japanese will be reported in this poster presentation. There are three areas defined in emergency medical care: pre-hospital care, in-hospital clinical care and emergency medical information system (Shimazaki, Asai, Aruga, Maekawa, Mashiko & Yukioka 2009). In this research, medical communication in a pre-hospital care setting has been examined. In pre-hospital care, emergency medical crews aim to conduct appropriate treatment to patients in casualties and transport them to hospital as quickly as possible. Activities of two emergency medical teams in a pre-hospital care training have been video-recorded and interactions between emergency crews and patient's family in relation to transition of the phases have been examined. From the results, placements and voices which inform future activities to a patient's family can be highlighted as one of the factors related to the quality of pre-hospital care although there are many limitations in this preliminary research with small data.

The training has conducted at Tokyo Medical University Hospital cooperating with Tokyo Fire Department. The research is a part of a medical communication research project founded by the Japan Science and Technology Agency.

### Keywords

Conversation analysis, Medical communication, Pragmatics, Corpus linguistics

### Introduction

In medical communication, a number of studies have been conducted in doctor-patient interactions (Candlin & Candlin 2003, Coulter, Entwistle & Gilbert 1999, Heritage & Maynard 2006, Pappas 1990) and nursing training (Brown, Crawford &

Carter 2006, Crawford 2009, Crawford, Brown & Nolan 1998). Research on a pre-hospital care setting, however, might be an area remaining uncovered in medical communication research. Based on theories discussed in the previous studies in medical communication, the current research attempts to analyse phase transitions of the pre-hospital emergency care in relation to placements and voices of information sharing between crews and patient's family with two methodologies; conversation analysis (Heritage 1997, Sacks 1992, Schegloff 2007) and corpus linguistics (Adolphs, Brown, Carter, Crawford & Sahota 2004, Carter & McCarthy 2006, McEnery, Xiao & Tono 2006, Sinclair 1996).

### 1 Theoretical frameworks

Heritage and Maynard (2006) has introduced three layers of medical communication analysis in their study on primary care.

- (1) The overall structure of the primary care visit.
  - (2) The sequence structures through which its particular component activities and tasks are realized
  - (3) The designs of the individual turns at talk that make up those sequences
- (Heritage & Maynard 2006: 13)

This current research can be categorised into the second layer of these three approaches. Based on Austin(1962) and Searle(1969)'s theory, the speech act of each utterance has been annotated for analysis and how sequences of speech acts orient realisation of activities in a pre-hospital setting has been investigated.

In a nurse-patient interaction study, Proctor, Morse and Khonsari (1995) have conducted a research on language of nurses in a trauma centre with

video-recorded data and recognise four pragmatic functions of 'comfort of talk';

- (1) *Holding on*: serves to praise, to let the patient know they can get through, to support, to instruct or distract the patient.  
Example) 'big girl', 'you're doing great'.
- (2) *Assessing*: concerns obtaining information about patient's condition.  
Example) 'How are you?' questions
- (3) *Informing*: involves warning the patient or explaining procedures.  
Example) 'it's gonna hurt', 'we'll be inserting a catheter'.
- (4) *Caring*: functions as some sort of reassurance, empathy or caring comment.  
Example) 'relax', 'OK sweetie', 'it does hurt, doesn't it?'

(Proctor, Morse & Khonsari 1996: 1673)

These classifications cannot be directly applied to this current study on interaction between medical crews and patient's family in a pre-hospital care setting, hence, considerations on 'informing' and 'caring' can be worth noting in activities in a pre-hospital care setting. As reported in Brown, Crawford and Carter (2006), 'habitus' and 'routines' in health care communication is also concerned.

## 2 Research Focus

In the experimental pre-hospital care training, the situation was set as the medical teams help a seventy-two years old male patient of CPA (cardiopulmonary arrest) becoming unconscious at home. The patient (indicated as P, hereafter) is a CPR(cardiopulmonary resuscitation) training manikin and a nurse acted as a daughter-in-law of the patient (referred as patient's family, PF, hereafter). Each team has three members and one of them is an emergency medical technician (EMT), who is licensed to perform emergency treatments, such as laryngeal mask airway insertion and IV line placement under direction of doctors. At the same time, EMT is assumed to control the activities as a leader of the team (indicated as L, hereafter). There are two other members (described as M1 and M2, hereafter) in the emergency medical teams. Under the same conditions, two emergency medical teams, namely Team1 and Team2, were requested to conduct pre-hospital care activities in a room where video and audio recordings were on.

Three phases can be recognised in pre-hospital care: 1) medical care setting creation, 2) emergency treatments and 3) pullout. In this research, focus has been placed on analysing placements and voices of sharing information between emergency medical crews and PF at a transition from the second phase, 'emergency treatments' to the third phase, 'pullout'.

Three aspects of use of language at a transition from treatments to pullout will be investigated;

- (1) Placement: When PF are informed,
- (2) Voice: From whom PF are informed, and
- (3) Act: In what way PF are informed.

From these three aspects, use of language in the two emergency medical teams in a pre-hospital care training have been analysed and compared.

## 3 Research Methods

The activities of two medical emergency teams in an experimental pre-hospital care setting have been video-recorded. Before the training, the general information about the project has been provided to the crews without telling them the purpose of the research precisely. Conversation analysis on the collected conversation data has been conducted using a multimedia annotation interface, Transana, and a qualitative analysis software, Atlas.ti. Placement, voice and act of information sharing at a transition from the treatment phase to the pullout phase have been analysed with time related corpus approach.

### 3.1 Transcription and data modification

The collected data has initially been transcribed by secretaries. The transcripts have been modified and time-stamped by a researcher with a time-based multimodal annotation software, Transana. Transana was developed by Chris Fassnacht at the University of Wisconsin. A video viewer, sounds wave bar, annotation field and project control panel are included as shown below.

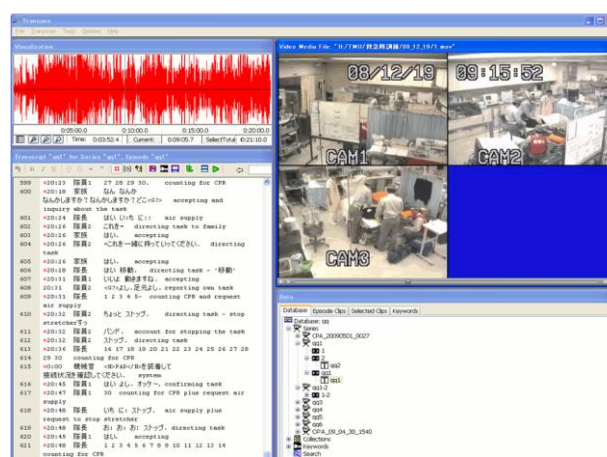


Figure 1: Transana

The time-stamped transcript has been exported into rich text files and imported into Atlas.ti. Atlas.ti is a qualitative analysis software developed by a German technology company called ATLAS.ti GmbH. With Atlas.ti, pieces of utterances have been

clipped as ‘quotations’ and ‘codings’ of speech acts, such as ‘summon to PF’, ‘inform to PF’ and ‘reporting own task’, have been added to each quotation.

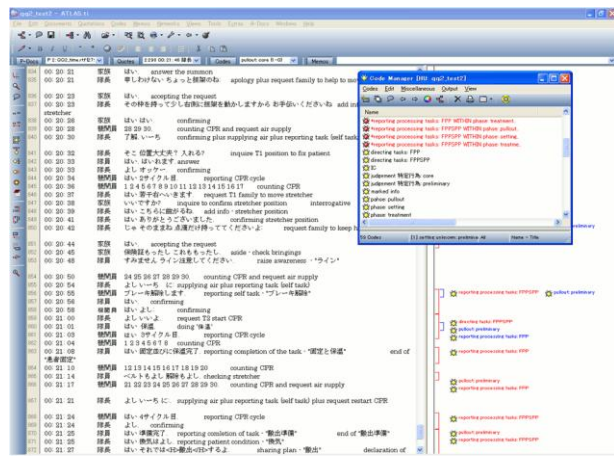


Figure 2: Atlas.ti

### 3.2 Lead time

Lead time is a concept introduced in Tsuchiya (2008)’s study on response tokens in conversation in academic tutorials. Lead time has been implemented as a scale to measure time distance from a floor taking point to use of particular response tokens in that study. The concept of lead time has been applied to this current study. The datum time has been fixed in the timing where the action of pullout has been observed.

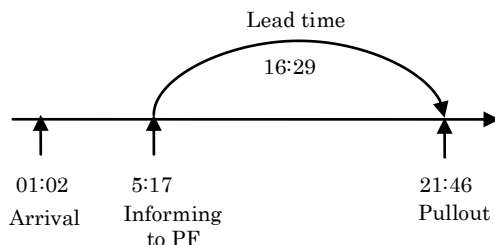


Figure 3: A sample of lead time

In the sample of lead time illustrated above, the datum time of pullout action is 21 minutes 36 seconds and lead time between the pullout action and the timing of the first information sharing is 16minutes 16 seconds. Time distance between the timing of informing pullout to PF and the pullout action have been examined.

### 3.3 Retrospective view

To have a retrospective view of the experimental pre-hospital care activities, comments and a simple assessment have been given by PF after the training.

## 4 Findings

Findings from the research will be reported in this section in reference to placement, voice and act of information sharing. From these three aspects, comparison between Team 1 and Team 2 has been made. PF’s assessments to each team will also be taken into consideration.

### 4.1 Placement of information sharing

The total time length from arrival to pullout of the pre-hospital activities in these two teams has been summarised in the table below. The timing when medical crews entered the room is defined as ‘arrival’ and ‘pullout’ is the timing when pullout action has been observed.

Table 1: Total length of pre-hospital care

	Arrival (MM:SS)	Pullout (MM:SS)	Total length (MM:SS)
Team 1	02:00	20:29	18:29
Team 2	01:02	21:46	20:44

The total time length of team 1 is 18:29 and team 2 is 20:44. Team 1 has conducted the course of pre-hospital activities in less time than team 2 by 2 minutes 15 seconds.

Table 2 shows that the lead time, a time distance between the first information sharing with PF and pullout action, in team 1 is 13 minutes 45 seconds while team 2’s lead time is 16 minutes 29 seconds.

Table 2: Lead time of pullout

	1 <sup>st</sup> Info share (MM:SS)	Pullout (MM:SS)	Lead time (MM:SS)
Team 1	06:44	20:29	13:45
Team 2	05:17	21:46	16:29

Note: ‘1<sup>st</sup> Info share’ means the timing when information sharing with PF about pullout has been observed at the first time.

The result can be interpreted that PF could have more time to get ready to go to hospital both mentally and physically in the case of team 2 although team 1 processed the whole activity faster than team 2, which can be another aspect to assess activities in pre-hospital care.

As illustrated in the figure below, three phases, namely ‘setting’, ‘treatment’ and ‘pullout’, have been recognised as primary components of pre-hospital care activities. These three phases can be divided into several subsidiary activities. In the setting phase, ‘setting and survey’ on a patient will be conducted. There are four activities in the treatment phase, such as ‘informed consent (IC, hereafter)’, ‘doctor’s instruction’, ‘Airway Mask’, ‘IV line’. In the EMT manual, IC to PF before proceeding medical treatments and

transporting a patient an emergency medical centre is instructed. L is also required phone a doctor at the dispatch centre to have a permission to conduct medical treatments.

The first information sharing with PF has been observed during the activity of ‘airway mask’ in team 1 while team 2 informed PF soon after ‘doctor’s instruction’ at the first time.

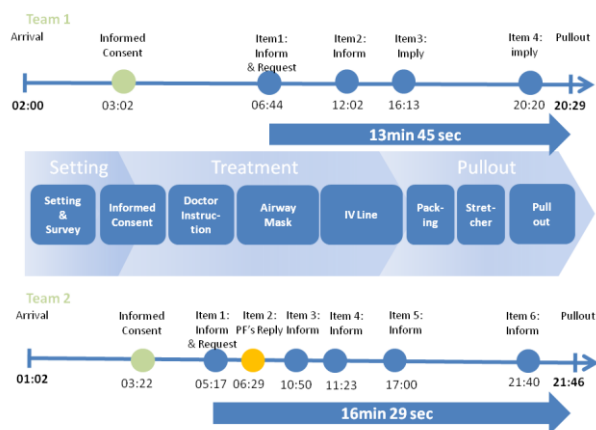


Figure 4: Placements of information sharing

In terms of the frequency of information sharing between crews and PF, team 1 has informed PF to go to hospital four times in total. The second information sharing in Team 1 occurs in the process of the ‘IV line’ insertion. The third item is allocated after the ‘IV line’ insertion and the last item has been placed soon before the action of pullout.

Team 2 has informed PF more frequently than team 1 and the total number of information sharing between crews and PF is six times. In addition to that, in the case of team 2, PF has replied back to the crews and informed that she is ready to go to hospital at Item 2. The other four items of information sharing in Team 2, have been placed at the boundaries of the activities, such as after the ‘Airway mask’, the beginning and the ending of ‘IV line’ and soon before the ‘pullout’ action.

## 4.2 Voice and act of information sharing

With transcripts described in extract 1 and extract 2 shown below, voice and act of information sharing for pullout in a pre-hospital care setting have been analysed in detail. The first column of the transcripts shows the timeline. Voice of the utterance appears in the second column. In the third column, original transcription in Japanese is described and translation in English is shown in the fifth column. The last column describes the act of each utterance.

## Extract 1: Transcript of Team 1

timeline	voice	original transcript in Japanese	translation in English	act
[Team1-Item 1: Inform]				
00: 06: 44	M1	ごめんなさいこれ、済んだら車内収容して行きますんで=	excuse me after this [Airway mask] we will transfer him to hospital so=	summon plus inform PF to go to hospital after treatment [airway mask]
00: 06: 47	PF	はい	yes	acknowledgement
00: 06: 48	M1	ご家族の方は火の元と戸締りをしてあと 家族の方でも	could you lock doors and windows and turn off cookers for fire safety and erm you can=	request PF to prepare for leaving to hospital
00: 06: 51	PF	はい	mhm	acknowledgement
00: 06: 51	M1	あの救急車の中でもPHSとか携帯は使えますから: それは: だいようぶですか=	use mobile in ambulance car so=: it will be okay	add info - availability of mobile use
[Team1-Item 2: Inform]				
00: 12: 02	M1	大変申しわけないです。今ね ちよとまだ=	erm sorry but erm we can't go now=	explain current process [IV line] to PF
00: 12: 04	PF	はい	yes	acknowledgement
00: 12: 04	M1	=あの、この穿刺 輸液が終わってから出たいと思いますので=	erm after the IV line we are leaving to hospital so=	inform PF to go to hospital after the treatment
00: 12: 07	PF	はいはい	yeah okay	confirming
[Team1-Item 3: Imply]				
00: 16: 13	M1	ご家族の方 すみません	excuse me, Ma'am	summon
00: 16: 18	M1	ちよとお借りした台なんですけど=	erm this table=	apology for wetting a table
00: 16: 19	PF	はい	yes	acknowledgement
00: 16: 20	M1	=濡れちゃったんですけど=	this table has got wet but=	apology for a wet table
00: 16: 21	PF	はい	yeah	acknowledgement
00: 16: 21	M1	=そのままちよと行っちゃいますのでごめんなさいね。すみません	=we gonna leave as it is so=: sorry for that	explain PF that they leave the table as it is plus imply they will go to hospital soon
00: 16: 22	PF	ああいいですいいです。はいはい	oh it's okay yeah yeah	accepting
[Team1-Item 4: Imply]				
00: 20: 20	M2	ご家族の方 申しわけございません	excuse me, Ma'am	summon
00: 20: 21	PF	はいはい	yes	answer the summon
00: 20: 22	M2	ちよと荷物(1つ)持っていたらいいですか?	could you help us to take this[infusion bag] with you?	request PF to take infusion bag with her
00: 20: 23	PF	なんなんかなんかしますか? なんかなんか? <G?> これを	wh= what= what can I do? What?	clarify the request
00: 20: 26	M2	はい	Where?	request PF to take infusion bag again
00: 20: 27	M2	=これを一緒に持って行ってください	this bag=	request PF to take infusion bag plus imply they are leaving
00: 20: 28	PF	はい	take this bag with you please.	accepting
[Team1-pullout]				
00: 20: 29	L	はい 移動	okay let's go	request M2 to move stretcher plus request M1 and PF to come along

Note: L = leader of the emergency medical team, PF= patient's family, M1= member one in the medical team, M2 = member two in the medical team, '<G?>' = inaudible sounds, '=' = unfinished sentence, ':' = prolongation of the immediately prior sound

## Extract 2: Transcript of Team 2

timeline	voice	original transcript in Japanese	translation in English	act
[Team2-Item 1: Inform and request]				
00: 05: 17	L	お 病院は救急センターというところをお願いしました	erm I asked an emergency medical centre to take a patient=	inform PF that L has requested emergency medical centre to accept a patient
00: 05: 21	PF	はい	okay	acknowledgement
00: 05: 22	L	で 私のほうで今 チューブとを準備してね	now I am going to do some treatments such as airway tube=	explain about treatments to be conducted
00: 05: 24	PF	はい	yeah	acknowledgement
00: 05: 25	L	え: 病院に少しでも早く行けるようにします	and trying to transport to hospital as fast as possible	give account for the treatment
00: 05: 28	PF	はい お願いします	yes please	request and accept L to conduct the treatments
00: 05: 30	L	え: え: ちよと病院に行く準備をしてくだい	so could you prepare for leaving to hospital please	request PF to prepare for leaving to hospital
[Team2-Item 2: PF's reply]				
00: 06: 29	PF	一応準備できましたので <G?> はい	erm I'm ready <G?> yeah	inform L that PF is ready to go to hospital
00: 06: 31	L	はい わかりました	right okay	confirming
[Team2-Item 3: Inform]				
00: 10: 43	L	家族の方 今ね チューブを固定しますからね	now I am fixing this tube yeah	explain current process [fixing airway tube] to PF
00: 10: 45	PF	はい	okay	acknowledgement
00: 10: 47	L	病院のほうは119番センターのほうで連絡をしてくれてますので	and the dispatch centre is trying to find a hospital to go	add info about hospital
00: 10: 49	PF	はいはい	yeah yeah	acknowledgement
00: 10: 50	L	私の処置が終わる次第	after the treatment	inform plan after the treatment
00: 10: 52	PF	はい	mhm	acknowledgement
00: 10: 52	L	すぐね	immediately	inform plan after the treatment
00: 10: 52	PF	はい	mhm	acknowledgement
00: 10: 53	L	病院のほうへ行きますからね	we are going to hospital okay	inform plan after the treatment
00: 10: 54	PF	はい	okay	confirming
[Team2-Item 4: Inform]				
00: 11: 15	L	じゃあね 今度 点滴しますからね	now I am going to insert an IV line	inform next process [IV line] to PF
00: 11: 17	PF	はい	okay	confirming
00: 11: 19	L	病院に行つてすぐお薬が使えるように ちよと必要な処置なのでね	it's needed for giving a drug after we get a hospital=	account for the treatment
00: 11: 22	PF	はい	mhm	confirming
00: 11: 23	L	え: 一回一人で実施して もし私の力では無理なようでしたらすぐ中断して病院へ行きますからね	erm I will try it now then if the condition doesn't change we gonna go to hospital immediately	inform next process and plan to go to hospital after treatment
00: 11: 30	PF	はい	okay	confirming
[Team2-Item 5: Inform]				
00: 17: 00	L	今点滴終わりましたからね	now IV line has fixed so	reporting completion of treatment [IV line] to PF
00: 17: 01	L	間もなく病院へ行くように	we gonna go to hospital soon	inform to go to hospital soon
00: 17: 03	PF	はい よろしくお願いします	oh yes please	acknowledgement and request to take patient to hospital
00: 17: 03	L	あの: 救急隊の担架に行きますよ	so we gonna move him to the stretcher	add info - move P to stretcher
00: 17: 05	PF	はい	okay	confirming
[Team2-Item 6: Inform]				
00: 21: 40	L	じゃ 家族の方 行きますよ	okay ma'am we gonna go	summon plus inform to go to hospital plus request PF to come
00: 21: 41	PF	はい <G?> はい	yeah <G?> okay	accept the request
[Team2-Pullout]				
00: 21: 46	L	はい 移動	okay let's go	request M1 to move stretcher plus request M2 and PF to come along

Note: L = leader of the emergency medical team, PF= patient's family, M1= member one in the medical team, M2 = member two in the medical team, '=' = unfinished sentence, ':' = prolongation of the immediately prior sound

In team 1, the addressers to inform PF vary. In item 1, 2 and 3 of team1, M1 has spoken to PF, and in item 4, M2 has requested PF to help them. At the pullout action, L has directed the crews and PF to move the stretcher in order to head for an emergency medical centre. Several speech act expressions have been recognised in information sharing in Team 1. In the first two items, PF has been informed directly. M1 has explained to PF that they are going to take P to the ambulance care after the treatment and requests PF to prepare for leaving to hospital with P. In the third and forth items in team 1, however, acts of 'apology' or 'request' have been observed. In the item 3, M1 apologises PF that they have got the table wet during the IV line insertion treatment but they are going to leave to hospital without cleaning. In the item 4, M2 has requested to take the infusion bag with her. These utterances imply that they are going to leave soon, whereas, PF needs to presume the actual timing of pullout from the utterances and activities that the crews are engaged in.

In contrast, in the case of team 2, PF has been informed by L all the time. The timing of pullout and how far they are from the point have been explained to PF from time to time.

#### 4.3 Comments and assessments from PF

After the experimental pre-hospital care activities, comments and a simple rating have been given to each team by PF. Three ranks, 'A', 'B' and 'C' are used for rating and A is the highest and C is the lowest mark. Team 1 is rated as 'C' and Team 2 is 'A'. Frequent use of technical terms in their explanation and ambiguity of roles in medical crews in Team 1 have been pointed out referring the fact that PF has been instructed by different crews in PF's comments.

### 5 Conclusion and implications

In this research, conversation analysis in a pre-hospital setting has been reported in reference to a transition from the treatment phase to the pullout phase. Sequences of acts for sharing information between medical crews and PF towards the pullout action have been described in reference to time placement and addressers. Both Team 1 and Team 2 have informed about the timing for going to hospital to PF several times in the course of activities, however, differences in placement, voice and act of information sharing have been observed. Team 1 has conducted the whole process faster than Team 2, however, PF has rated Team 2 higher than Team 1. From the findings and a retrospective view from PF, it can be said that use of language in activities seems to be involved in the issues of

quality of pre-hospital care.

Not only a patient in a causality but also the patient's family are facing an unusual situation in a pre-hospital care setting, which they might not expect when the day comes and they do not know how to deal with the unfamiliar 'reality'. For emergency medical crews, on the other hand, this 'unusual situation' is the field they work in daily life and there seems to be many 'habitual' behaviours, 'routines' in activities and terms and styles of communication shared only among the emergency medical professionals. Emergency crews, patients and their family, however, need to contribute to realise the activities in a pre-hospital care setting together although they have different perspectives on the reality they have faced. Use of language can play an important role in order to fill the gap between these two sides: medical professionals and ordinal people in emergency care. In a recent lecture on medical communication, Paul Crawford (2009) advises medical professionals with the phrase 'make them [patient] visible' and suggests that just eye contact and greetings to a patient can change the situation and improve patients' satisfaction to the treatments. In their study on doctor-patient communication in primary care consultations, Heritage and Stivers (1999) reports that use of 'online commentary', which is doctors' descriptions about patients' conditions during physical examinations, helps patients and patients' family beside to 'forecast' and 'reassure' the diagnostic evaluations coming after the examinations especially in the case of 'well visit'. Although there are many limitations in this small case study and further studies need to be conducted, three aspects in use of language highlighted in this research, namely timing, addresser and act of information sharing, can be one of the factors to improve the quality of pre-hospital care.

#### Acknowledgement

We thank the Tokyo Fire Department for their contribution to the research.

#### References

- Adolphs, S., Brown, B., Carter, R., Crawford, P. & Sahota, O. (2004). Applying Corpus Linguistics in a health care context. *Journal of Applied Linguistics* 1: 9-28.
- Austin, J. L. (1962). *How To Do Things With Words*. London: The Clarendon Press.
- Brown, B., Crawford, P. & Carter, R. (2006). *Evidence-based health communication*. Maidenhead: Open University Press.
- Candlin, C. N. & Candlin, S. (2003). Health care communication: a problematic site for applied linguistics research. *Annual Review*

- of *Applied Linguistics* **23**: 134-54.
- Carter, R. & McCarthy, M. (2006). *Cambridge Grammar of English: A Comprehensive Guide Spoken and Written English Grammar and Usage*. Cambridge: CUP.
- Coulter, A., Entwistle, V. & Gilbert, D. (1999). Sharing decisions with patients: is the information good enough? *British Medical Journal* **318**: 318-22.
- Crawford, P. (2009). *Nursing our hospitals*. Potcast lecture. University of Nottingham. [Accessed: 20 July 2009].
- Crawford, P., Brown, B. & Nolan, P. (1998). *Communicating Care: The Language of Nursing*. Cheltenham: Stanley Thornes.
- Heritage, J. (1997). Conversation analysis and institutional talk. In Silverman, D. (ed.) *Qualitative Research: Theory, Method and Practice*. London: Sage publications. 222-265.
- Heritage, J. & Maynard, D., W (2006). *Communication in Medical Care: Interaction between primary care physicians and patients*. Cambridge: CUP.
- Heritage, J. & Stivers, T. (1999). Online commentary in acute medical visits: a method of shaping patient expectations. *Social Science & Medicine* **49**: 1501-1517.
- McEnery, T., Xiao, R. & Tono, Y. (2006). *Corpus-Based Language Studies: an Advanced Resource Book*. London: Routledge.
- Pappas, G. (1990). Some implications for the study of the doctor-patient interaction. *Social Science & Medicine* **30**: 199-204.
- Proctor, A., Morse, J., M. & Khonsari, E. S. (1996). Sounds of Comfort in the Trauma Center: How Nurses Talk to Patients in Pain. *Social Science & Medicine* **42**: 1669-1690.
- Sacks, H. (1992). *Lectures on Conversation, Volume 1*. Massachusetts: Blackwell Publishers.
- Schegloff, E., A (2007). *Sequene Organization in Interaction: A Primer in Conversation Analysis, Volume I*. Cambridge.
- Searle, J. R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: CUP.
- Shimazaki, S., Asai, Y., Aruga, T., Maekawa, T., Mashiko, K. & Yukioka, T. (2009). *Hyojun kyukyu igaku(Standard Acute Medicine)*. Tokyo: Igaku-shoin.
- Sinclair, J. M. (1996). *Preliminary recommendations on corpus typology*. Online documents: <http://www.ilc.cnr.it/EAGLES/typology/no-del.html>. University of Birmingham. [Accessed: 8 January 2007].
- Tsuchiya, K. (2008). Toward a cultural-sensitive taxonomy of backchannels: How teachers and students use *erm* in conversation? In Lee, H. & N. Lee (eds.) *The 13th Conference of Pan-Pacific Association of Applied Linguistics*. Manoa: Pan-Pacific Association fo Applied Linguistics. 145-148.

# A Study of English Language Textbooks in Japan: With a focus on consistency

Koji Uenishi

Faculty of Foreign Studies, Setsunan University  
uenishi@ilc.setsunan.ac.jp

## Abstract

In this paper primary and junior and senior high school textbooks in Japan were analyzed and discussed. The authorized textbook *Eigo Noto* (English Workbook) by the Ministry of Education, Culture, Sports, Science, and Technology should be discussed for future English language teaching at primary school in Japan. This discussion will contribute to a future formal textbook which will be offered to all primary schools and be utilized there. Therefore, *Eigo Note* introduced for primary schools and popular textbooks used at junior and senior high schools were analyzed in terms of English expressions with wh-interrogative questions. Based upon the data results obtained, the consistency of the learning content from primary school to junior and senior high school was discussed. As a result, it was identified, on the whole, that the textbooks could not be said to be consistent or connected in terms of the English expressions including wh-questions that students learn at school.

## Keywords

English textbooks, wh-questions, primary school

## 1 Introduction

The Ministry of Education, Culture, Sports, Science and Technology (MEXT) announced last year that *Gaikokugo Katsudo* (foreign language activity) would be introduced as a lesson for the fifth and sixth grades at all primary schools in Japan in 2011.

Informally, English has already been introduced, in principle, once a week at some primary schools. Some or many teachers are very concerned about English teaching at primary school for various reasons: 1) curriculum, 2) teaching staff, 3) textbooks, and 4) connections with junior high school.

Regarding connections with junior high school, based upon the practices of Narita primary school, Niizato (2009) points out the following:

“Unfortunately, even at the base school of ‘foreign language activity’ at primary school they cannot afford to take into consideration the English education at junior high school in terms of time or spiritual burden. Primary school teachers are very busy thinking over teaching their own English lessons. Therefore, they will not be aware how what they teach is connected with the teaching contents at junior high school. ... On the other hand, I wonder if teachers at junior high school know about what primary school teachers teach—the goals of teaching English and the activities for achieving the goals. Probably the teachers hardly know about them.” [translated by the author]

According to Okuno (2009), it was already indicated that there was a wide range of problems on introducing English education at primary schools during the Meiji Era in Japan. Some of these are the relationship between English and

native tongue, needed teaching staff, and the integration with junior high school textbooks. However, similar problems remain to be solved at present.

With regard to textbooks, the MEXT proposed that teachers utilize the so-called *Eigo Note* (English Notebook) textbooks at primary schools. Unfortunately, the use of these textbooks during English lessons is left to individual primary schools. Therefore, "... some Boards of Education have opted not to have schools use the textbooks, while schools in other areas plan to use them selectively. Therefore, it will likely be difficult to iron out differences in teaching approaches among schools nationwide." (Torikai, 2009). The MEXT's announcement will result in a confusion about how to teach future junior high school students.

In this paper, the *Eigo Note* textbooks produced by the MEXT and primary school textbooks from Thailand were focused on and analyzed. Among the teaching contents at schools, the data results obtained from Japanese textbooks were compared with those from Thai textbooks. Regarding the research paper on the contents of English teaching at primary school, Nishina, Fujiwara, & Matsuoka (2009) discussed English expressions taught at primary schools, especially prepositions and key phrases including wh-interrogative questions. As a result, they found that chunks or key phrases including interrogative questions are vital for learners to enhance communication skills.

## 2 Research Objectives

In this paper, with a focus on phrases including wh-interrogative questions used in English teaching at primary schools in Japan, the formal textbooks offered by the MEXT were analyzed and discussed. Accordingly, below is the research objective.

How well are materials integrated between Japanese primary school and junior / senior

high school textbooks in terms of wh-interrogative questions and English phrases including them?

## 3 Research Methods

Each textbook was first digitized and then analyzed in terms of wh-interrogative phrases and expressions. The *Eigo Note* textbooks used in 'foreign language activity' at Japanese primary schools were analyzed, aiming at investigating English phrases and expressions, including wh-interrogative questions, that were taught to pupils. Below are the interrogative questions and English interrogative-phrases and -expressions analyzed in this research.

### *Eigo Note 1*

What do you want? / What do you ...? /

What would you like? / How many ...?

### *Eigo Note 2*

When (is your birthday)? / Where (is the station)? / What time do you (get up)?

Below are the textbooks analyzed in this paper.

- (1)*Eigo Note* textbooks at primary schools in Japan (5<sup>th</sup> Grade and 6<sup>th</sup> Grade)
- (2)*New Horizon* at junior high schools in Japan (Book 1 through Book 3)
- (3)*Prominence* at senior high schools in Japan (Book 4 and Book 5)

## 4 Results and Discussion

### 4.1 English Expressions Including Interrogatives

#### 4.1.1 Expressions Including 'What'

Based upon the aforementioned research question, the textbooks from primary to senior high school were analyzed and discussed. In that case, the analyses and implications about the data results were performed on every English phrase including each wh- interrogative question.

#### 4.1.1.1 ‘What do you want?’

The English expression “What do you want?” appears three times in Lesson 3 in the fifth-grade *Eigo Note* textbook. However, the same expression never appears in the junior and senior high school textbooks. In Unit 4 of Book 3 “What do you want to do on the weekend?” is used as the expression including “to verb.”

When it comes to the sentence pattern “What do you ...?”, the pattern is used twice in Book 1 in Japan. This phrase should be used more constantly in order for learners to memorize it longer as a viable expression. Besides, this pattern appears

once or twice in all textbooks, from Book 1 through Book 3, after the fifth-grade *Eigo Note* textbook. Given the situation of the appearance of the phrase “What do you ...?” it can be said that after a fashion the phrase is learnt rather continuously, though the frequency of its appearance is very infrequent. Even so, students will be able to memorize the phrase if it is taught using extra materials from time to time. In this kind of case, it depends on the teacher whether the students can learn the sentence pattern as a viable expression or not.

Table 1 ‘What do you ...?’ in Japanese Textbooks

What-Interrogative	EN 1	EN 2	Book 1	Book 2	Book 3	Book 4	Book 5
What do you want?	3	0	0	0	0	0	0
What do you want to...?	0	0	0	0	1	0	0
What do you ...?	3	0	2	1	2	2	2

#### 4.1.1.2 ‘What would you like?’

The English expression “What would you like?” including the word ‘would’ is discussed here. The word ‘would’ is often used in expressing oneself in a euphemistic and polite way. The expression is frequently seen in shopping and meal scenes on a daily basis.

It appears three times in the last lesson, Lesson 9 of the fifth-grade *Eigo Note* textbook. However, it is never used in junior high textbooks, Book 1 through Book 3 (Table 5). In this respect it can be said that the problem of the integration of learning materials between primary school and junior high school is identified. Furthermore, ‘would’ never appears in senior high school textbooks, Book 4 and

Book 5.

Regarding the pattern “What would you like to...?” including ‘to infinitive’, the sentence “What would you like to drink?” appears only once in Book 3. It seems that the frequency of the expression is not so low in daily life in English-speaking countries. Therefore, it is undeniable that there is a problem about the situation where the expression is hardly or never taught in this way. In addition, since it never appears in the senior high school textbook, *Prominence*, we need to discuss it, taking into consideration the connections between primary school and junior and senior high schools.

Table 5 ‘What would you like ...?’ in Japanese Textbooks

Useful Expressions	EN 1	EN 2	Book 1	Book 2	Book 3	Book 4	Book 5
What would you like?	3	0	0	0	0	0	0

What ...?	9	2	10	18	5	12	10
-----------	---	---	----	----	---	----	----

#### 4.1.1.3 ‘What time do you...?’

The English expression “What time do you...?” is used in daily life, and appears twice in Lesson 6 in the sixth-grade *Eigo Note* textbook. However, it appears in neither junior nor senior high school textbooks (Table 7).

The pattern “What time do you ...?” is another pattern that fails to never appear in junior and senior high school textbooks. As a result, students

who use the official textbooks never learn this phrase when asking questions about ‘time’ in daily life. As long as teachers don’t offer questions of “What time do you ...?” type in a supplementary way, the students never learn the sentence pattern in junior and senior high schools. When you think about learning the pattern, I think it is vital to teach it even at the beginning of junior high school.

Table 7 ‘What time ...?’ in Japanese Textbooks

What-Interrogative	EN 1	EN 2	Book 1	Book 2	Book 3	Book 4	Book 5
What time do you ...?	0	2	0	0	0	0	0
What time ...?	0	2	2	0	0	0	0
What time is it (...)?	0	0	2	0	0	0	0

#### 4.1.2 Expressions Including ‘When’

Interrogative questions including ‘when’ is discussed in this section. Although “When is your birthday?” appears twice in Lesson 3 of the sixth-grade *Eigo Note* textbook, it is never re-introduced in junior or senior high school textbooks. Besides, in the pattern of “When ...?” after it is treated twice in the sixth-grade *Eigo Note*

textbook, the pattern is used only once or twice in Books 1, 3, and 4. This means that the direct interrogative sentence using ‘when’ never appears in Book 2 and Book 5. Also, it appears around the end of the book in Unit 10 of Book 1 for the first time in junior high school. Then it is used again in Unit 11 of Book 1. However, in Unit 3 of Book 3 the expression “when ...?” appears after that.

Table 9 ‘When ...?’ in Japanese Textbooks

When-Interrogative	EN 1	EN 2	B1	B2	B3	B4	B5
When is your birthday?	0	2	0	0	0	0	0
When ...?	0	2	2	0	1	2	0

It is assumed that junior high school students in Japan need to have more chances to read and use interrogative questions including ‘when’. In the same manner as the expressions including ‘what’, it seems that the expressions using ‘when’ need to be treated in class by teachers.

#### 4.1.3 Expressions Including ‘Where’

The sentence pattern “Where is ...?” such as “Where is the station?” appears once in Lesson 5 of the sixth-grade *Eigo Note* textbook. After that, the pattern appears three times in Book 1. Unexpectedly, the sentence “Where’s my ticket?” appears in Unit 8 of Book 1 for the first time. In Book 2 the interrogative ‘where’ is treated once in Unit 1 using the sentence “Where were you?” However, after

that the interrogative question never appears in Book 3 through Book 5. This implies that as grades proceed from junior to senior high school, learning contents are more difficult and more abstract, and

indirect interrogative sentences and relative adverb ‘where’ are studied rather than the direct interrogative ‘where’.

Table 11 ‘Where ...?’ in Japanese Textbooks

Where-Interrogative	EN 1	EN 2	Book 1	Book 2	Book 3	Book 4	Book 5
Where is / Where's ...?	0	1	3	0	0	0	0
Where ...?	0	1	3	1	0	0	0

#### 4.1.4 Expressions Including ‘How’

The interrogative phrase “How many...?” appears twice in Lesson 3 of the fifth-grade *Eigo Note* textbook and twice in Book 1 in junior high school. The first appearance of this pattern in junior high school is in Unit 5 of Book 1; the expression using

‘how many’ is “How many CDs do you have?” In this respect it seems difficult to judge whether or not it is enough to learn the pattern like this. However, considering the retention of this pattern, it is assumed that it will be better to learn it again in Book 2, as well.

Table 13 ‘How ...?’ in Japanese Textbooks

How-Interrogative	EN 1	EN 2	Book 1	Book 2	Book 3	Book 4	Book 5
How many ...?	2	0	2	0	0	2	0
How are you?	1	0	0	0	0	0	0
How ...?	3	0	7	9	2	6	3

## 5 Conclusion

Prior to the introduction of ‘foreign language activity’ at primary school in 2011, the *Eigo Note* textbooks and junior and senior high school textbooks were compared and analyzed mainly from the perspective of the expressions including interrogatives. Since the *Eigo Note* textbooks were published only recently, these textbooks should have been produced thinking more about the English expressions (vocabulary) in junior and senior high school textbooks. However, this research identified that in fact the *Eigo Note* expressions failed to appear in a consistent or continuous way.

In the future, as the authorized textbooks are produced and English teaching has progressed in other EFL countries such as Thailand and China,

formal textbooks used by all primary schools should be published in Japan, as well. In that case, learning contents need to be discussed more by taking into account logical language development from primary school to junior and senior high school. Moreover, junior high school daily-life, high-frequency phrases and expressions, including vocabulary should be adopted. This will make learning contents closer to real situations and lead to more natural, interactive English learning. As a result, the students’ interest and enthusiasm about English learning should be enhanced. In that case, Nishida’s (2009) research focusing on prepositions and interrogatives in workbooks for primary school provides a foundation for materials development.

Henceforth, reflecting on the above research results, textbooks need to be analyzed and

discussed. In order to produce better primary school textbooks, further research should be done, such as, comparing Japanese English textbooks with other

Asian EFL materials. For that purpose, analyses and discussion of primary school textbooks from other EFL countries should be promoted from now on.

## 6 References and appendices

### 6.1 References

- Okuno, H. (2009). Shougakkou gakushuu shidou youryou ni okeru gaikokugo katsudou no hihanteki kousatsu — shougakkou eigokyouiku no rekishi to genjou wo fumaete — [Critical discussion about 'foreign language activity' in The Course of Study at Primary School—Based on history and situations of English language education at primary school] *Studies in English Language Teaching*, 32, 53—62.
- Kanashige, N., & Naoyama, Y. (2008). *Shougakkou shin gakushuu shidouyouryou no tenkai gaikokugo katsudou hen* [Development of New Course of Study at primary school: foreign language activity] Meiji Toshosha.
- Matsukawa, R. & Oshita, K. (2007). *Shougakkou eigo to chuugakkou eigo wo musubu: eigokyouiku ni okeru shouchuu renkei*. [Connecting primary school English with junior high school English: Cooperation between primary and junior high schools in English language teaching]. Kouryousha Publishing Company.
- Ministry of Education, Science and Culture, 2009. *Eigo Noto 1* [English Workbook 1]. Tokyo: Kyoiku Shuppan.
- \_\_\_\_\_. 2009. *Eigo Noto 2* [English Workbook 2]. Tokyo: Kyoiku Shuppan.
- Niizato, M. (2008). Chuugakkou no renkei wo dousurunoka — naritashi no shouchuu ikkan no torikumikara [How should we connect English teaching at primary school with English teaching at junior high school—Regarding consistent English language teaching between primary and junior high schools at Narita City] *English Language Magazine*, September, 17-19.
- Nishina, Y., Fujiwara, Y., & Matsuoka, Y. (2009). Shougakkou no tamenon juuyou fureizu oyobi teikeihyougen no chuushutsu: zenchishi wh-gimonshi no baai. [Extraction of key phrases and fixed expressions for English teaching at primary school: with a focus on prepositions and wh-interrogative expressions]. JACET KANSAI JOURNAL No.11 14—32.
- Torikai, K. (2009). Seeds of English learning: Primary classes aiming for children to enjoy learning. The Daily Yomiuri, May 10, p.10.
- The Daily Yomiuri. (2009). Half of middle school teachers unaware of grade-school English. May 10, p.10.
- Suzuki, Y. (n.d.). 「tai ni okeru shougakkou eigo kyouiku no genjou to kadai」 [Present situation and problems at primary school in Thailand.] Retrieved May 19, 2009, from [www.mext.go.jp/b\\_menu/shingi/chukyo/chukyo3/015/siryo/05120501/s004\\_4.pdf](http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo3/015/siryo/05120501/s004_4.pdf)

# Acquisition of New Phonemes in Learner Speech: A Study of Korean English

Eun sil Bang, Oh Seon Kweon, and Jong-mi Kim

Department of English Language and Literature, Kangwon National University  
bespretty@hanmail.net, rnjsdhtjs89@naver.com, \*kimjm@kangwon.ac.kr

## Abstract

The acquisition order of new phonemes in adults' language learning were investigated using utterances previously shown to be problematic in loanword adaptation. A large scale of English speech data was collected from 124 students of English as a foreign language (EFL) at five ranks of proficiency. To contrast, two different groups of native speech data were collected from 12 native speakers of American English and 11 native speakers of Korean. The learner speech data was compared with the L1 Korean pronunciation of loanwords in Korean context and L1 English pronunciation of phones in loanwords. We examined the acoustic features that are relevant to the phonemes in question.

Learners acquired faster a new L2 phone that was allophonic in L1 for free variation than for complementary distribution. A new L2 vowel was faster acquired when phonetically dissimilar in terms of tongue height than reduced energy with schwa quality. Orthographic representation played a role in that a complex mapping of letter to sound caused difficulty in learning. The error rate drops drastically after instruction to indicate that an explicit instruction is an effective method to enhance the phonological acquisition in language learning.

## Keywords

English loanwords into Korean, Korean learner speech of English, Acquisition order of sounds, Learner speech, Loanword phonology

## 1 Introduction

Is there a principled order of acquisition of new phones in learner speech of a second language? Long-standing observation in literature is that new L2 phones are most often assimilated to their 'phonetically similar' L1 counterparts (Best, 1995; Flege, 1995; and, Strange, Akahane, Kubo, Trent, Nishi, & Jenkins, 1998), of which some non-native contrasts are difficult to discriminate when they are

strongly identified with a single native category (Fledge, 1995). Fledge (1995) predicts that a phonetic difference that distinguishes contrasting foreign sound but does not also distinguish contrasting native sounds will be poorly detected and reproduced. The study has been proposed to account for the changes that occur in both the perception and production of the first and second language of an L2 learner.

This observation, however, is not specific enough in the similarity metric, and the degree of difficulty to predict the acquisition of non-native contrasts. The similarity metric has been set in Strange et al. (1998) using the traditional features of "tongue-height" and "backness." They report that American English vowels were most often assimilated to their "phonetically-similar" Japanese counterparts as defined by these distinctive features. Likewise, studies on consonants in L2 acquisition also used such traditional distinctive features such as voicing, lateralness, and fricative/stop (Best, McRoberts, & Goodell, 2001). Then the question remains on what features determine the acquisition order of L2 phones whose phonetic differences distinguish contrasting foreign sound but do not also distinguish contrasting native sounds?

To answer these questions, we investigate the assimilation process of some English phones by Korean learners whose phonetic features are defined by continuity ([f]/[p]), lateralness ([l]/[r]), and vowel height ([i]/[e] / [a]). We also investigate the vowel reduction feature ([ə]) and the orthographic composition (mapping of one sound to two letters). We then compare the learners' L2 speech data with their L1 loanwords as well as the L1 target speech. Specifically, we collect the speech data of the English loanwords that are produced in learner speech and analyzed the process of acquisition of new sounds.

The answer to the question of segmental acquisition process may depend on the developmental aspects of learner speech, because studies have shown that learner speech progressively assimilate to the value of the target

speech. We investigate the acquisition order by using both a cross-sectional and longitudinal methods. The longitudinal method is to test cognitive and affective factors in adult language learning -- such as the role of attention and the importance of motivation in learning, as Schmidt (1995) claims, “In order to acquire phonology, one must attend to phonology” (Schmidt, 1995a, 17).

## 2 Method

In order to accomplish our objective, we administered the phonetic experiment that comprises four sub-experiments on native and learner speech in cross-sectional and longitudinal studies. The four sub-experiments are on: 1) Korean native speech of English loanwords in Korean context, 2) English native speech of the English loanwords in English context, 3) Korean learner speech of the English loanwords in English context by a cross-sectional study, 4) Korean learner speech of the English loanwords in English context by a longitudinal study. Each sub-experiment used different sets of speech materials, participants, and the data acquisition steps.

### 2.1 Speech materials

Three types of speech materials and the recording lists were used: 1) recorded L1 Korean speech materials of English loanwords to Korean in Korean context, 2) recorded L1 and L2 English speech materials of English loanwords to Korean in English context, 3) listening test materials in English for Korean learners to serve as co-variance reference to the development in speech production.

The listening test materials were the listening component of *Test Of English for International Communication* (TOEIC). Two TOEIC listening tests were used as a standardized test to recognize the improvement of listening comprehension before and after instruction.

For recorded speech materials, the native and learner participants read words in isolation, and those embedded in English or Korean contexts. Table 1 is the recording list, in which the segments with different sound quality in Korean and English to contrast are underlined. The Korean context is romanized according to the *Official Romanization System* of Korean.

Table 1: Recording list for missing target phonemes in L2 English words, spoken by L1 Korean speakers

(a) Words in isolation	(b) Words in English context	(c) Words in Korean context
1. <u>a</u> dd [ɒ] 2. <u>a</u> dditional [ɒ] 3. <u>l</u> ight [r] 4. <u>p</u> late [r] 5. <u>d</u> e <u>f</u> ine [p] 6. atmos <u>ph</u> ere [p]	1. I <u>a</u> dded one more figure to my paper. [ɒ] 2. I need an <u>a</u> dditional remark. [ɒ] 3. Where do you <u>l</u> ive in Korea? [r] 4. Where are they <u>p</u> laying tennis? [r] 5. We are going to define the principles of <u>ph</u> ysics.[p] 6. I can't <u>f</u> ind the place. [p]	1. Oneul <u>pe</u> ulei eottaesseyo? (How was the <u>play</u> today?) 2. Je bijaneun imi <u>ik</u> seupaieo dwaesseoyo. (My visa has already <u>expired</u> .) 3. Eoje “ <u>ro</u> seuteu” bwanni? (Did you see the “ <u>lost</u> ” yesterday?) 4. <u>pe</u> oseuteu keulaessue anjatseo. (I sat in the <u>first</u> class.)

The underlined phonemes of English words in Table 1 do not have phonemic status in Korean phoneme inventory. For instance, the English phoneme [ɪ] is absent in Korean vowel inventory, whose front unrounded vowels comprise [i], [e], and [æ] (Lee, 1999). The expected assimilation forms to their “phonetically-similar” Korean counterparts are transcribed by IPA symbols in square brackets.

The recording lists in columns (a) and (c) were used to obtain L1 Korean speech materials of English loanwords in Korean context, while the lists in columns (a) and (b) were to obtain L1 and L2 English speech materials of English words in English context.

### 2.2 Participants

Research participants consisted of 11 Korean native speakers who read English loanwords in Korean context, 12 native speakers of American English and 124 Korean learners of English who read the English loanwords to Korean in English context.

All native speakers spoke the General American English, while learners the Standard South Korean as their native language. The learner's English proficiency level varied with respect to the scale of five ranks in accordance to TOEIC score. We then putatively used the scaling of score intervals in *TOEIC Can-Do Guide* (2000) as

co-variance units for our analysis.<sup>1</sup> This follows the developmental effects that the more proficient the speaker is, the more target-like the perception of the target sound is for the listener.

### 2.3 Data acquisition procedure

The data acquisition procedure varied for all four sub-experiments. For the first sub-experiment, where we acquired Korean native speech of English loanwords, we used the following three stages of eliciting the reading speech. At the first stage, the Korean native speakers were given the English word list in English alphabet, and asked to transcribe the words into Korean alphabet. The transcription was done for the words in isolation. At the next stage, the speakers were given the Korean sentence list where the English loanwords are embedded hidden in empty blanks. The speakers are then asked to fill in the blanks by copying down their own previously transcribed English loanwords in Korean alphabet. At the last stage, the speakers were asked to read the completed Korean scripts for recording the words in isolation and the words embedded in sentences..

For the second sub-experiment, where we acquired English native speech of the English loanwords in English context, the native speakers of American English underwent only one stage of the recording of the production stimuli.

For the third sub-experiment, where we acquired Korean learner speech of the English loanwords in English context by a cross-sectional study, the learners took listening and production tests after a brief listen and repeat practice of the recording stimuli, on a very first week of any class instruction to avoid experimental intervention.

For the fourth sub-experiment, where we acquired Korean learner speech of the English loanwords in English context by a longitudinal study, the procedure consisted of three stages: 1) pre-listening and production tests after a brief listen

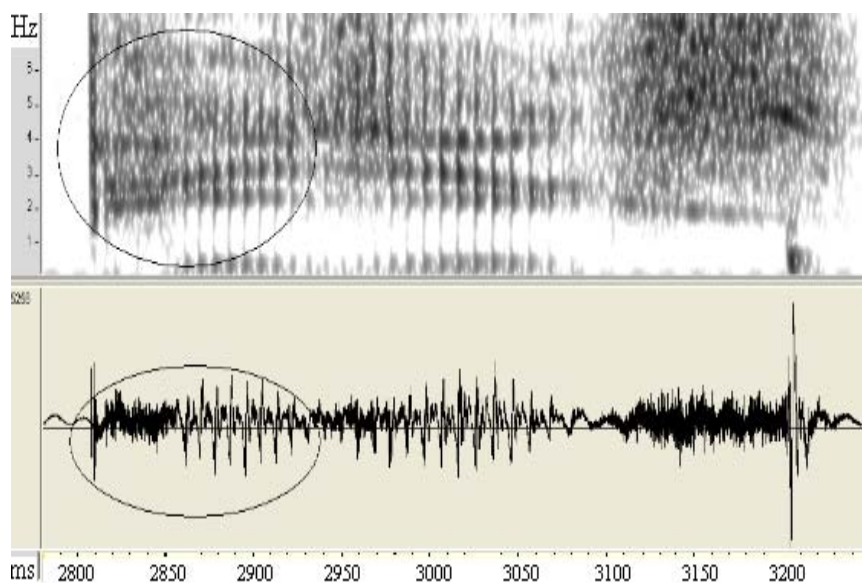


Figure 1: Illustration of an incorrectly produced phone [p] for the target phone [f] in L2 English word *physics*, spoken by a male Korean learner of English. The word was embedded in a sentence, “We are going to define the principles of **physics**”. The L2 phone demonstrates several acoustic cues of [p]: 1) silence (stop gap), 2) release burst indicated by a strong vertical spike, and 3) a long VOT of 50 ms as short frication noise (scattered marks after the release) before vowel formants begin.

and repeat practice of the recording stimuli, 2) pronunciation instruction in class for 7 weeks that included many different aspects of pronunciation including, but not focused on, avoiding the Korean l-alternation and applying the English de-aspiration rules, and 3) post-listening and production tests on the same data after 7 weeks, and again on the different data after next 7 weeks.

### 2.4 Analysis

We measured the acoustic features that are relevant to the contrasting phonetic features for the segments as outlined in Table 1. We illustrate an erroneous production of learner speech in Figure 1.

Figure 1 shows a spectrogram of an erroneous production of English by a Korean learner at level 3 (mid-intermediate proficiency) before instruction. The English phone [f] in the word “physics” in this learner speech is assimilated to the “phonetically similar” counterpart [p] in Korean phoneme inventory. The stop quality of the given phone is detected by several acoustic cues: 1) silence (stop gap), 2) release burst indicated by a strong vertical spike, and 3) a long VOT of 50 ms as short frication noise (scattered marks after the release).

For analysis, we compared the learner speech with the native speech. For less clear cases of acoustic features, we used the additional means of the perceptual judgment by researchers and native speakers of English. We counted only the expected

<sup>1</sup> This edition of the *Can-do Guide* fits to our learner speech data, as we used the previous version of TOEIC listening test using the General American English. The next edition was published on 2009 to reflect the recent modification in TOEIC to contain different English accents around the world.

errors that are shown in the loanwords for the different L2 phones in question.

### 3 Results

A total of 2085 word data were acquired from the recorded corpus of native and learner speech in word-level and sentence level production by 124 Korean learners of English and 11 Korean and 12 English native speakers.

The developmental results for the acquisition order of new L2 phones are shown for word level production in Figure 2 and sentence level production in Figure 3.

For the speech data of words in isolation, the acquisition order of new L2 phones are shown in

the cross-sectional study (2a) and in the longitudinal study (2b): [f] in define, [l] in plate, [ɹ] in add, [ɹ] in light, [ɹ] in additional, and [p] in atmosphere.

Learners acquired faster a new L2 phone that was allophonic in L1 for free variation than for complementary distribution, since the discrimination of [f] from a single native category [f/p] was earlier learned before the discrimination of an onset [l] from a single native category of a liquid phoneme [l/r] which is phonotactically constrained of an onset [r] and coda [l]. A new L2 vowel was faster acquired when phonetically dissimilar in terms of tongue height than reduced energy with schwa quality, since the discrimination

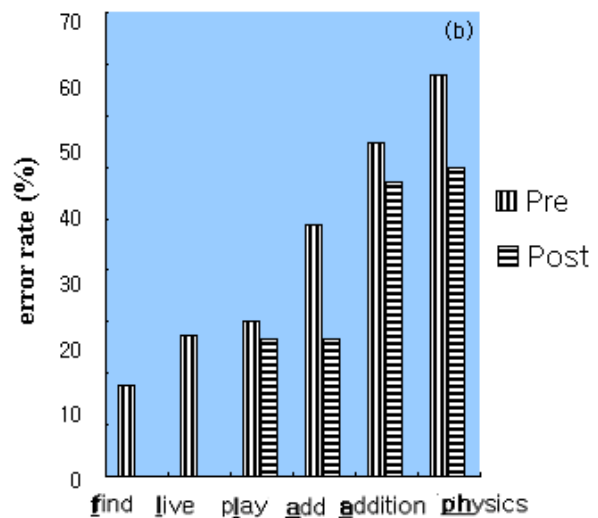
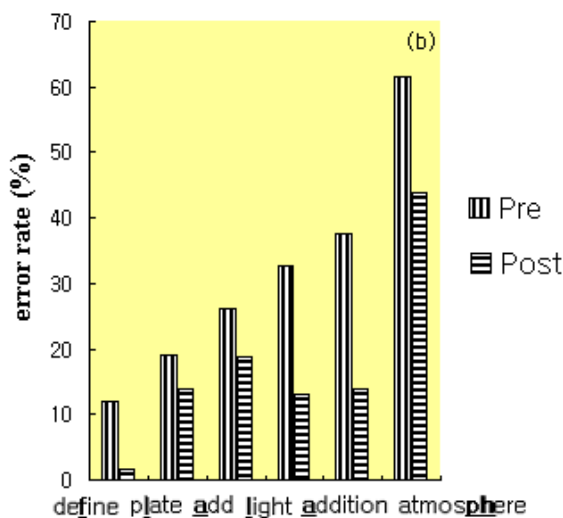
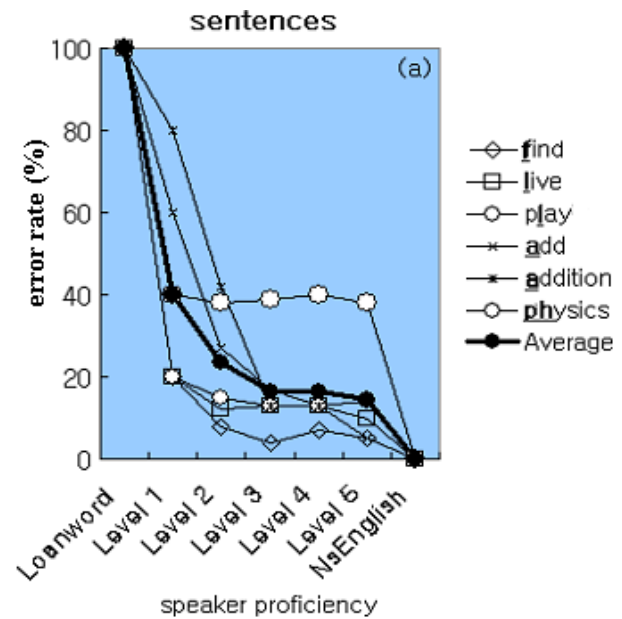
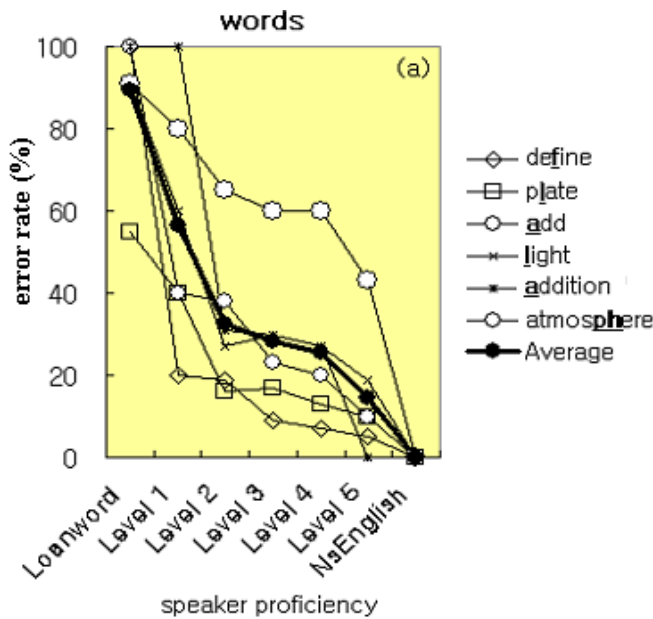


Figure 2: Results of the acquisition order for non-native phones in word-level production from a cross-sectional study (a) and a longitudinal study (b).

Figure 3: Results of the acquisition order for non-native phones in sentence-level production from a cross-sectional study (a) and a longitudinal study (b).

of the L2 phone [ɹ] with a single native category [ɹ]. [ɹ] is faster acquired before the discrimination of the [ɹ] with a single native category [ɹ]. Orthographic representation played a role in that a complex mapping of letter to sound caused difficulty in learning, since “f” in *define* is faster learned than “ph” in *atmosphere*. Both of the developmental results in Figures (2a) and (2b) support that the error rate decreases as the learner proficiency increases.

Similar results are obtained for the speech data of words in sentence context. Figure 3(a) is the result for the cross-sectional study, and the acquisition order of new L2 phones are: [f] in *find*, [l] in *live*, [l] in *playing*, [ɹ] in *add*, [ɹ] in *additional*, and [f] in *physics*.

Learners acquired faster a new L2 phone that was allophonic in L1 for free variation than for complementary distribution, since the discrimination of [f] from a single native category [f/p] was earlier learned before the discrimination of an onset [l] from a single native category of a liquid phoneme [l/r] which is phonotactically constrained of an onset [r] and coda [l]. A new L2 vowel was faster acquired when phonetically dissimilar in terms of tongue height than reduced energy with schwa quality, since the discrimination of the L2 phone [ɹ] with a single native category [ɹ] is faster acquired before the discrimination of the [ɹ] with a single native category [ɹ]. Orthographic representation played a role in that a complex mapping of letter to sound caused difficulty in learning, since “f” in *find* is faster

learned than “ph” in *physics*. Both of the developmental results in cross-sectional study (3a) and in longitudinal study (3b) support that the error rate decreases as the learner proficiency increases.

To see whether the developmental improvement takes effect in a new data, we tested the learners after 7 weeks gap, and the result is shown in Figure 4. The post-test results show the decrease of error rates regardless when the test is with the same data and the different data. The post-test was taken after 7 weeks of interval; with the same data after the first 7 weeks, and with the new data after the next 7 weeks of instruction. The results are similar both for the words in isolation and the words embedded in sentences.

#### 4 Discussion

The results confirm some old findings while they present new findings in this study. Confirmation to old findings comprises the following: 1) As Fledge (1995) predicts, a phonetic difference that distinguishes contrasting foreign sound but does not also distinguish contrasting native sounds are poorly detected and reproduced, as in our data of learner speech. 2) Among these data with phonetic similarity, the traditional features did distinguish their developmental order, as shown in Strange et al. (1998) and Best et al. (2001). 3) The longitudinal method supports that cognitive and affective factors in adult language learning involve the role of attention to the given new speech sounds, as Schmidt (1995) claims.

New findings in this study are the following.

First, acquisition order of new L2 English phones by Korean learners are found in that 1) a new L2 phone that was allophonic in L1 for free variation earlier than for complementary distribution, 2) a new L2 vowel phonetically dissimilar in terms of tongue height earlier than reduced energy with schwa quality; 3) Orthographically one-to-one simple mapping of letter to sound earlier learned than complex mapping. Secondly, the cross-sectional study shows the significant developmental transition of the acquisition of new L2 phones. Thirdly, learner speech is far more distant from loanwords, and much closer to target phonology. In other words, learner speech is far more faithful to target phonology than loanword phonology. Fourthly,

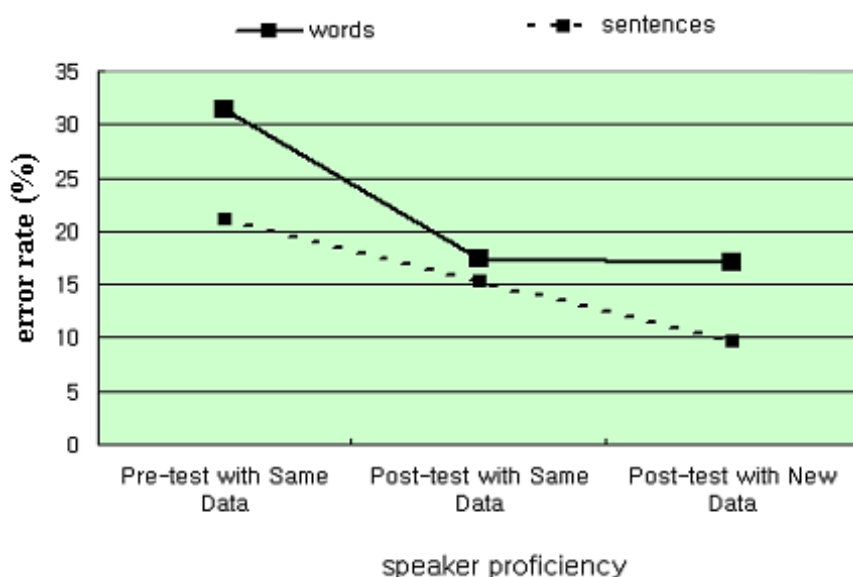


Figure 4: The post-test results show the decrease of incorrect production of non-native phones, regardless whether the test is on the same data, on the different data, for the words spoken in isolation, or for the words spoken in sentences.

another difference between learner speech and loanword phonology is that learner speech develops toward target phonology, while loanword phonology is rather constant. Fifthly, needless to mention is an additional fact that loanwords conform to native phonotactics, whereas learner speech does not.

Currently, the discussions in this study are based on the average values of a large number of data. We plan to do a more sophisticated analysis with statistical significance.

## Acknowledgements

This work was supported in part by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (KRF-2006-321-A00933).

## References

- Best, C. T. (1995). A direct realist view of cross-language speech perception. *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, ed. by W. Strange, 171-203. Timonium, MD: York Press.
- Best, C. T., McRoberts, G. W., & Goodell, E. (2001). American listeners' perception of non-native consonant contrasts varying in perceptual assimilation to English phonology. *Journal of the Acoustical Society of America*, 109, 775-794.
- ETS. (2000). *TOEIC Can-Do Guide: Linking TOEIC Scores to Activities Performed Using English*. (2000). Princeton, NJ: Author. [TOEIC\_Can\_Do.pdf, www.ets.org]
- Flege, J. E. (1995). Second language speech learning: Theory, Findings, and problems. *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, ed. by W. Strange, 233-277. Timonium, MD: York Press.
- Lee, H-B. (1999). Korean. *Handbook of the International Phonetic Association*, 120-123. London: The Association.
- Official Loanword Transcription Regulations*. (1995). [The current edition in official use]. Ministry of Culture and Tourism, Republic of Korea. [www.korean.go.kr].
- Official Romanization system*. (2000). [The current edition in official use]. Ministry of Culture and Tourism, Republic of Korea. [www.korean.go.kr].
- Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in foreign language learning. *Attention and awareness in foreign language learning* ed. by R. Schmidt, 1-63. Honolulu, HI: Univ. of Hawai'i. Second Language Teaching and Curriculum Center.
- Strange, W., Akahane, R., Kubo, R., Trent, S., Nishi, K., & Jenkins, J. (1998). Perceptual assimilation of American English Vowels by Japanese listeners. *Journal of Phonetics*, 26, 311-344.

# College Teacher Training for Tertiary Students Towards Becoming Effective English Teachers

Byung-Bin Im

Kongju National University

rainbow@kongju.ac.kr

## Abstract

This study relates to the field of teaching and second language acquisition (hereinafter SLA) with respect to college teacher training for students of English Education at a tertiary level. As such, the purpose of this study is to establish a number of recommendations (4 in total) of key areas deemed by the author as necessary to improve the effectiveness of a college teacher training course, and expand upon them in detail. This article focuses upon recommendations towards the development of effective teacher training in relation to the following: (1) the need for Orientation training in SLA, ESP and target culture, and (2) teacher awareness of their attitudes. Additional concepts include (3) building trust between the supervisor and trainee teachers, and (4) further related ongoing education for teacher-trainers emphasized during in-service training to assist such prospective teachers in developing their model classes and professional development whilst also the promotion of lifelong education for teachers in general. Such recommendations have been expanded upon in detail throughout this article and based on supported research conducted in respect of a number of teacher education studies.

## Keywords

second language acquisition, teacher training, teacher awareness, lifelong education

## Introduction

SLA is a complex learning environment requiring teachers to be attentive and reflective in order to best serve learners. Attitudes and awareness of teachers are integral factors towards their successful adoption of various teaching strategies and students' (recipients) own learning styles. Awareness and, in many cases where it is called for, changes in own attitudes and beliefs, enable teachers to better adapt to the students' various learning styles and lead to adopting appropriate and non-biased beliefs and attitudes applicable towards a particular cultural, gender-specific, class-based and socio-economic

group.

Continuing education, independent observation, cultural training, mutual trust and reflection are valuable aids and tools for dedicated ESL teachers to achieve their best. Following the conclusion of that section, this writer has also provided a detailed summary of the subsequent resulting implications, conclusions and theories as they relate to the need for certain elements to be implemented as part of the university teacher training specifically for Korean English teachers (KETs) who are Non-native Speaking language teachers (hereinafter referred to as NNS) within the field of English Education.

## 1 Theoretical framework

### 1.1 Orientation training in SLA, ESP, and Target culture

Zinkgraf (2003) noted in her study of prospective teacher trainees (university level education) that the learners' critical analysis was too subjective and that they need to maintain an open and flexible mind that is crucial for a teacher and researcher. The critical awareness and assessment of language was important to the learner and assisted their deeper understanding of the target culture and hidden ideologies. Thus a deeper understanding enables lessons to be culturally content based (Celce-Murcia, 2001). Kumaravadivelu (1993) criticized the education of language teachers, specifically those engaged in communicative language teaching (CLT), as being responsible for failing to provide adequate tools when such educators are unable to create genuinely communicative classrooms (Kumaravadivelu, 1993). Warschauer (2000) noted that teachers in Korea can create opportunities for communication based on values, cultural norms and needs of learners rather than focus on the imperialistic dictates of the Native English speaking countries. Educators should take into account the social, economic, cultural and linguistic consequences of the spread of English.

## **1.2 Teacher awareness and reflection-Attitudes and beliefs**

There is a further crucial awareness factor to be considered in which educators are able to identify and critically assess the impact of their own attitudes and beliefs within the classroom. Altan (1997) proposes that teachers' beliefs are a valuable complement to traditional approaches to the study of teaching. Such values and beliefs are the basis for much of the teachers' decision-making and classroom action and as a result constitute the culture of teaching.

As such, unanswered questions remain in terms of the following:

- (i) Can study refrain teachers from reaching interpretations and biased decision-making?
- (ii) How does an interactive internet site replace interpersonal guidance and counseling?
- (iii) Can we rely on teachers to make those changes in awareness themselves without any follow-up evidence to reaffirm their commitment to making use of their education?

Reflection is emphasized as a tool for aiding teachers develop context-specific, personal theories of L2 teaching (McDonough, 2006). As the complexity in the classroom changes, teachers must be prepared to adjust both their teaching styles and strategies in order to respond effectively. Risks will yield positive results which this writer would seek to offer as advice to prospective teachers (Brown, 2000). Negative results are construed as positive given the learning experience they provide.

## **1.3 Concept of building trust: supervisor/trainee and teacher/student**

Trust is that one aspect of teaching that cannot be automatically granted or assumed. Rather, it is achieved through a number of ways. Chamberlin (2000) observed the trustworthiness of supervisors towards teaching candidates and suggested that supervisors who use non-verbal behaviors of affiliation will be perceived as more trustworthy than supervisors who use non-verbal behaviors of dominance. As supervision is a fundamental part of pre-service and in-service professional development for teachers of ESL, attention has been given to creating meaningful learning experiences for teachers that encourage reflection. Reflection should be encouraged (reflective supervision) and non-verbal displays of affiliation were more effective towards building trust than dominance cues. Listening, office arrangement and amount of space likewise had an effect upon the level of trust. It is suggested that for new

teachers to reflect upon their teaching, supervisory practices that value learning to teach as a continual process is needed.

## **1.4 In-service training of prospective teachers in model classes, classroom management: Continuing education**

Standards to which pre-service KETs receive training and ultimately work towards in relation to their organization of material, activity focus, and language goals are essential. Such standardized teaching standards as adopted throughout South Korea for all subjects in general may not always be directly applicable for SLA teaching though obviously still required. Due to the complexity of SLA, it is a matter of determining the best and most practical standards that can be adhered to and having knowledge and practice of such standards implemented into the university training programs.

It has become accepted that language teachers should conduct research on their classrooms and beliefs about teaching. However, less emphasis has been placed upon the educators themselves of those teachers. A brief survey conducted by the author of 20 linguists revealed that none of them had conducted research upon their teaching, testing and students' learning. The reason for this reveals that linguists have traditionally applied analysis and reflection rather than research upon their students' learning of KAL. A further reason that the author suggests, is that due to the high level of qualification needed (usually doctorate level) for KAL educators, their own professional education has not prepared or encouraged them to engage in such research. Sert (2006) prescribes to the theory that increased awareness of autonomous learning and its benefits will enhance the teachers' own capacity for self-governing and contribute towards higher achievement and motivation.

## **2 Issues**

There has been an ongoing issue of whether they are the most appropriate to teach a foreign language in comparison to the native speaker. Whilst the native speaker's linguistic competence represents an advantage, language proficiency alone does not compensate for other advantages that a NNS teacher is able to utilize. Nemtchinova (2005) notes the NNS teacher has certain linguistic and pedagogical resources that include: conscious knowledge of grammar, second language learning experience, good learner role model, rapport and empathy with the target audience. The attitudes and awareness of a teacher are integral factors to their successful implementation of various teaching strategies. Continuing education, independent

observation and reflection are valuable aids and tools for effective teachers.

There has to be general agreement as to the importance of teachers' awareness and beliefs and a need to raise awareness in order to assist teachers in becoming aware of own beliefs and attitudes as they reflect in their teaching styles and understanding of the classrooms and students. There are many courses available for further education for teachers though few have been designed to raise awareness sufficiently whereby they are encouraged to self-assess themselves on a level that might impose upon their own of beliefs and attitudes. It may be enough to raise awareness of the impact that their held beliefs might affect teaching and their response to learners particularly in the L2 field of varying cultures and student needs. There are varied opinions as to how that education might take place. Though, it is clearly stated, how, when and where it might be applicable. Obviously, there are a number of schools of thought on the subject due to the continuing complexity of SLA. How do we achieve this education?

I would recommend the following 4 key areas to improve teacher training:

- (i) Provision of orientation studies that reflect cultural awareness
- (ii) Courses which provide a framework to systematically observe, evaluate and reflect upon L2 teaching whilst encouraging critical analysis of awareness and self-inquiry.
- (iii) Teacher educators to establish a trusting relationship with teacher trainees.
- (iv) Adoptions of standards that are determined to be the most practical standards that can be adhered to specifically related to the complex subject of SLA.

### **3 Implications for second language teaching**

There is a need to determine what the teachers' as educators suggest would be appropriate forms of learning. Questionnaires and regulated interviews are required on a large scale to address irregularities in the L2 education system. Teachers need to have more voice as to the curriculum's format for which they are the ultimate distributors. Allowing teachers to address their own concerns as a whole encourages reflection and open debate over a whole

spectrum of issues.

In respect of teacher trainees in Korea there should be an increase of critical language awareness into the curriculum. Spoken fluency is not always the main target. There is an increasing role for the teaching of critical and analytical skills by ELTs to be addressed within the courses for teacher trainees. The greatest advantage a learner possesses is their ability to learn autonomously and that autonomy is often passed through the teacher. Teaching the students how to learn and become autonomous learners is more often going to be successfully achieved if the teachers themselves have that capacity for self learning. Teachers are by nature, individuals, and that individuality reflects negatively and positively upon the learners. Teachers themselves can't be expected to always self-regulate their own actions and awareness and take on responsibility. Assistance is required from the students and parents as to what is successful in their own learning and when is the teaching most beneficial. Student evaluation and parent groups will serve to highlight problems encountered in the teaching field. L2 teachers need to individually assessed for their understanding of the practices of L2 teaching and their level of knowledge of the students and their respective cultures (or at least respect for foreign cultures). Rigorous enforcement of teaching values and autonomous capability needs to be addressed at the L2 level.

SLA is an expanding field met by varying degrees of success and failure. Specific L2 teacher education in awareness training could provide an understanding of the difficulties faced by L2 learners, techniques for self-reflection and analysis of bias, awareness of learning styles, and subsequent importance of the linkage between self beliefs and motivation, and that of the learner. From a socio-cultural perspective, cognitive development in teachers progresses from an external, socially-mediated activity to internalization (Johnson, & Golombek, 2003). Education in tools such as reflective journals aid teacher trainees in understanding problems in the classroom and subsequent relationships with students.

### **4 Conclusion**

As teachers and professionals, L2 educators must come to the realization that attitudes and awareness that complement a person's personality are integral factors towards the successful implementation of various teaching strategies and capability for appreciating the students' learning styles, culture and individuality. Factors including reflection, supervisory trust, and ongoing research by KAL educators were discussed at length. 4 Key

recommendations by this author were put forward and these issues highlighted the need for further training in respect of awareness and reflective practices, professional research and an undertaking to build the aspect of trust between the trainees and education supervisors. As a result of the discussion above, this writer has further recommended that awareness training be implemented during university or college and continue during the teachers' careers including random classroom observation, reflective journals, and self-assessment. Finally, in order to gain a greater understanding of their students, L2 teachers in particular should be required to undergo cultural studies and possess L2 experience.

## References

- Altan, M.Z. (1997). *The culture of the English language teacher in an EFL context*. (Report No. FL024-533). Orlando, FL: Paper presented at the Annual meeting of teachers of English to speakers of other languages. (ERIC Document Reproduction Service No. ED438706).
- Brown, H.D. (2000). *Principles of language learning and teaching*. (4th ed). White Plains, New York: Pearson Education, Longman.
- Celce-Murcia, M. (Ed). (2001). *Teaching English as a second or foreign language*. (2nd ed). United States of America: Heinle & Heinle.
- Chamberlin, C.R. (2000). TESL degree candidates' perceptions of trust in supervisors. *TESOL Quarterly*, 34(4), 653-672.
- Johnson, K.E., & Golombek, P.R. (2003). Seeing Teacher Learning. *TESOL Quarterly*, 37(4), 729-737.
- Kumaravadivelu, B. (1993). Maximizing learning potential in the communicative classroom. *ELT Journal*, 47(1), 12-21.
- McDonough, K. (2006). Action research and the professional development of graduate teaching assistants. *The Modern Language Journal*, 90(1), 33-47.
- Nemtchinoya, E. (2005). Host teachers' evaluations of nonnative-English-speaking teacher trainees - A perspective from the classroom. *TESOL Quarterly*. 39(2), 235-261.
- Sert, N. (2006). EFL student teachers' learning autonomy. *ASIAN EFL Journal*, 8(2), Article 8.
- Warschauer, M. (2000). The changing global economy and the future of English teaching. *TESOL Quarterly*, 34(3), 511-535.

# Effect of Using English Storybooks for Improving Korean English Learners' Reading and Writing Skills and Their Learning Attitudinal Factors

Mun-koo Kang

Department of English Education Kongju National University

kangmunkoo@hanmail.net

## Abstract

This paper explores the effect of using children's storybooks through a variety of activities for improving Korean primary school EFL students' reading and writing skills and their learning interest, confidence, and motivation. It focuses on why and how children's storybooks serve as an essential context for increasing students' reading and writing skills and their learning attitudinal factors. One hundred and twenty Korean 5th grade English students were randomly chosen and participated in this study. Data was gathered through pre- and post-tests in terms of reading and writing skills and pre- and post- questionnaires with regard to their interest, confidence, and motivation. The analysis revealed that: (1) the students who received the lessons through children's English storybooks performed better on reading and writing skill post-tests than those who learned only through the English course book; and (2) the students in the experimental group showed more significant increase in their interest, confidence, and motivation than those in the control group. The results of this study conform to the educational values of using English storybooks for Korean primary school EFL students. The pedagogical implications for EFL educators and teachers are described.

## Keywords

English storybook use/reading and writing skills/ learning attitudinal factors

## 1 Introduction

The introduction of children's storybooks into the children's language learning settings has been supported by a number of researchers and teachers, making an effort to produce wholistic literacy instructional approaches (Gilles, 1991; Huck, 1989; Hunt, 1991; Peterson & Edes, 1990; Rha, 2002; Smith, 1998). Stories in the EFL primary school classroom are enjoyable. Stories can expose children to other cultures, and they can be used to

introduce and recycle language (Brewster, Ellis, & Girard, 2002; Slatterly & Willis 2001). Moreover, stories help children develop a sense of the target language (Wright, 1995). Rha (2002) states, through introducing literature into the classrooms, children can learn about the fundamental nature of language in terms of the importance of imagination and keep children's natural curiosity and love of learning.

In order to increase the learners' language skills of the target language, learners need more exposure to it. Storybooks are common language learning instruments that provide and stimulate a very large amount of that language. As Krashen (1985) notes, reading stories to children with limited skills and knowledge of the target language on a frequent basis can be beneficial because they help students acquire language, figure out the grammar, and make predictions. Literature is a powerful form of communication and is considered as one of the most efficient tools teachers can use to help students of all ages learn a new language.

The Korean 7th National English Curriculum is in line with the emphasis of CLT in that the target language needs to be taught in the learning environment where L2 learners are able to immerse themselves and to be exposed to the target culture. Storybooks are recommended to be employed to in every 4th class of each chapter of the primary English textbooks in the 7th National English Curriculum. However, the contents and language seem not to be natural, authentic, and sufficient. The 3rd and 4th graders might have fun with simple dialogue provided in their course books, but as they become the 5th and 6th graders, they often get bored and lose motivation. Hill (2001) points out how surprising it is that modern ELT course books and syllabi do not make more use of stories which are a fundamental and enjoyable aspect of the target language.

The purpose of this study is to investigate the effect of using English storybooks for improving Korean 5th-grade English learners' reading and

writing skills and their English learning attitudinal factors including motivation, confidence, and interest. The research hypotheses are as follows:

1. Korean 5th-grade English learners whose teacher uses English storybooks through a variety of activities will exhibit higher improvement on their English reading and writing skills than those whose teacher does not use English storybooks.
2. Korean 5th-grade English learners whose teacher uses English storybooks through a variety of activities will exhibit higher improvement on their English learning attitudinal factors including interest, confidence, and motivation than those whose teacher does not use English storybooks.

## 2 Method

This present study used an experimental design in order to investigate if the difference in teaching methods has a positive effect on the development of Korean 5<sup>th</sup> grade English learners' reading and writing skills and their interests, confidence, and motivation. This chapter presents the methodology used in the present study. The information on the participants, instruments and research design and procedures is provided.

### 2.1 Participants

The participants were 120 students who attended one of the primary schools located in Gong-ju, Chungnam. Among six classes at the fifth grade level, four classes were chosen to participate, depending on the teacher's willingness to participate in the research study. Initially, there were 130 students attending at the four classes of the 5th grade, but the researcher needed to discard the data of 10 students since 5 students were absent on one of the testing days and 5 students did not respond to the test questions with usable answers as data.

These four classes were randomly divided into two groups, experimental and control. The two classes with 30 students each were employed in the experimental group, the other two with 30 students each in the control one. Before the treatment, the researcher administered the first questionnaire regarding the students' English learning experience during the first week of the 7 month's research study session.

### 2.2 Instruments

The following instruments to gather data were used in this study

- 1) The English language proficiency tests regarding reading, and writing skills administered as the pre- and post-tests to explore if the difference in teaching methods influences the improvement of students' reading and writing ability. Each test consisted of 10 to 13 questions, selected from English proficiency test provided by KICE (Korea Institute for Curriculum and Evaluation) and scored 2 points for each correct item.

- 2) The two kinds of questionnaires were developed. The first one was developed to find out what English learning experiences the participants have had regarding how long they have studied English and for what purposes they are studying English. It was administered for the first week of the research session. This questionnaire was only administered once before the treatment. The second questionnaires, pre- and post-, were developed to investigate if the difference in teaching methods influences the participants' interest, confidence, and motivation toward English learning.

### 2.3 Procedures

This research was conducted for about 7 months from the first week of May, 2006 to the second week of December, 2006, excluding 4 weeks from the last week of July till the third week of August, during which the school was on the summer vacation.

The teacher for the experimental classes was an experimenter and the other for the control classes was a volunteer teacher who has majored in English education at Kongju National University and has been teaching English part-time at several colleges for many years. The experimenter met the volunteer teacher before and after the classes to have mini-workshops about classes.

The Korean English curriculum requires 2 hours of English lesson each week for the fifth grade students. Therefore, during the research session, every week the experimental group was given one hour of their regular English lesson based on the English course book and one hour of the experimental English lesson based on the research design, that is a story-based syllabus. The control group was given the two hours of their regular lesson, based on their course book each week. The researcher utilized English storybooks for two different purposes: one for in-class reading, the other for the out- class listening and reading activity and student's self-selected reading. Storybooks were displayed in the classroom book corner.

The list of the in-class reading storybooks were

*Meg and Mog, The Snowman, Pat the Cat, The Very Hungry Caterpillar, and The Turnip*. The researcher chose those books since the stories included the appropriate level of language, that is vocabulary or structures, any language features such as interesting rhyme, onomatopoeia, intonation, visual support, and repetition. The storybooks also came with the cassette recording.

### 3 RESULTS

#### 3.1 The Results of the Questionnaire on the Students' Previous English Learning Experience

Based on the students' responses from the first questionnaire before the treatment, about the half of the students answered that they have been studying English for about 1 year, about 20 percents of the students for about 2 years, about 10 percents of the students for 3 years, and another 10 percents of the students for 4 years. With regard to the students' primary purpose of studying English, about 40 percents of the students answered that it's because English is an international language. About 20 percents of the students answered, they study English in order to communicate with native speakers of English. From the first questionnaire it was revealed, even though the majority of the participants have experienced English study for only one year, they have recognized the importance of knowing English and being able to communicate in English these days.

#### 3.2 The Relationship between the Teaching Method and the Students' Reading and Writing Skills

The first hypothesis of this study was to determine whether students' reading and writing skills would be improved depending on the teaching methods or not. In order to collect the data, the language proficiency pre- and post-tests were administered and the mean differences of each test between two groups were compared based on ANOVA (Analysis of Variance) and ANCOVA (Analysis of Covariance).

##### 3.2.1 The Results of Reading Pre-test

Before the treatment, in order to find out if the experimental and control groups were the same in terms of their reading skill, the two groups were administered reading pre-test. ANOVA results of the students' reading pre-test are shown in Table 2. The mean of the experimental group was 16.22, and the mean of the control group was 19.12, which means the students' reading proficiency in

the control group was higher than the one in the experimental group. ANOVA revealed that there was significant difference in mean scores between two groups ( $F=6.48$ ,  $p=.012$ ). This means that before the treatment the two groups were not at the same level in terms of reading proficiency. Therefore, for the analysis of the students' reading post-test, it is necessary to use ANCOVA which is used to analyze the mean differences of the reading post-test by keeping the mean differences of two groups' reading pre-test under control.

Table 2: ANOVA Results of the Reading Proficiency Pre-test

	Group	N	M	SD	F	P
Pre-test	Control	60	19.12	6.00	6.48	.012
	Experimental	60	16.22	6.48		

##### 3.2.2 The Results of Reading Post-test

After the treatment, in order to find out if there is any difference between the experimental and control groups in terms of their reading skill, the two groups were administered reading post-test and the ANCOVA results of the two groups' reading post-test are presented in Table 3. The mean of the experimental group was 20.33, and the mean of the control group was 19.73. ANCOVA revealed that there was significant difference in mean scores between two groups ( $F=33.56$   $p=.000$ ). Therefore, after the treatment the experimental groups' reading proficiency were improved higher than the control ones and this means that the treatment has a positive effect on the improvement of the students' reading proficiency.

Table 3: ANCOVA Results of the Reading Proficiency Post-test

	Group	N	M	SD	F	P
Post-test	Control	60	19.73	5.64	33.56	0.00
	Experimental	60	20.33	4.65		

##### 3.2.3 The Result of Writing Pre- and Post-tests

As seen in Table 4, the mean of the experimental group of the writing pre-test was 20.30 and the mean of the control group was 19.87. ANOVA revealed that there was no significant difference in mean scores between two groups ( $F=.24$ ,  $p=.624$ ). Therefore, before the treatment the two groups were at the same level in terms of writing proficiency.

After the treatment, the mean of the experimental group was 22.73 and the mean of

the control group was 20.80. ANOVA revealed that there was significant difference in mean scores between two groups ( $F=17.03$   $p=.000$ ). Therefore, after the treatment the experimental groups' writing proficiency were improved higher than the control ones and this means that the treatment has a positive effect on the improvement of the students' writing proficiency.

Table 4: ANOVA Results of Writing proficiency Pre- and Post-tests

Categories	Test	Control			Experimental			F	P
		N	M	SD	N	M	SD		
Writing	Pre-	60	19.87	5.49	60	20.80	4.42	.24	.624
	Post	60	20.80	4.42	60	22.73	3.29	17.03	.000

### 3.3 Footnotes The Relationship between the Teaching Method and the Students' Attitudinal Factors

The second hypothesis of this study was developed to determine whether students' attitudinal factors including interest, and confidence in English learning, and also motivation to learn English would be improved depending on the teaching methods or not. In order to collect the data, the pre- and post-questionnaires were administered and the mean differences of each questionnaire between two groups were compared using ANOVA

#### 3.3.1 The Results of Students' Interest in Learning English

As seen in Table 5, the mean of the experimental group before the treatment was 3.74 and the mean of the control group was 3.76. ANOVA revealed that there was no significant difference in mean scores between two groups ( $F=.01$ ,  $p=.911$ ). Therefore, before the treatment the two groups were at the same level in terms of their interest in English learning.

However, the mean of the experimental group was 4.32, and the mean of the control group was 3.78. ANOVA revealed that there was significant difference in mean scores between two groups ( $F=45.13$   $p=.000$ ). Therefore, after the treatment the experimental groups' interest in English learning was improved higher than the control one and this means that the

treatment has a positive effect on the improvement of the students' interest in English learning.

#### 3.3.2 The Result of Students' Confidence in English Learning

As seen in Table 5, the mean of the experimental group was 3.13 and the mean of the control group was 3.19. ANOVA revealed that there was no significant difference in mean scores between two groups ( $F=.11$ ,  $p=.743$ ). Therefore, before the treatment the two groups were at the same level in terms of their confidence in English learning.

After the treatment, the mean of the experimental group was 3.62 and the mean of the control group was 3.24. ANOVA revealed that there was significant difference in mean scores between two groups ( $F=21.61$   $p=.000$ ). Therefore, after the treatment the experimental groups' confidence in English learning was improved higher than the control one and this means that the treatment has a positive effect on the improvement of the students' confidence in English learning.

#### 3.3.3 The Result of Students' Motivation to Learn English

As seen in Table 5, the mean of the experimental group was 3.73 and the mean of the control group was 3.69. ANOVA revealed that there was no significant difference in mean scores between two groups ( $F=.07$ ,  $p=.799$ ). Therefore, before the treatment the two groups were at the same level in terms of their motivation to learn English.

Table 5: ANOVA Results of the Students' Attitudinal Factors

Categories	Questionnaire	Control			Experimental			F	P
		N	M	SD	N	M	SD		
Interest	Pre-	60	3.76	.82	60	3.74	.81	.01	.911
	Post-	60	3.78	.87	60	4.32	.55	45.13	.000
Confidence	Pre-	60	3.19	.96	60	3.13	1.03	.11	.743
	Post-	60	3.24	.98	60	3.62	.78	21.61	.000
Motivation	Pre-	60	3.69	.74	60	3.73	.87	.07	.799
	Post-	60	3.71	.83	60	4.19	.58	24.03	.000

After the treatment, the mean of the experimental group was 4.19 and the mean of the control group was 3.71. ANOVA revealed that there was significant difference in mean scores between two groups ( $F=24.03$ ,  $p=.000$ ). Therefore, after the treatment the experimental groups' motivation to learn English was improved higher than the control one and this means that the treatment has a positive

effect on the improvement of the students' motivation to learn English.

#### 4 CONCLUSION AND SUGGESTION

So far, this study aimed to examine the effect of using English children's storybooks on the improvement of Korean primary school EFL learners' reading and writing skills and their interest and confidence in English learning, and their motivation to learn English and yielded the following major findings:

First, there was a significant difference in the degree of participants' reading and writing skill improvement between the experimental group and the control one. The students who received the lessons through English children's storybooks performed better on their reading and writing skill posttests than those who learned through only the English course book.

Second, there was a significant difference in the degree of the increase in students' interest, confidence, and motivation between the two groups. The results of the students who received the experimental treatment showed more significant increase in their interest, confidence, and motivation than those who got the control one.

Third, for this study, the researcher were able to provide only about 30 books for each class every month, since providing more books was too costly for the researcher. There was not a single book written in English available not only in the classrooms but in the school where this study was conducted. As Brewster et al. (1991) claim, having a book corner or a classroom library where students can read books of their own choice and at their own pace promotes a positive attitude towards reading and create enthusiasm among children for books. The problem many teachers or schools face is how to acquire a sufficient quantity of reading materials. Stocking a classroom library corner of the room or a school library with books purchased with money from school budgets or district grants is not an easy task. Nevertheless, it is necessary for teachers and educators to search for any available funds. Start stocking only 10-20 books each classroom and share books among classrooms can be a beginning step to invite storybooks to our classroom. Once reading materials are obtained, students choose what to read from the collection and set aside about 20

minutes of the school day for individual reading.

Fourth, as mentioned previously, children were recommended to read books or listen to the stories using the cassette recording whenever they want at school or to check them out to use at home. There were the book corner record cards to be recorded by students and the personal reading logs to be kept by the individual student. The researcher asked the homeroom teacher to recommend students to use the book corner as frequently as possible and to use it voluntarily. Although the researcher checked them once a week and gave two students who used the book corner the most some kind of prize each week, the number of the students who used frequently was not satisfactory for what the researcher expected. Throughout the research period, about 7 to 8 students in the each experimental group, and about 2-4 students in the each control group used the book corner regularly. The researcher tried to figure out and came up to an underlying reason that the students were recommended to use the book corner freely and the book corner was managed by the homeroom teacher. In order for the book corner to be efficiently utilized, it would be better for the students to have the teacher's guidance through which the whole class set aside about 20 minutes of the school day for individual reading with self-selected books from the book corner.

Several positive findings from the previous studies mentioned so far have been encouraging teachers to employ English storybooks for teaching English. The results of this study also conform to the educational values of using English storybooks for Korean primary school EFL learners. Yet, further research should explore more practical teaching models of story-based instruction for teachers, depending on learner's proficiency levels and their characteristics.

#### 5 REFERENCES

- Brewster, J., Ellis, G., & Girard, D. (2002). *The Primary English Teacher's Guide (New Edition)*. Harlow: Penguin Books, Longman Pearson Education Limited.
- Brumfit, C. j., & Carter, R. A. (1987). *Literature and language teaching*. New York: Oxford University Press.

- Cameron, L. (2001). *Teaching languages to young learners*. Cambridge: Cambridge University Press.
- Cho, K. S. & Krashen, S. K. (1994). Acquisition of vocabulary from the Sweet Valley Kids series: Adult ESL acquisition. *Journal of Reading*, 37(8), 662-667.
- Cho, K. S. & Krashen, S. K. (1995). Becoming a dragon: progress in English as a second language through narrow free voluntary reading. *California Reader*, 29(1), 9-10.
- Cho, K. S., & Seo, S. Y. (2001). The effects of reading aloud to children on interest and vocabulary acquisition in EFL classes. *Primary English Education*, 7(2), 222-242.
- Danielson, K. E., & Labonty, J. (1994). *Integrating reading and writing through children's literature*. Boston: Allyn & Bacon.
- Ellis, G., & Brester, J. (1991). *Storytelling Handbook for Primary Teachers*. New York: Penguin Press.
- Elly, W. B. & Mangubhai, F. (1983). The impact of reading on second language learning. *Reading Research Quarterly*, 19, 53-67.
- Elly, W. B. (1991) Acquiring Literacy in a Second Language: The effect of book-based programs. *Language Learning*, 41(3), 375-411.
- Elly, W. B. (1997). Raising literacy levels in third world countries: A method that works. Unpublished paper.
- Gilles, C. J. (1991). *Negotiating the meanings: the uses of talk in literature study*

# The Development of English Passive Constructions by Korean Learners

Hyesook Park

Dept of English Language of Literature, College of Humanities, Kunsan National University

sapark@kunsan.ac.kr

## Abstract

This study investigates how Korean speakers develop their interlanguage of English passive constructions with a reference to learners' grammar proficiency levels. College students of different levels of English participated in this study. They were asked to complete a sentence-completion task. Their production was classified into accurate passives, malformed passives, pseudo-passives, and overgeneralization of passive constructions. And then they are further analyzed depending on the subjects' levels of grammar proficiency by three main factors: L1 transfer, the English voice system, and universal semantic factors. Based on the analyses of the subjects' responses, the interlanguage passive construction system of Korean learners of English is shown and some implications are suggested for effective teaching of English.

## Keywords

English passive constructions, L1 transfer, universal semantic factor, English middle voice

## 1 Introduction

One of the interlanguage constructions SLA researchers have examined with interest is the IL passives produced by L2 learners of English. Here are some characteristic examples of IL passive constructions produced by different L1 backgrounds.

- (1) a. \*New cars must keep inside.  
b. \*These ways almost classify two types.  
(Chinese. Yip & Mathews, 1995, pp.17, 22)  
c. \*First, the change of life-style will be happened. (Korean, Min Kyung Ju, 1997)  
d. \*My mother was died when I was just a baby.  
(Thai, Zobl, 1989)

The above (1a) and (1b) are ungrammatical sentences with theme subjects followed by transitive verbs in active forms. They are typically termed IL(interlanguage) pseudo-passives because

their intended English structure is believed to be the passive (Yip & Matthews, 1995, pp.17, 22 cited in Simargool, 2008). On the other hand, (1c) and (1d) are ungrammatical sentences where learners inappropriately generalize the English passive rules to intransitive verbs. These overpassivization errors are prevalent even among advanced L2 learners of English and are observed across learners of different native languages.

With this background, this study attempted to examine IL passive constructions from Korean learners of different levels of English to see how Korean learners acquire English passives and what the main factor affecting the acquisition of English passive constructions is.

## 2 Background of the research

### 2.1 L1 transfer

Some researchers claimed that the IL passives result from L1 transfer, or reflection of the typology of the learner's L1. For example, the above pseudo-passives are claimed to reflect learners L1 background: pragmatic word order sentence structure (Han, 2000; Schachter & Rutherford, 1979). Schachter & Rutherford (1979) claimed that the IL pseudo-passive is a carryover of native language function-form characteristics, a type of discourse-syntactic transfer, by analyzing written English samples produced by Chinese and Japanese learners of English.

The Korean language has a similar pragmatic word order as well. Basically, the Korean language has a SOV word order different from English basic word order, SVO. Furthermore, Korean has properties of a practical word order language: the lack of articles, dummy subjects, and the subject-creating constructions (e.g., passive and raising constructions). In addition, the constituents in the sentences can be relatively freely moved unlike English or Chinese, as shown in (2) .

- (2) Younghee-ga Chulsu-lul saranghan-ta.  
(Younghee Chulsu loves.)  
Chulsu-lul Younghee-ga saranghan-ta.

(Chulsu Younghee loves.)  
Younghee loves Chulsu.

The meaning of the sentences is determined not by the order of the constituents, but by the case endings of the constituents.

Now let us look at the characteristics of the Korean passive construction. First, English differentiates passives from unaccusatives by marking passives morphologically. However, there isn't any distinction between the constructions for Korean as in (3).

- (3) English  
The window **was broken by** the boy. (passive)  
The window **broke**. (unaccusative)

Korean  
Changmun-I gu sonyun-ay uhaesu  
**ccay-ci-uss-ta**. (passive)  
(The window was broken by the boy.)  
.  
Cahngmun-I **ccay-ci-uss-ta**. (unaccusative)  
(The window broke.)

Second, the Korean passive constructions are generated by lexical and derived passive morphemes “-i, -hi, -li, -gi, -u, -gu, -ci, or -chu” as in (4) (No & Chung, 2006, p.25).

- (4) Lexical passives in Korean  
saracita disappear  
ttelecita fall

Derived passives in Korean  
ccay-ta break → ccay-ci-ta be broken  
baccwu-ta change → baccwu-i-ta be changed  
tat-ta close → tat-hi-ta be closed

No & Jung (2006) suggested that learners' L1 influenced their judgment of grammaticality on unaccusatives, showing Korean learners tended to judge ungrammatical passives containing “break, close, change” which have the derived passive morphemes in Korean as grammatical.

It can be predicted on the basis of the characteristics of the Korean language that Korean learners of English may produce IL pseudo-passives at the early stage of development of English passive constructions and prefer to overpassivize unaccusative verbs such as “break, change, close” which have derived passive morphemes in Korean.

## 2.2 The English voice system

Generally, the English voices can be simply classified into the active voice and the passive voice. In the active voice, the subject of a clause is usually the agent, or doer, of some action. On the other hand, in the passive voice, the subject of a clause is the receiver, or undergoer of the action. However, there is a middle voice between active and passive voices. The middle voice allows the subject of sentence to be nonagentive, as in the passive voice, but the morphology of the verb to be in the active voice.

Ergative, or change-of-state, verbs like “bake, break, change, close, open, boil, melt” can appear in all three voices and take either agents or undergoers of the action as subjects, as shown in (5).

- (5) I'm baking cookies.  
The cookies are baking.  
The cookies are being baked by her husband.

On the other hand, some other verbs like “happen, appear, disappear, arrive” take the focus of the process as subjects and do not have the passivized counterparts, as shown in (6).

- (6) The diamond disappeared.  
\*The man disappeared the diamond.  
\*The diamond was disappeared.

Likewise, English has a special middle voice and different categories of ergative verbs. It can be predicted based on the English inherent features that Korean learners may produce more passive constructions to ergative verbs with transitive alternations than ergative verbs without transitive counterparts.

## 2.3 Universal semantic factors

Still others argue that the learnability problem of the English passive constructions may be attributed to universal factors of the subjects in sentences. Research findings report that L2 learners from different backgrounds appear to either overpassivize unaccusatives or reject grammatical unaccusative constructions.

Kellerman (1979) reported that Dutch learners of English tended to reject grammatical sentences with some English unaccusatives (with transitive counterparts) as in (7a), while they preferred either causative constructions or agentless passives as in (7b) and (7c).

- (7) a. The cup broke.

- b. Someone broke the cup.
- c. The cup was broken.

This shows that L2 learners resist having nonagents as subjects for unaccusatives even when their L1 allows it.

Min Kyung Ju (2000) introduced a universal semantic factor to account for the overpassivization phenomenon of learners' passive constructions. According to her, learners are more likely to make overpassivization errors in externally caused events than internally caused events. Although no agent is posited semantically in sentences, from real-world knowledge in some contexts learners assume that there must be an entity that is responsible for an event. The entity that causes the event to occur is a pragmatically conceptualizable agent of the predicate. So, in the context where learners are likely to assume that there must be an implicit agent, learners may tend to apply passivization to the unaccusative verbs. Generally human beings have a tendency to naturally accept that any change of state occurs with external causation, but not to accept that any change of state occurs without any external causation. As a result, learners may use the passive when the sentence has an inanimate subject and an implicit agent, which results in overpassivization to unaccusative sentences.

Therefore, learners will be more likely to produce passivized constructions when the subject is inanimate or there is an implicit agent in using the verbs like 'break, change, close, boil, freeze, burn, disappear, fall, happen, appear'.

### 3. Research Design

#### 3.1 Research questions

- 1) What kinds of IL passive constructions do Korean learners of English produce?
- 2) What are the main factors that cause Korean learners to produce IL English passive constructions?
- 3) Are they dependant on their levels of English?

#### 3.2 Subjects

Sixty two Korean college students participated in this study. Their majors included English, math, social science, natural science and engineering, and their age ranged from 19 to 23. Most students began learning English in elementary school, so their average length of learning English was around 9 years.

For the baseline of the analysis, a control group was composed of 10 native teachers of English: two

female teachers and eight male teachers. They studied different majors such as communication, education, business, science, foreign languages, TESL and linguistics. Most of the teachers have taught English more than 7 years while one teacher has taught it since the spring semester, 2009 and two teachers have taught it around for 2 years.

#### 3.3 Task and procedures

The participants were given a sentence-completion task and a test of their grammar knowledge. The sentence-completion task was comprised of 24 pairs of nouns and verbs which was adapted from Simargool (2008) (Appendix 1), and the grammar test included 15 items which was intended to serve for judging learners' grammar proficiency levels.

The sentence-completion task included 24 verbs: 10 transitive verbs (read, write, paint, push, steal, sing, hit, win, catch, cancel), 10 unaccusative verbs (break, change, close; boil, freeze, burn; disappear, fall; happen, appear), and 4 unergative verbs (jump, die, swim, run). The ten unaccusative verbs included subclasses of unaccusative verbs: with/ without transitive counterparts and with/without the Korean passive morphemes (-ci, -i, -hi, -li, -gi) in Korean translations of the verbs. The distribution of the verbs in the task is as follows.

**Table 1 Distribution of the Verbs**

Verb types	Verbs used
Transitive	read, write, paint, win, push, steal, sing, hit, catch, cancel
Ergative I	break, change, close
Ergative II	boil, freeze, burn
Ergative III	disappear, fall
Ergative IV	happen, appear
Unergative	jump, run, swim, die

Ergative I – unaccusative verbs with transitive counterparts and Korean translation passive morphemes

Ergative II – unaccusative verbs with transitive counterparts and without Korean translation passive morphemes

Ergative III – unaccusative verbs without transitive counterparts and with Korean translation passive morphemes

Ergative IV – unaccusative verbs without transitive counterparts and Korean translation passive morphemes

### 4. Results and discussions

The 62 participants were divided into 2 groups depending on their scores of the grammar test: the

lower group and the higher group. All the analyses of the results will be given according to the two different groups. In addition, the responses of the native control group will be given as well.

#### 4.1 Results of transitive verbs

The participants' responses were classified into 5 categories: well-formed passives, malformed passives, possible pseudo-passives, actives, and others. Table 2 shows the results of 10 transitive verbs by the different groups.

**Table 2 Responses to transitive verbs**

Construction	Instances (%)	
	Lo-G	Hi-G
Well-formed passive	113(37.2)	153 (49.5)
Malformed passive	47 (15.5)	10(3.2)
Pseudo-passive	48 (15.8)	18(5.8)
Active	59 (19.4)	87 (28.2)
Others	37 (12.2)	41(13.3)
Total	304	310

The well-formed passives are native-like passives as in *the newspaper was read*, while the malformed passives include the ones with subject-verb agreement errors and errors in past participle markers, as in *\*the tiger was catch*, and *\*the tiger was catched*. Pseudo-passives are the ones with theme subjects and active verbs, as in *\*the cart pushed*, whereas active sentences are those with agent subjects and active verbs, as in *he is pushing the cart*. Other constructions include ones that don't belong to the above categories as in *don't push the cart*.

The students of the lower group showed 113 well-formed passives (37.2%), 47 malformed passives (15.5%), 48 pseudo-passives (15.8%), 59

actives (19.4%) and 37 others (12.2%), while those of the higher group exhibited 153 well-formed passives (49.5%), 10 malformed passives (3.2%), 18 pseudo-passives (5.8%), 87 actives (28.2%), and 41 others (13.3%). Compared to Simargool's study (2008) in which there were only 3 pseudo-passives among 380 instance, the present lower group formed many more pseudo-passives, 48 among 304 instances. This may be the evidence that the learners' L1 has an influence on their earlier stage of acquisition of passive constructions.

The chi-square test was performed to see if the different groups made a difference in sentence completion responses to transitive verbs, and it indicated that there was significant difference between the two groups ( $\chi^2=39.044$ ,  $df=4$ ,  $p=.000$ ).

#### 4.2. Results of ergative verbs

The ergative verbs with transitive counterparts and the Korean passive morphemes (break, close, change) were first analyzed according to unaccusative constructions, passive constructions, malformed unaccusative constructions, malformed passive constructions and others.

The unaccusative constructions are ones with a theme subject and active verbs like *the dish broke*, while the passive ones are sentences with a theme subject and be+-en as in *the dish was broken*. The malformed unaccusative constructions are unaccusatives with incorrect forms of past tense as in *the dish breaked*, while malformed passives are passive sentences with incorrect forms of past participle as in *the dish was broke*. Others include all the remnant, for example, actives with agent subject and imperatives as in *I broke a dish and Break the dish*.

**Table 3. Responses to ergative verbs I and II**

Constructions	Ergative I Instances (%)			Ergative II Instances (%)		
	Lo-G	Hi-G	Con-G	Lo-G	Hi-G	Con-G
Unaccusative	13 (14.1)	21(22.8)	17(56.6)	20(22.5)	46(49.5)	27(67.5)
Passive	52 (56.5)	46 (50.0)	9(30.0)	45(50.6)	27(29.0)	9(22.5)
Malformed unaccusative	6 (6.5)	0 (0.0)	0(0.0)	4(4.5)	1(1.1)	0 (0.0)
Malformed passive	8 (8.7)	4 (4.3)	0(0.0)	14(15.7)	6(6.5)	0(0.0)
Others	13 (14.1)	21(22.8)	4 (13.3)	6(6.7)	13(14.0)	4(10)
Total	92	92	30	89	93	40

As shown in Table 3, both the lower group and the higher group formed passive constructions more than unaccusative constructions when using the verbs “break, close, change”, although the use of unaccusative constructions increased from 14.1 % to 22.8 %. The native group did form 9 passive constructions out of 30 instances. Among the three verbs, both the learners and natives produced the most passive constructions with the verb “break”, while they made the least passive constructions with the verb “change”. The chi-square test indicated that there was a significant difference between the two learners groups ( $X^2=11.465$ ,  $df=4$ ,  $p=.022$ ).

Let us look at the other subclass of ergative II verbs which have transitive counterparts, but don’t have the Korean derived passive morphemes such as “boil, freeze, burn”.

The lower group formed more passive constructions than unaccusative constructions (50.6% - 22.5%), while the higher group used unaccusative constructions 49.5% of the time and passive constructions 29.0%, showing that learners develop their interlanguage towards the native speakers’ system of English. Korean learners used more passive constructions in the verbs “boil, freeze, burn” than accusative constructions, in particular, for the lower group, although these verbs don’t have the Korean passive morphemes.

The chi-square test indicated that the two groups differently responded to the ergative verbs II ( $X^2=32.299$ ,  $df=4$ ,  $p=.000$ ).

For the lower group, they similarly responded to ergative verbs I and ergative verbs II, showing they preferred to use passive constructions to unaccusative constructions regardless of the subclasses of the ergative verbs. However, the learners of the higher group formed more unaccusative constructions than passive constructions to the ergative verbs II.

Let us consider the ergative verbs which don’t have transitive counterparts. First, let us look at ergative III verbs with Korean passive morphemes such as “happen, appear”. The lower group used unaccusative constructions 47.5% and passive constructions 42.6% to the verbs “happen, appear”, while the higher group used unaccusative constructions 83.9% to the verbs. The subjects of the lower group used much more passive than those of the higher group. This is a remarkably interesting result as we consider they don’t have transitive counterparts and Korean passive morphemes.

Now let us consider the ergative IV verbs (‘disappear, fall’) which don’t have transitive counterparts, but have the Korean translation passive morphemes to see if there is any influence of L1 transfer. As shown in Table 4, both the lower group and the higher group used more unaccusative constructions than passive constructions even though the verbs have the

**Table 4 Responses to ergative verbs III and IV**

Constructions	Ergative III			Ergative IV		
	Instances (%)			Instances (%)		
	Lo-G	Hi-G	Con-G	Lo-G	Hi-G	Con-G
Unaccusative	29(47.5)	52(83.9)	20(100)	27(45.8)	49(79.0)	20(100)
Passive	26(42.6)	8(12.9)	0	12(20.3)	7(11.3)	0(0.0)
Malformed unaccusative	1(1.6)	0(0.0)	0	13(22.0)	2(3.2)	0(0.0)
Malformed passive	2(3.3)	2(3.2)	0	5(8.5)	2(3.2)	0(0.0)
Others	3(4.9)	0(0.0)	0	2(3.4)	2(3.2)	0(0.0)
Total	61	62		59	62	20

Korean translation passive morphemes. In particular the percentage of passive constructions for “disappear, fall” was even lower than those for “happen, appear.”

### 5.3 Results of unergative verbs

The subjects’ responses to the unergative verbs were divided into 5 categories: actives, passives, malformed actives, malformed passives, and

others. Sentences like the boy jumped were given to the actives, while the passives were sentences like *the boy was jumped*. Malformed actives were sentences with tense and agreement errors, as in *the boy jump into swimming pool*, while malformed passives were ones with past participles errors or ‘subject- be’ agreement errors, as in *the dog was die* and *the boy were jumped*.

**Table 6 Responses to unergative verbs**

Constructions	Instances (%)		
	Lo-G	Hi-G	Con-G
Active	73(58.9)	105(84.7)	37 (92.5)
Passive	19(15.3)	7(5.6)	0 (0.0)
Malformed active	14(11.3)	3 (2.4)	0 (0.0)
Malformed passive	7(5.6%)	0 (0.0)	0 (0.0)
Others	11(8.9)	9 (7.3)	3 (7.5)
Total	124	124	40

It was shown that the subjects of the lower level chose 29% and 25.8% passives for the verbs “die, run” respectively, and even the more advanced subjects formed 16.1 % passives to the verb “die”, as in *my dog was died because of virus*.

The chi-square test showed that there was a difference in the responses to the unergative verbs between the two groups ( $X^2=25.609$ ,  $df=4$ ,  $p=.000$ ).

## 6. Conclusions and implications

The followings are the finding from analyses of the results:

1) Korean learners of English produced malformed passives, pseudo-passives, and passives to ergative verbs and unergative verbs. In particular, the learners of lower level showed malformed passives, pseudo-passives, and overpassivization remarkably more than did those of higher level. In particular, they formed sentences like pseudo-passives and tended to choose passive voice to change of state verbs.

2) The analyses of the results revealed that as the learners had a lower level of English, they used more pseudo-passives, showing the subjects' L1 have an influence on their earlier stage of acquisition of the English passive constructions. The subjects preferred to use passives to ergative verbs which have transitive counterparts and may imply an external agent, and even to the unergative verb of ‘die’ in particular, in an adversative context. This shows universal semantic factors play an important role in the acquisition of English passive constructions. However, the existence of Korean passive morphemes does not seem to have as much influence as No & Jung (2006) suggested.

3) Nevertheless, the learners of higher level formed fewer passive constructions to ergative verbs than did those of lower level. The chi-square tests indicated that there were significant differences in the formation of passive

constructions between the two different groups. This can be explained that through extensive exposure to English, they are restructuring their interlanguage system of English passives.

Some implications to SLA can be drawn based on the findings:

- 1) Positive evidence may be not sufficient and effective for L2 learners of English.
- 2) With positive evidence, proper negative evidence should be given to help L2 learners acquire English passive constructions and avoid the overproduction to ergative and unergative verbs.

## References

- Balcom, Patricia. (1997). Why is this happened? Passive morphology and unaccusativity. *Second Language Research*, 13(1), 1-9.
- Ju, Min Kyong. (2000). Overpassivization errors by second language learners. *Studies in Second Language Acquisition*, 22, 85-111.
- Han, Z. (2000). Persistence of the implicit influence of NL: The case of the pseudo-passive. *Applied Linguistics*, 21(1), 78-105.
- Kellerman, E. (1979). Transfer and nontransfer: Where we are now. *Studies in Second Language Acquisition*, 2, 37-57.
- No, GyeongHee & Chung, Taegoo. (2006). Multiple effects and the learnability of English unaccusatives. *English Teaching*, 61(1), 19-39.
- Schachter, J. & Rutherford, W. J. (1979). Discourse function and language transfer. *Working Papers in Bilingualism*, 19, 1-12.
- Simargool, Nirada. (2008). Interlanguage passive construction. *Journal of PAAL*, 12(1), 97-119.
- Yip, V. & Matthews, S. (1995). I-interlanguage and typology: The case of topic-prominence. In L. Eubank, L. Selinker, & M. Sharwood Smith (Eds.), *The current state of interlanguage: Festschrift* (pp.17-30). Amsterdam: John Benjamins.
- Yip, V. (1995). *Interlanguage and learnability from Chinese to English*. Philadelphia: John Benjamins.
- Zobl, H. (1989). Canonical structures and ergativity. In S. M. Gass & J. Schachter (Eds.), *Linguistic perspectives on second language acquisition* (pp. 203-221). New York: Cambridge University Press.

# The Association between EFL Learners' Reading Motivation and Their L2 Reading Behavior

Eun-Mi Yang

Kkottongnae Hyundo University of Social Welfare

emyang@kkot.ac.kr

## Abstract

This paper examined how the Korean EFL students' reading motivation is related to their L2 reading behavior. The construct of the students' L2 reading motivation was investigated with the data obtained through the questionnaire. A principal components analysis was conducted to extract the major factors and determine the interrelationship among items in the questionnaire. The Interrelatedness between the results of the pleasure reading (reading amount and time) and motivation factors was calculated by correlation coefficients.

## Keywords

extensive reading, pleasure reading, L2 reading motivation

## 1 Introduction

Even though motivation to read has been a very popular research topic in the field of language education (Gardner, 1985, 1988; Kim, 2004; Noels, Pelletier, Clement, & Vallerand, 2000), motivation to read in an L2 and its relationship with reading achievement has not been explored that much. Some researchers reported that intrinsic motivation was the stronger index of the amount of reading than extrinsic motivation. However, most of the studies have been conducted only to see the relationship between L1 and L2 reading motivation with extensive reading. There still seems to be room to study the relationship among L2 reading motivation, reading amount, and reading achievement in the EFL context of Korea. So the purpose of this study is to elucidate the relationship between the kinds of L2 reading motivation of the students and their L2 reading behavior, such as actual amount of reading.

## 2 Method

### 2.1 Subjects

The subjects in this study are 120 second year

students at a university in Korea. Their majors are nursing science (N=40) and social welfare (N=80) and approximately 80% of the students are female. Most of them had received formal English education at least for eleven years when the research was begun. They were divided into four classes on a voluntary course registration basis. The participants took a 2-hour English class per week for 16 weeks. They studied reading, speaking, listening, and writing in the class, and the contents of the reading were mostly related to social issues. The students' English proficiency levels ranged from low-intermediate to high intermediate based on their Suneung (SAT) scores (Their English scores ranged between level 3 and 6 out of 9 levels, and their average level was about 4.).

### 2.2 Procedures

#### 2.2.1 Setting up the Pleasure Reading Assignment

The pleasure reading assignment was designed and implemented considering its learning potential to the students who basically intensively read only the reading material provided in the textbook for the English course. The students were required to read English story books for assignment. On the first day of the semester they were advised to choose books among the graded readers which interest them. It was emphasized that they would read books for having pleasure as well as learning English. I gave them some tips to choose books which are appropriate for their English reading level and told them it is important to choose a book which is not difficult for pleasure reading.

For the sake of having pleasure it must be desirable not to relate the assignment for the course grade, but formative evaluation of the reading assignment was also needed since it was given as a course assignment. So they were asked to read at least two hours a week, to

write a reading log whenever they read, and to submit the reading record sheet at the end of each month. They promised to be honest when they proceed the reading log, and the basic score was given to everybody whoever read at least two hours a week. The items they had to record about their reading on the reading record sheet were: the book title, publisher, reading level of the book, date, reading time, reading level they perceived, and page numbers they read. They also summed up and recorded the total reading time and total page numbers on their reading record sheet at the end of each month. The researcher recorded the students' data on an SPSS data sheet for analysis.

### 2.2.2 Construction of the Questionnaire

A forty-item 4-point Likert-scale questionnaire was constructed based on studies in the field of L1 and L2 reading motivation (Mori, 2002; Takase, 2007; Wigfield & Guthrie, 1997) in order to find out the motivational components toward L2 reading of the Korean university students. The questionnaire consisted of 40 items. Among them, thirty items are related to motivation and attitude toward English (L2) reading and they replicated the questionnaire items Mori (2002) developed. Mori's subjects, Japanese university students, were in an EFL situation similar to mine, so her questionnaire items were adopted for my study. Four out of 40 items are regarding motivation toward Korean (L1) reading, and 6 questions are about general English study.

The questionnaire written in Korean was administered to the subjects at the beginning of the semester in each class. The students answered each item on a 1 to 4 scale, with answer choices from *strongly disagree* to *strongly agree*.

### 2.2.3 Construction of the Questionnaire

Reading speed was tested twice at the beginning and at the end of the semester. Two pages of a book was copied and distributed to each student and the students read the passage for two minutes. They marked on the spot of the passage when they finished reading for 2 minutes at the beginning of the semester. The same copy was distributed to each student again at the end of the semester, and they repeated the 2-minute-reading and marked the spot when they finished. The numbers of the words were recorded and compared to examine the existence of any change of their reading speed.

### 2.2.4 Data Analyses

The questionnaire scores and the results of the reading amount (total number of pages) and the reading time of the subjects were analyzed using SPSS 12.0. First, Cronbach's coefficient alpha was calculated to estimate the internal consistency reliability of the questionnaire for measuring students' motivation in L2 reading. The reliability estimate was .623 and considered acceptable for this study. Second, a principal components analysis was conducted to extract the major factors and determine the interrelationship among items in the questionnaire. The result of KMO and Barlett test indicated the correlation among items were significant as .720 ( $p = .000$ ). Correlation coefficients were calculated to find out the interrelatedness between the results of the pleasure reading (reading amount and time) and motivation factors. The relationship between the results of the pleasure reading and motivation toward reading in L1 was measured through regression analysis.

## References

- Crawford Camiciottoli, B. (2001). Extensive reading in English: Habits and attitudes of a group of Italian university EFL students. *Journal of Research in Reading*, 24(2), 135-153.
- Day, R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge: Cambridge University Press.
- Im, B. (2007). An approach to improve college students' EFL reading comprehension through rapid reading and pleasure reading techniques. *English Language & Literature Teaching*, 13(1), 181-210.
- Jeon, Jongmin. (2008). Extensive reading in a formal English reading class. *English Teaching*, 63(4), 49-84.
- Mori, S. (2002). Redefining motivation to read in a foreign language, *Reading in a Foreign Language*, 14(2). Retrieved January 15, 2009 from the World Wide Web: [http://nflrc.hawaii.edu/r 2/october2002/mori/mori.html](http://nflrc.hawaii.edu/r%20/october2002/mori/mori.html).
- Takase, A. (2007). Japanese high school students' motivation for extensive L2 reading. *Reading in a Foreign Language*, 19(1). Retrieved December 10, 2008 from the World Wide Web: <http://nflrc.hawaii.edu/rfl/april2007/takase/takase.html>

## A Discussion of Understandable English

Kyunghee Choi

English Language Teaching, Department of International Tourism,

Division of Literal and Social Arts, Hanyang Women's College

khchoi@hywoman.ac.kr

### Abstract

The purpose of this study is to discuss the result of a judgment test of 24 inadequate English sentences, conducted toward 24 college students and 3 college English teachers coming from a so-called native English speaking country. Among the 24 students of a course offered in English in a Korean university in Seoul, 17 are Korean nationals. However, 8 of them lived abroad for the period of 10 months and up to 14 years. 7 of them are from the countries including USA, UK, Canada, and Malaysia. The 3 English teachers came from New Zealand. A questionnaire has 24 sentences, each of which has a more or less grammatical problem. It asks the participants to indicate whether they think each question is correct, incorrect, or fine to understand even if it is problematic. The study analyses whether and how each question is understood and judged differently by the participants depending on different linguistic backgrounds, depending on their jobs – students or teachers, depending on experience of living abroad even though they are from the same Korean linguistic background, and depending on different native English speaking nationalities. In addition, the study aims to discuss some characteristics of Korea English found in the answers of the Korean participants.

### Keywords

Native speaker, non-native speaker, World Englishes, Asia Englishes, Korea English

### Introduction

English is becoming an ever more powerful language in the Korean society. Even though English is not yet an official language and still used as a foreign language in the nation, the situation is different in major universities in Seoul. Many courses are offered in English in those universities in which a lot of foreign students come to study either on their own expenses or as exchange students. In addition, the number of Korean students who have lived abroad and return to study in their own country is increasing. It is no

exaggeration to say that some major Korean universities are indeed becoming global universities. The Korean students in those universities who have never lived abroad have to communicate in English with their professors and classmates. Practically English is used as an official language in some major universities in Korea.

### 1 Theoretical Background

Many researchers (Choi, 2007a, 2007b; Hughes & Trudgill, 1979; Jenkins, 2003; Trudgill & Hannah, 1994) argue that Standard English is only a social variety of language. Even among the so-called standard varieties of English, there are standard and non-standard varieties and it is hard to claim that there is one and only Standard English. Jenkins (2003) even argues that the distinction between a native speaker and a non-native speaker is blurred since there are many English speakers coming from the outer circle countries including Singapore who claim themselves to be native speakers.

Local varieties of English influenced by local languages and cultures have been studied by many researchers around the world (Alsagoff & Lick, 1998; Bao, 2003; Bao & Wee, 1999; Bautista, 2004; Benson, 2000; Bolton, 2000; Bolton & Bautista, 2004; Bolton & Butler, 2004; Gupta, 1991; 1994; Hung, 2000; Kachru, 1992; Kirkpatrick & Zhichang, 2002; McArthur, 2003, 2004; Peng & Ann, 2004; Tayao, 2004; Wee, 2003, 2004; Yang, 2005; Zhang, 2002). Introduced some 130 years ago, English has become an important part of the Korean society. Now, it is time to consider whether it has developed its own type of English – Korea English.

Instead of insisting on the correct English and Standard English which our students may never be able to attain but rather makes them feel frustrated by being far from the standard, it is more practical and efficient to give more opportunities to use the language without fear of making mistakes (Choi, 2007b; Honna & Takeshita, 2000). Instead of insisting on the right English for our students, we need to teach understandable English.

Some of the characteristics of Asia Englishes

especially in the grammatical structures can be considered incorrect and problematic by the so-called native English speakers. But are these unique features of local Englishes or faulty features? This study does not attempt to conclude whether they are unique characteristics of local varieties or wrong features. Instead the study is designed to analyze how students and teachers, and non-native speakers and native speakers evaluate and perceive some faulty sentences. Do the participants in the study think the sentences are incorrect or correct or fine to understand even though they are problematic? By analyzing the result of the study, it intends to suggest how far English teachers can be lenient with their students when they produce inadequate sentences while learning speaking English for every day communicative purposes.

## 2 Research Methods

### 1.1 Participants

As Appendix 2 shows, among the 27 participants, 14 are female and 13 are male. 24 of them are students and 3 of them are English teachers all coming from New Zealand. 17 Korean nationals participated in the survey. Among them, 9 have never lived abroad but only travelled for a short period of time. And 8 of them have lived abroad from 10 months to up to 14 years. Most of them lived in such English speaking countries as USA, UK, Australia, New Zealand, and Singapore, etc. There are 7 foreign students – 2 Malaysian students, 2 UK students, 2 Korean-American students, and 1 Korean-Canadian student.

### 1.2 Questionnaire

The questionnaire (see Appendix 1) includes 24 inadequate sentences. Some sentences are quoted from the data of the cross cultural distance learning (CCDL) between Korea University and Waseda University (Lee, 2007; Park, 2005). Many of them can be seen and heard easily in East Asia Englishes (Choi, 2007a).

(1) “Today was very difficult but exciting” (Lee, 2007; Park, 2005). Where a dummy subject should be used, a special subject ‘Today’ is used. This kind of sentence is often used both in Korea English and China English (Yun & Jia, 2003). (2) “Thanksgiving always is celebrated in November” (Park, 2005). It is an example of a wrong word order. (3) “No wonder you can’t sleep when you have coffee too much.” Instead of ‘too much,’ ‘so much’ has to be used.

(4) “know how to go” (Lee, 2007). This is an example of a subject deletion. In China English, similar examples can be seen. Xu Xi (2000) illustrates that null-subject utterance is a feature of

China English. For example, “Sometimes \_\_\_ just play basketball, and sometimes \_\_\_ go to the Beijing library.” (5) “In Korea have four season” (Lee, 2007). It is another example of a null-subject sentence. The sentence should be “In Korea, there are four seasons,” or “Korea has four seasons,” or “There are four seasons in Korea.” (6) “I thanked for” (Lee, 2007). This is an example of an object deletion. Xu Xi (2000) also indicates that null-object sentences are quite often observed in the utterances spoken by Chinese people.

The sentences (7), (8), and (9) are examples of wrong usage of articles. (7) “They had big dinner” (Lee, 2007). In this sentence, the indefinite article ‘a’ is omitted. It should be inserted. (8) “Although it is a hard work, I enjoy it” (Lee, 2007). The unnecessary article ‘a’ should be deleted. (9) “The place is near the Gangnam Station.” In this sentence, an unnecessary definite article is used.

(10) “I graduated gyeonggi high school” (Lee, 2007; Park, Lee, & Ju, 2003). In this sentence, the preposition, ‘from’ is omitted. (11) “I want to study about many fields” (Lee, 2007; Park, Lee, & Ju, 2003). In this sentence the preposition ‘about’ should be omitted. (12) “I’m leaving after two hours.” This is an example of wrong usage of preposition. Instead of ‘after,’ ‘in’ has to be used.

(13) “Her skin color was changed.” (14) “I was confusing.” (15) “I disappointed.” For Asian English speakers including Koreans, passive and active voices are also quite confusing. Sometimes they use passive sentences where active voices are need as in the sentence (13), and vice versa as in the sentence (14). Sometimes they omit ‘be’ as illustrated in the sentence (15).

(16) “I cut my hair.” Since you do not usually cut your hair by yourself, you have to go to a barber shop and have your hair cut. Therefore the sentence should be “I had my hair cut” or “I had a haircut.” (17) “I had a horrible toothache. I pulled my tooth.” Again, you cannot pull your tooth by yourself. You have to go to a dentist and have your tooth pulled. (18) “I dry-cleaned my jacket.” People usually have their jackets dry-cleaned in a dry-cleaner. These types of sentences are particularly difficult for Korean people. However, some native speakers sometimes are confused as the result of this study shows.

(19) “He seems to be pleased when he saw me.” (20) “But as time goes by I was accustomed to it.” (21) “So in the second semester, I attended every class and sometimes go to the lab and listen to tapes many times.” All these sentences are examples of wrong usage of tense. There is no agreement between verb tenses. The improper uses of tenses are quite often observed in Korea English.

(22) “She gave me many helpful advices about

living abroad” (Lee, 2007). (23) “Please pass me two pencils.” (24) “Please send me more informations about your company” (Lee, 2007). Non-countable nouns are incorrectly pluralized in Korea English. On the other hand, when a count noun should be pluralized, it is not properly pluralized.

The participants are asked to indicate whether they think each sentence is incorrect (hard to understand), or fine (not quite correct but understandable) or correct. And additionally at the end of each question, the participants are requested to indicate whether they would use the sentence by ticking (yes) or (no).

### 1.3 Research Questions

The four major research questions are:

- (1) Is each question understood and judged differently by the participants depending on different linguistic backgrounds?
- (2) Is each question understood and judged differently depending on their jobs – students or teachers,
- (3) Is each question understood and judged differently depending on experience of living abroad even though they are from the same Korean linguistic background?
- (4) Is each question understood and judged differently depending on different native English speaking nationalities?

### 1.4 Analysis and Discussion

There are 24 questions in the questionnaire answered by the 27 participants. If a group of 19 responded to the questionnaire, there are 456 answers to count. In Table 1, there are 19 participants in the non-native group. The percentage of the answers marked incorrect is 28.5% since there are 130 answers out of the total 456 answers. The following percentages in the tables are calculated this way.

As Table 1 shows, in terms of the ability of recognizing the correct sentences, the native speaker group does far better than the non-native group. However, even though the NS group understands that some sentences are incorrect, they tend to categorize them as fine structures while the NNS group tends to mark them incorrect. It is interesting to note that even though NS group considers the inadequate sentences fine to understand, they marked that they would not speak these ways.

Table 1: Different linguistic backgrounds (NNS-NS)

	NNS	NS
Group	1, 2, 3	4, 5, 6, 7

Number	19		8	
Incorrect	130	28.5%	14	7.3%
Fine	251	55.0%	159	82.8%
Correct	75	16.5%	19	9.9%
No	283	62.1%	172	89.6%
Yes	173	37.9%	20	10.4%

There are no significant differences between the two non-native groups, the Korean group and the Malay group.

Table 2: Different linguistic backgrounds

Lang.	Korean		Malay		English	
Group	1, 2		3		4, 5, 6, 7	
Number	17		2		8	
Incorrect	119	29.2%	11	22.9%	14	7.3%
Fine	225	55.1%	26	54.2%	159	82.8%
Correct	64	15.7%	11	22.9%	19	9.9%
No	253	62.5%	30	62.5%	172	89.6%
Yes	155	37.5%	18	37.5%	20	10.4%

Even though all the sentences are problematic, the 3 native speaker teachers marked 6 answers incorrectly by marking them correct. However, they did better than the student group.

Table 3: Students (NNS & NS) vs. Teachers (NS)

	Students		Teachers	
Group	1, 2, 3, 4, 5, 6		7	
Number	24		3	
Incorrect	138	24.0%	6	8.3%
Fine	350	60.8%	60	83.4%
Correct	88	15.2%	6	8.3%
No	388	67.4%	17	93.1%
Yes	188	32.6%	5	6.9%

In the comparison between the native speaker student group and the native speaker teacher group, there seems to be no significant difference. However, when asked whether they would use the sentences, more teachers answered that they would not use them than students.

Table 4: Students (NS) vs. Teachers (NS)

	Students		Teachers	
Group	4, 5, 6		7	
Number	5		3	
Incorrect	8	6.7%	6	8.3%
Fine	99	82.5%	60	83.4%
Correct	13	10.8%	6	8.3%
No	105	87.5%	17	93.1%
Yes	15	12.5%	5	6.9%

The Korean students who have lived abroad did better in recognizing incorrect and correct sentences than the Korean students who have never lived abroad and only traveled for a short period of time. Also those who have lived abroad tend to use the inadequate sentences slightly less than those who have never lived abroad.

Table 5: Koreans (never abroad) vs. Koreans (abroad)

	Never lived abroad		Lived abroad for more than 6 mths	
Group	1		2	
Number	9		8	
Incorrect	50	23.2%	69	35.9%
Fine	127	58.7%	98	51.1%
Correct	39	18.1%	25	13.0%
No	126	58.3%	127	66.1%
Yes	90	41.7%	65	33.9%

There seems to be no significant differences between different English speaking countries. The two students from the USA, however, did slightly better than the three NZ teachers in indicating incorrect and correct sentences.

Table 6: Different English speaking countries

	Canada		UK		USA		NZ	
Group	4		5		6		7	
Number	1		2		2		3	
Incorrect	1	4.2%	2	4.2%	5	10.4%	6	8.3%
Fine	18	75.0%	39	81.3%	42	87.5%	60	83.4%
Correct	5	20.8%	7	14.5%	21	2.1%	6	8.3%
No	18	75.0%	40	83.3%	47	97.9%	67	93.1%
Yes	6	25.0%	8	16.7%	1	2.1%	5	6.9%

Both Korean groups who have lived abroad and who have never lived abroad seem to have a few problems in using articles, prepositions, and passive and active voices. They are also confused with tenses and pluralization of non-countable nouns. The sentences like (17) and (18) seem to be difficult for Koreans. For the question (18), 7 out of 17 said that they would say the sentence “I dry-cleaned my jacket” while only 3 out of the 17 Korean participants incorrectly marked it correct.

The native speakers those who came from the commonwealth countries considered the sentence (1) “Today was very difficult but exciting” correct. All the 6 participants marked them as correct sentences even though only 2 of them said that they would use the sentence. On the other hand, the two American participants answered that it is fine but they would not use it.

For the 3 New Zealand teachers, the sentences (16) and (18) are confusing. 2 of them think the sentence (16) “I cut my hair” is a correct sentence but they would use it while 1 of them thinks the sentence (18) “I dry-cleaned my jacket” is correct but 2 of them said they would use it.

### 3 Conclusion

In recognizing the sentences correct or incorrect, the native participants did far better. However, they seem to be more lenient in the incorrect sentences even though they would not use them. From the

native speakers’ point of view, the inadequate sentences included in the questionnaire are not too critically problematic to understand. If they can understand, why can’t the non-native speakers. The knowledge that whether such a sentence is right or wrong is helpful in many ways in understanding and using English. However, if one is excessively obsessed with the correctness, he/she may lose chances to practice the language.

One of the shortcomings of this study is a small number of participants. There are only 2 Malay students who answered the questionnaire. Thus, it is hard to compare the two non-native groups. The two Asian student groups seem to show no significant differences. There are many common features in Asian Englishes suggested by some researchers recently (Choi, 2007a; Kirkpatrick, 2007; Lee, 2007)

Again a small number of native speaker participants is problematic. However, 9.9% of the total answers that the NS groups responded were wrong answers. Even the 3 native speaker teachers marked 8.3% answers incorrectly by marking them correct. It is interesting to note that they are also confused with such passive sentences as (16) and (18). The 6 participants coming from the commonwealth countries marked the sentence (1) “Today was very difficult but exciting” correct while the 2 American participants regarded them fine but they would not use it. This difference may be influenced by British English.

Those Koreans who have experiences of living abroad did a slightly better in marking correct and incorrect sentences. Regardless of the fact that which group they belong to, Koreans seem to have difficulties with using articles, prepositions, passive voice and tenses. These common errors may reflect the influence of their first language, Korean. Since there are only 17 Koreans participated in the study, it is hard to conclude that these common errors are unique features of Korea English. However, they may be useful data in the investigation of characteristics of Korea English.

## 4 References and appendices

### 4.1 References

- Alsagoff, L., & Lick, H. C. (1998). The relative clause in colloquial Singapore English. *World Englishes*, 17, 127-138.
- Bao, Z. (2003). Social stigma and grammatical autonomy in nonnative varieties of English. *Language in Society*, 32, 23-46.
- Bao, Z., & Wee, L. (1999). The passive in Singapore English. *World Englishes*, 18, 1-11.
- Bautista, M. L. S. (2004). The verb in Philippine

- English: A preliminary analysis of modal *would*. *World Englishes*, 23, 113-128.
- Benson, P. (2000). Hong Kong words: Variation and context. *World Englishes*, 19, 373-380.
- Bolton, K. (2000). The sociolinguistics of Hong Kong and the space of Hong Kong English. *World Englishes*, 19, 265-285.
- Bolton, K., & Bautista, M. L. S. (2004). Philippine English: tensions and transitions. *World Englishes*, 23, 1-5.
- Bolton, K., & Butler, S. (2004). Dictionaries and the stratification of vocabulary: Towards a new lexicography for Philippine English. *World Englishes*, 23, 91-112.
- Choi, K. (2007a). East Asia Englishes. In K. Park & M. Nakano (Eds.), *Asia Englishes and Miscommunication* (pp. 115-137). Seoul: Korea Univ. Press.
- Choi, K. (2007b). A study on students' attitude towards World Englishes and non-native English teachers. *English Teaching*, 62(4), 47-68.
- Gupta, A. (1991). Acquisition of diglossia in Singapore English. In A. Kwan-Terry (Ed.), *Child language development in Singapore and Malaysia* (pp. 119-160). Singapore: Singapore Univ. Press.
- Gupta, A. (1994). *The step-tongue: Children's English in Singapore*. Clevedon: Multilingual Matters.
- Honna, N., & Takeshita, Y. (2000). English language teaching for international understanding in Japan. *English Australia Journal*, 18(1), 60-78.
- Hughes, A., & Trudgill, P. (1979). *English accents and dialects*. London: Arnold.
- Hung, T. T. N. (2000). Towards a phonology of Hong Kong English. *World Englishes*, 19, 337-356.
- Jenkins, J. (2003). *World Englishes: A resource book for students*. London: Routledge.
- Kahane, H. (1992). American English: from a colonial substandard to a prestige language. In B. B. Kachru (Ed.), *The other tongue. English across cultures*. Urbana, IL: Univ. of Illinois Press.
- Kirkpatrick, A. (2007). *World Englishes: Implications for international communication and English language teaching*. Cambridge: Cambridge Univ. Press.
- Kirkpatrick, A., & Zhichang, X. (2002). Chinese pragmatic norms and 'China English.' *World Englishes*, 21, 269-279.
- Lee, K. (2007). Korea English. In K. Park & M. Nakano (Eds.), *Asia Englishes and Miscommunication* (pp. 241-315). Seoul: Korea Univ. Press.
- McArthur, T. (2003). English as an Asian language. *English Today*, 19, 19-22.
- McArthur, T. (2004). Singapore, grammar, and the teaching of 'internationally acceptable English.' *English Today*, 20, 13-19.
- Park, K. (2005). *Korean English as a glocalized English*. Paper presented at the Waseda University Digital Campus Consortium.
- Park, K., Lee, H., & Ju, Y. (2003). *Syntactic transfer by Korean learners of English*. Paper presented at the Waseda University Digital Campus Consortium.
- Park, K., & Nakano, M. (Eds.). (2007). *Asia Englishes and Miscommunication*. Seoul: Korea Univ. Press.
- Peng, L., & Ann, J. (2004). Obstruent voicing and devoicing in the English of Cantonese speakers from Hong Kong. *World Englishes*, 23, 535-564.
- Trudgill, P., & Hannah, J. (1994). *International English*. Cambridge: Cambridge Univ. Press.
- Wee, L. (2003). The birth of a particle: *know* in Colloquial Singapore English. *World Englishes*, 22, 5-13.
- Wee, L. (2004). Reduplication and discourse particles. In L. Lim (Ed.), *Singapore English* (pp. 105-126). Amsterdam: John Benjamins.
- Xu, X. (2000). Writing the literature of non-denial. *World Englishes*, 19(3), 415-428.
- Yang, J. (2005). Lexical innovations in China English. *World Englishes*, 24, 425-436.
- Yun, W., & Jia, F. (2003). Using English in China. *English Today*, 19(4), 42-47.
- Zhang, H. (2002). Bilingual creativity in Chinese English: Ha Jin's *In the Pond*. *World Englishes*, 21, 305-315.

## 4.2 Appendix 1

### Questionnaire:

#### Are these sentences understandable?

- (1) Your first language: \_\_\_\_\_
- (2) Duration of your stay in Korea: \_\_\_\_\_
- (3) Duration of your stay in foreign countries and where? \_\_\_\_\_ & \_\_\_\_\_
- (4) Male ( ) Female ( )
- (5) What do you think of your English speaking ability?
  - ① very poor
  - ② fine (understandable, no difficulties in communicating)

- ③ very good  
④ excellent  
⑤ a native speaker

(6) Department: \_\_\_\_\_

□ What do you think of the following sentences?  
Please circle the number that you think is the most appropriate and underline any problematic word or phrases if there are any.

- ① incorrect (hard to understand)  
② fine (not quite correct but understandable)  
③ correct  
\* I would speak this way. ( yes ) ( no )

1. Today was very difficult but exciting.  
① ② ③ \* ( yes ) ( no )
2. Thanksgiving always is celebrated in November.  
① ② ③ \* ( yes ) ( no )
3. No wonder you can't sleep when you have coffee too much.  
① ② ③ \* ( yes ) ( no )
4. know how to go.  
① ② ③ \* ( yes ) ( no )
5. In Korea have four season.  
① ② ③ \* ( yes ) ( no )
6. I thanked for.  
① ② ③ \* ( yes ) ( no )
7. They had big dinner.  
① ② ③ \* ( yes ) ( no )
8. Although it is a hard work, I enjoy it.  
① ② ③ \* ( yes ) ( no )
9. The place is near the Gangnam Station.  
① ② ③ \* ( yes ) ( no )
10. I graduated gyeonggi high school.  
① ② ③ \* ( yes ) ( no )
11. I want to study about many fields.  
① ② ③ \* ( yes ) ( no )
12. I'm leaving after two hours.  
① ② ③ \* ( yes ) ( no )
13. Her skin color was changed.  
① ② ③ \* ( yes ) ( no )
14. I was confusing.  
① ② ③ \* ( yes ) ( no )

15. I disappointed.

- ① ② ③ \* ( yes ) ( no )

16. I cut my hair.

- ① ② ③ \* ( yes ) ( no )

17. I had a horrible toothache. I pulled my tooth.

- ① ② ③ \* ( yes ) ( no )

18. I dry-cleaned my jacket.

- ① ② ③ \* ( yes ) ( no )

19. He seems to be pleased when he saw me.

- ① ② ③ \* ( yes ) ( no )

20. But as time goes by I was accustomed to it.

- ① ② ③ \* ( yes ) ( no )

21. So in the second semester, I attended every class and sometimes go to the lab and listen to tapes many times.

- ① ② ③ \* ( yes ) ( no )

22. She gave me many helpful advices about living abroad.

- ① ② ③ \* ( yes ) ( no )

23. Please pass me two pencil.

- ① ② ③ \* ( yes ) ( no )

24. Please send me more informations about your company.

- ① ② ③ \* ( yes ) ( no )

### 4.3 Appendix 2

Sex	Job	Nationality	Language	English Proficiency	Duration in Korea
F	14 Student	24 Korea	17 Korea	Very poor	4 Less than 1 year
M	13 Teacher	3 Malaysia	2 Malay	Fine	10 1-5 yrs
		2 UK	7 English	Very good	3 6-10 yrs
		2 USA	1 Korean/English	Excellent	2 11-15 yrs
		1 Canada		Native Speaker	8 16-20 yrs
		3 New Zealand			3 21-25 yrs
					1 26-30 yrs

# Development of Negation in Child Language and Learning Principles

Ho Han

Ajou University

hhan@ajou.ac.kr

## Abstract

Exploring the development of negation with respect to the answer to negative interrogative questions, I draw attention here to the explanatory adequacy of learning principles in child language acquisition, and present why Subset Principle does not hold for relevant acquisition data. Choi and Zubin (1985) report that both Korean children and English speaking children take three developmental stages in the acquisition of negation, regarding the answer to affirmative or negative interrogative questions. Both groups of children step through the same first and second stage of production of full propositional negation but diverge at the third stage, at which Korean children adhere to full propositional negation while English speaking children start to produce core propositional negation. Offering theoretical accounts of the developmental stages, Jang and Kwon (2008) deny Subset Principle and support Markedness Hypothesis as well as the Growth Theory in Chomsky's (1987) sense. In this article, I will point out potential problems of Jang and Kwon's arguments through reinterpretation of Choi and Zubin's findings, and argue that the Subset Principle does not hold in this case for a different reason.

## Keywords

Negation, subset principle, markedness

## Introduction

When we consider children's amazing ability of learning their mother tongue, we are puzzled as to what principles work so successfully in spite of poverty of input, immature perceptual and/or cognitive capacity, and poor communication strategies. This puzzling ability led generative linguists to one of the key tenets in their theories of language that children have innate linguistic knowledge, along with several broad assumptions<sup>1</sup>. This tenet is more or less based on the 'logical

problem of language acquisition' (Roeper & Williams, 1987) which questions how children acquire such complex linguistic system with scarce input of relatively poor quality.

There are several issues relevant to innateness of linguistic knowledge, some of which have been paid a lot of attention in the field. For instance, there has been much controversy over what learning principles are involved in the operation of learning mechanism. This study explores if Subset Principle, a learning principle, holds for the development of a particular type of negative constructions in child Korean and child English. The relevant data were originally presented in Choi & Zubin's (1985) longitudinal study and they were further discussed by Jang & Kwon (2008), who argue that Subset Principle does not work and markedness offers a better account for the case.

Presenting a detailed account of the data, I will support that Subset Principle is not operative for the case, but reject Jang & Kwon's line of reasoning. In addition, I will discuss which better explains the cases, Continuity or Maturation Hypothesis and support the latter, against Jang & Kwon's argument.

## 1 Subset Principle

Children with innate linguistic knowledge that is specified with Universal Grammar (UG) can overcome the 'logical problem of language acquisition' while processing triggering input. Children take a limited set of hypotheses about the target language and thus they entertain only a limited number of possibilities. This procedure follows the Subset Principle (SP) that children start with the most restrictive grammar and later change it when they encounter positive evidence in the input that the target grammar is different from their previous, hypothetical grammar. If children start with grammar larger than the target grammar, they could not get to the target grammar because there is no negative evidence. (Berwick, 1985; Wexler & Manzini, 1987; Becker, 2005). This is illustrated below in Figure 1.

<sup>1</sup> See Fodor and Sakas (2005: 514-515) for an extensive list of those assumptions.

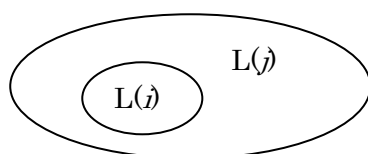


Figure 1: Subset Relation

First of all, it is obvious that children get to the value  $j$  directly in case the input that belongs to  $L(j)$  is given. However, encountering the input of  $L(i)$  and  $L(j)$ , children select  $i$  since it yields the fewer possibilities than  $j$ , and then they can stay with it or select  $j$  if they are given the input belonging to  $L(j)$ . But if  $j$  is selected, then  $i$  can not be selected because there is no negative evidence that allow children access to the more restricted set.

The effect of the SP is well demonstrated by Wexler & Manzini (1987) on setting a parameter of binding relation which is conditioned by the nature of binding domain. It is the governing category parameter that children have to deal with to process binding relation in their target language. The parametric values of the governing category are as follows:

- (1) The governing category parameter (Wexler & Manzini, 1987:53)

$\gamma$  is a governing category for  $\alpha$  iff  $\gamma$  is the minimal category which contains  $\alpha$  and

- a. has a subject; or
- b. has an Infl; or
- c. has a Tense; or
- d. has an indicative Tense; or
- e. has root Tense

Languages vary through (a) to (e). For instance, English is subject to (a) while Korean takes (e), which implies Korean is the language with more possibilities in governing categories than English. In other words, there is a subset-superset relationship between Korean and English in this regard. Both groups of children start with the smallest subset with the value (a). Children learning English stick to (a) because they are not given any input of other values, while Korean children go further toward (e) owing to positive input of all the possible values.<sup>2</sup>

The SP has been the most powerful learning principle in first language acquisition, although it

continues to be criticized in terms of the cost of obeying the SP in incremental learning (Fodor & Sakas, 2005), the existence of intermediate grammar (Angluin, 1980), and several cases of parameters with intersection relation rather than subset-superset relation between two languages (MacLaughlin, 1995). Now we turn to acquisition of negation to see if the SP is operative or not for the case.

## 2 Negation in Child Korean

Korean has two types of syntactic negation, as shown in (2)<sup>3</sup>:

- (2) a. Sue-ka Ed-lul mana-ss-e  
Sue-Nom Ed-Acc meet-Past-Decl  
'Sue met Ed.'
- b. Sue-ka Ed-lul **an** mana-ss-e  
Sue-Nom Ed-Acc not meet-Past-Decl  
'Sue didn't meet Ed.'
- c. Sue-ka Ed-lul mana-ci **an** ha-ess-e  
Sue-Nom Ed-Acc meet-Nmz not do-Past-Decl  
'Sue didn't meet Ed.'

(2a) is an affirmative sentence and (2a) and (2b) are its negative counterparts. (2a) is a short-form, preverbal negation (SFN), and (2b) a long-form, postverbal negation (LFN). SFN is yielded through insertion of *an* 'not' before the verb while LFN is formed through *ha* 'do'-support, attachment of the nominalizer *ci* to the verb stem, and insertion of *an* 'not' just before the *ha*- complex, which is often involved with contraction of *an* and *ha* to yield *anh*-.

Child Korean data show that children start to produce SFN around the age of 1;7 and LFN around the age of 3;5, although children at the SFN-only production stage can comprehend LFN (Cho & Hong, 1988; Hagstrom, 2002). The late production of LFN is attributed to its structural complexity (Han & Park, 1995) and relatively low frequency in motherese (Choi & Zubin, 1985)<sup>4</sup>.

A similar developmental order is also observed in child English as seen in (3) (Lason-Freeman & Long, 1991):

- (3) Stage 1: *no*+NP+VP  
Stage 2: NP+*no*+VP  
Stage 3: NP+*not*+VP  
Stage 4: NP+AUX+*not*+VP

<sup>3</sup> I will use the following abbreviations: Nom=Nominative, Acc=Accusative, Decl=Declarative, Nmz=Nominalizer

<sup>4</sup> In addition, semantic differences between SFN and LFN might be influential factors to the developmental order between them. See Han & Jang (2003) for this point.

<sup>2</sup> For other examples of the operation of the SP, see Hyams (1986) for the null subject parameter and Biberauer & Roberts (2007) for an SP account of diachronic change.

We see similarity in structure between LFN in child Korean and the structure at Stage 4 in child English, and late development of the both. Since frequency does not matter in the case of child English, processing difficulty due to structural complexity seems to decide the relevant developmental order in both child Korean and child English.

Standard syntactic negation in adult English includes the negator *not* and one in adult Korean the negator *an*. English takes the form similar to LFN in Korean whereas Korean allows two forms, SFN and LFN. The negation observed in Stage 2 and 3 in child English and SFN in child Korean is ‘simple’ negation, and one observed in Stage 4 in child English and LFN in child Korean is ‘complex’ negation, in a sense that the latter is involved with *do*-support. Now two possible subset relations (SRs) could be posited in accordance with the SP:

- (4) SR 1: Simple  $\supset$  Complex  
 SR 2: Complex  $\supset$  Simple

We posit SR1 when we account for language variation with respect to the form of structural negation and take the SP that children start with grammar with the fewest possibility. However, SR1 contradicts to the acquisition data in that children do not start with the complex negation in both languages. In contrast, SR2 conforms with the developmental order from ‘simple’ to ‘complex’ but contradicts subset relation regarding language variation. Above all, SR2 should explain why English speaking children start with the option that has never been given to them as input. Notice that selection of the correct SR converges on the basic question how children ‘know’ which are the subset and the superset value. I will get back to the issue later.

Either of SR1 and SR2 cannot be posited for the account of the development of structural negation in child language. There is another type of negation which appears as an answer to an interrogative question. We will see if the SP holds for this type of negation in the next section.

#### 4 Development of Yes-No Question-Answer Patterns

Considering *yes-no* questions and answers to them, there are four types of the *yes-no* question-and-answer pattern that are subject to the form of the question and the situation referred to, as shown in Table 1 (Choi & Zubin, 1985, p. 136).

Table 1: Possible Question-Answer Pairs

Question form	Situation depicted	
	(It is a cup.)	(It isn't a cup.)
a. <i>Is this a cup?</i>	True Affirmative (TA)	False Affirmative (FA)
b. <i>Isn't this a cup?</i>	False Negative (FN)	True Negative (TN)

The terms ‘True’ and ‘False’ refer to (non)correspondence between the literal form of the question and the situation depicted. The terms ‘Affirmative’ and ‘Negative’ describe absence and presence of the negative marker in the question, respectively.

Variation relevant to these patterns is observed between Korean and English, as follows (Choi & Zubin, 1985, p. 136):

Table 2: Answering Systems in Korean and English

	Question Form	(it is)	(it isn't)
Korean	a. i-ke kep-i-ya?	ing	ani
	b. i-ke kep-i an-ya?	ani	ing
English	a. Is this a cup?	yes	no
	b. Isn't this a cup?	yes	no

The difference resides in the answer to a negative question. In English the occurrence of *yes* or *no* in this case relies on the situation depicted, and thus the speaker is concerned with the question's core proposition (‘This being a cup’). On the other hand, in Korean it depends on whether or not the speaker agrees on the full proposition carried by the surface form of the question (‘This is a cup.’ or ‘This is not a cup.’)

Choi & Zubin (1985) reported a developmental pattern of this system in child Korean and in child English, presenting the data from their longitudinal study. This is shown in Table 3 (Choi & Zubin, 1985, p. 137).

Table 3: Developmental patterns

	Stage 1	Stage 2	Stage 3
Child English	No attention to neg. marker	Full proposition neg.	Core proposition neg.
Child Korean	No attention to neg. marker	Full proposition neg.	Full proposition neg

According to Choi & Zubin (1985), at Stage 1, since children do not pay attention to functional elements including the negative marker but content words, they take the negative questions as affirmative questions. At Stage 2, noticing the negative marker, both groups of children produce full proposition negation. At Stage 3 the two groups diverge and English-speaking children start to

produce core proposition negation. This is illustrated in the following table:

Table 4: Examples of the Patterns at Each Stage

	(Reference: a bird)	English	Korean
Stage 1	<i>Isn't this a cat?</i> <i>i-ke koyangi ani-ya?</i>	No	ani 'no'
	<i>Isn't this a bird?</i> <i>i-ke say ani-ya?</i>	Yes	ing 'yes'
Stage 2	<i>Is this a cat?</i> <i>i-ke koyangi-ya</i>	No	ani 'no'
	<i>Isn't this a cat?</i> <i>i-ke koyangi ani-ya</i>	Yes	ing 'yes'
Stage 3	<i>Is this a cat?</i> <i>i-ke koyangi-ya</i>	No	ani 'no'
	<i>Isn't this a cat?</i> <i>i-ke koyangi ani-ya</i>	No	ing 'yes'

Kwon & Jang (2008) account for the pattern illustrated in Table 3 and 4, and argue that since Korean is [+full proposition/-core proposition] and English [-full proposition/+core proposition] in answering patterns, no subset relation can be established between the two languages. They contend that the SP cannot explain the relevant developmental pattern.

I agree with Kwon & Jang (2008) that the SP does not operate in this case. But my argument is based on a different line of reasoning. First of all, take a look at the following examples:

(5) Question: (Referring to a bird)  
i-ke koyangi ani-ya?  
This-thing cat not-be  
'Isn't this a cat?'

Answer: (i) ing  
Yes  
'Yes/Right'  
(ii) ing. koyangi an-ya  
yes cat not-be  
'Yes/Right. It's not a cat.'  
(iii) koyangi an-ya  
cat not-be  
'It's not a cat.'

There are three alternatives for the answer to the negative question in Korean. Furthermore, in case of (ii), *ing* 'Yes' and *say-ya* '(It) is a bird' do not constitute one sentence, unlike in English, as shown in (6):

(6) Question: Isn't this a cat?  
Answer: (i) No.  
(ii) No, it's not.  
(iii) It's not a cat.

The (5)-(ii) in Korean actually conveys two separate propositional contents, 'I agree with you. (That's right)' and 'It is a bird', and in other types of

answers in (5), the part that expresses (dis)agreement is optional. There must be a [+/- (dis)agreement] feature working in the *Yes/No* question-answer pattern. Korean allows either of + or – value of the relevant parameter, while English is specified as a [-(dis)agreement] language. Regarding this sort of language variation we can posit a subset relation as in (7).

(7) Korean = { +(dis)agreement, -(dis)agreement }  
English = { -(dis)agreement }  
∴ Korean ⊃ English

The SP predicts that children start with the subset, the [-(dis)agreement] option. But this prediction is not borne out according to the relevant developmental order: Children start with the superset item, the [+(dis)agreement] option.

We might conclude that the SP is not operative in the development of negation including the structural negation and the *Yes/No* question-answer pattern. Then, how can we explain the developmental pattern discussed above? This is what we turn to.

## 5 Markedness and Growth Theory

As mentioned in Section 1, one key question about the SP is how children know which are the subset and superset value at the onset. Manzini & Wexler (1987) suggest that children have an inherent system of default values for parameters that correspond to the subset value. The default value seems to be the one that arises from crosslinguistic variation in their argument, considering the subset relation in the binding parameter. As we have seen above, however, this is not the case for the developmental pattern of child negation.

An alternative explanation is that children start with an unmarked option. What is the nature of an unmarked element or feature, in general? A broad claim is that a structural property is taken to be unmarked, if it has relatively wide distribution across languages, appears at the initial state of the diachronic change, or occurs at the initial state of language acquisition. Then, we might be able to contend that children start with the SFN-like negation construction and the [+(dis)agreement] question-answer pattern because they are unmarked.<sup>5</sup>

Kwon and Jang (2008) support the Growth Theory that language faculty grows through the developmental stage of language acquisition. Taking the developmental order of the *Yes/No*

<sup>5</sup> To prove this point, we may need to investigate their crosslinguistic distribution and diachronic change, which I will leave for the future research.

question-answer pattern, they argue that at Stage 1 the [+/- full proposition, +/-core proposition] parameter is unset since children's linguistic faculty is not mature, and later it is set as the faculty grows. But their argument do not explain why English speaking children erroneously set the parameter on [(dis)agreement] first with no triggering input. Have English speaking children not yet had matured linguistic faculty until stage 2? If not, why do they choose that option? Is it because the option is 'simple' and/or 'easy'?

I argue that what matures is children's processing capacity, not their linguistic faculty on a par with Han (1997). According to Fodor & Crain (1987), there is a positive correlation between simplicity and generality: the simpler, the more general. If children begin with unmarked options, they take simpler and more general options. At Stage 1, children could not 'recognize' the negative marker and at Stage 2, they set on a simpler unmarked option since it is easier to process. Later at Stage 3 when children's processing faculty grows and thus children can process triggering input, English-speaking children take the adult form with [-(dis)agreement].

The same line of reasoning applies to the development of structural negation. SFN-like constructions are syntactically easier to process than LFN-like constructions, and thus Korean children start with easier and simpler ones or English speaking children cannot produce the target form with *do*-supported involved.

## 6 Conclusion

There has been a consensus in previous research of child language acquisition that learning principles operate together with innate linguistic knowledge on triggering input (Wexler & Culicover, 1980). Learning principles, which are logical and computational in their nature, presumably apply crosslinguistically, predicated on its universal cognitive properties. In this study, SP, one of the learning principles, was attested in terms of its applicability, regarding the developmental pattern of certain types of negation in child language. I have drawn attention to problems about the operation of SP, as I argued that the children do not start with the subset but rather with unmarked option for the case.

On exploring learning principles, we hope that they can work overall in the course of language acquisition to attain a sort of generalization. We expect that the more generalizable and thus powerful a principle is, the nearer it gets to explanatory adequacy. SP is a powerful learning principle, which holds for many cases in language acquisition, but why is it not operative in the

acquisition of negation discussed above? There are some possible explanations to this question. First, a markedness hierarchy might override SP. As mentioned before, it is highly potential that unmarked elements constitute subset. But that might be not always the case, considering marked subset elements. Second, notice that SP in language acquisition is successfully applied to most of the cases which pertain to development of syntactic properties. Negation is also involved with syntactic mechanism under generative grammar, but it is more strongly associated with semantic/pragmatic properties. I conjecture that there might be a distinction between syntactic and pragmatic learning principles in terms of applicability of SP. I leave these possible explanations for future research.

## References

- Angluin, D. (1980). Inductive Inference of Formal Languages from Positive Data. *Information and Control* 45, 117–135.
- Biberauer, T. & Roberts, I. (2007). The return of the Subset Principle. Ms. University of Cambridge.
- Becker, M. (2005). Raising, control and the Subset Principle. In J. Alderete, C.-H. Han, & A. Kochetov (Eds.), *Proceedings of the 24<sup>th</sup> West Coast Conference on Formal Linguistics* (pp. 52-60). Somerville, MA: Cascadia Proceedings Project. [www.lingref.com](http://www.lingref.com), document #1206.
- Berwick, C. R. (1985). *The Acquisition of Syntactic Knowledge*. Cambridge, MA: MIT Press
- Cho, Y.-M. Y. and Hong, K.-S. (1988). Evidence for the VP constituent from child Korean. *Papers and Reports in Child Language Development* 27, 31-38.
- Choi, S. & Zubin, D. (1985). Acquisition of negation: How do children learn to answer yes-no questions in Korean and in English. In Kuno, S., Whitman, J., Lee, I.-H., & Kang, Y.-S. *Harvard Studies in Korean Linguistics I* (pp. 135-144). Cambridge, MA: Department of Linguistics, Harvard University.
- Goldberg, A. (2005). *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.
- Fodor, J. D. & Crain, S. (1987). Simplicity and generality in rules of language acquisition. In B. MacWhinney (Ed.), *Mechanisms in Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates
- Fodor, J. D. & Sakas, W. G. (2005). The Subset Principle in syntax: costs of compliance. *Journal of Linguistics* 41, 513-569.

- Hagstrom, P. (2002) Implications of child errors for the syntax of negation in Korean. *Journal of East Asian Linguistics* 11, 211-242.
- Han, H. (1997). Development of Functional Categories in Child Korean. Doctoral dissertation, University of South Carolina.
- Han, H. and Park, M.-K. (1995). The syntax of negation and its development in child language. In J. Fuller et al. (Eds.), *ESCOL '94* (pp. 152-162). Ithaca, NJ: CLC Publications, Cornell University.
- Han, H. & Jang, Y. (2003). An analysis of negation in child Korean: An alternative to the UCC account. *Korean Journal of Applied Linguistics*, 19(1), 41-58
- Jang, Y. & Kwon, Y.-K. (2008). Acquisition of negation in Korean and English. *Korean Journal of Linguistics* 32(4) 735-749.
- Larson-Freeman, D. & Long, M. H. (1991) *An Introduction to Second Language Acquisition*. New York: Longman.
- Legate, J. and Yang, C. (2002). Empirical Re-Assessment of Stimulus Poverty Arguments. *The Linguistic Review* 19, 151-162.
- MacLaughlin, D. (1995). Language acquisition and the Subset Principle. *Linguistic Review* 12(2), 143 – 191.
- Manzini, R. & K. Wexler (1987). Parameters, Binding Theory, and Learnability. *Linguistic Inquiry* 18, 413 – 444.
- Thompson, S. P. & Newport, E. L. (2007). Statistical learning of syntax: the role of transitional probability, *Language Learning and Development* 3, 1–42.
- Roeper, T. & Williams, E. (1987). *Parameter Setting*. Dordrecht: D. Reidel.
- Tomasello, M. (2003). *Constructing a Language: A Usage-Based Approach to Language Acquisition*. Boston, MA: Harvard University Press.
- Wexler, K. & Culicover, P. W. (1980). *Formal Principles of Language Acquisition*. Cambridge, MA: MIT Press.
- Wexler, K. & Manzini, R. (1987). Parameters and learnability in binding theory. In T. Roeper & E. Williams (Eds.), *Parameter Setting* (pp. 41-76). Dordrecht: D. Reidel.

# Insertion Preferred to Deletion in Learner Speech: A Study of Korean English

Gyu-jeong Han and Jong-mi Kim

Department of English Language and Literature, Kangwon National University

lemon0603@naver.com, kimjm@kangwon.ac.kr

## Abstract

Do language learners prefer insertion or deletion when forced to produce the target speech that is not permissible in native phonology? To answer this question, we analyze the speech data from 3,256 word data of L1 English, L2 English, and L1 English loanwords, which were elicited from 124 Korean adult learners of American English and 22 native speakers of Korean and English. We measured the acoustic features that are relevant to insertion and deletion in English. We focus on whether non-native English evidence effects of native Korean more in the deletion of a consonant as in [concept] for the English loanword 'concept', or in the insertion of a vowel as in 'concept' [concept]? The results indicate that the learner speech manifests significantly more insertion errors than deletion errors, in both word-level and sentence-level production. The error rate drops drastically after instruction to the extent that the difference between the insertion and deletion becomes less obvious. The results support that an explicit instruction is effective to phonological acquisition of the learner speech.

## Keywords

English loanwords into Korean, Korean learner speech of English, insertion and deletion in learner speech, loanword phonology

## 1 Introduction

We examine whether insertion or deletion occurs when a learner is forced to produce the target speech that is not permissible in native phonology. To argue for the insertion preference, adult language learners (L2 learners, henceforth) may have stronger force of adding speech elements when conveying the intended information to the listener, as their speech risks at all times of lack of intelligibility by non-native accents (Jenkins, 2000:69-98). Apparently, loanwords manifest insertion over deletion of the target segments (See *Official Loanword Transcription Regulations*,

1995). We thus make a working hypothesis that insertion is preferred to deletion in learner speech.

(1) HYPOTHESIS 1: Insertion is preferred to deletion in learner speech.

Hypothesis 1 is confirmed, or strictly speaking, remains provisional in Korean learners' speech of L2 English when two conditions are met: (1) the loanwords with an inserted segment interfere the learner speech more frequently than the loanwords with a deleted segment; and (2) the loanwords with both an insertion form and a deletion form observe more of inserted forms than those of deleted forms.

The validity of the Hypothesis 1 interplays with the developmental aspect of learner speech, as learner speech is assumed to include less of non-native accent as the L2 proficiency of the learner increases over time. We thus make another working hypothesis that learner speech observes reduced number of insertion or deletion cases according to the increase of the proficiency.

(2) HYPOTHESIS 2: The rate of insertion and deletion decreases in learner speech according to the increase of the proficiency.

Hypothesis 2 can be confirmed if a cross-sectional or longitudinal study in learner speech shows a developmental transition of decreased insertion and deletion in L2 words.

The hypothesis on developmental enhancement, however, is controversial in literature, as Scovel (1988) claims that adult learners improve insignificantly in foreign pronunciation. We, therefore, investigate our second question: Whether or not learner speech develops from the form of loanword phonology toward the form of target speech.

We investigate the acquisition order by using both a cross-sectional and longitudinal methods. The longitudinal method is to test cognitive and

affective factors in adult language learning -- such as the role of attention and the importance of motivation in learning, as Schmidt (1995) claims, "In order to acquire phonology, one must attend to phonology" (Schmidt, 1995: 17).

The following sections test these two hypotheses by means of a set of acoustic and perceptual phonetic experiments.

## 2 Method

In order to accomplish our objective, we administered the phonetic experiment that comprises four sub-experiments on native and learner speech in cross-sectional and longitudinal studies. The four sub-experiments are on: 1) Korean native speech of English loanwords in Korean context, 2) English native speech of the English loanwords in English context, 3) Korean learner speech of the English loanwords in English context by a cross-sectional study, 4) Korean learner speech of the English loanwords in English context by a longitudinal study. Each sub-experiment used different sets of speech

materials, participants, and the data acquisition steps.

### 2.1 Speech materials

Three types of speech materials and the recording lists were used: 1) recorded L1 Korean speech materials of English loanwords in Korean context, 2) recorded L1 and L2 English speech materials of English words in English context, 3) listening test materials in English for Korean learners to serve as co-variance reference to the development in speech production.

The listening test materials were the listening component of *Test Of English for International Communication* (TOEIC). Two TOEIC listening tests were used as a standardized test to recognize the improvement of listening comprehension before and after instruction.

For recorded speech materials, the native and learner participants read words in isolation, and those embedded in English or Korean contexts. Table 1 is the recording list, in which the insertion and deletion to contrast are underlined. The Korean context is romanized according to the *Official Romanization System* of Korean.

Table 1: Recording list for insertion and deletion in L2 English words, spoken by L1 Korean speakers

Loanword phonology	I. Words in isolation	II. Words in English context	III. Words in Korean context
(a) Insertion	1. <u>d</u> [ ]ry 2. <u>app</u> [ ]ly 3. <u>app</u> [ ]lication 4. bus[ ]ness	1. I'll <u>d</u> [ ]rink this wine. 2. I wat <u>ch</u> [ ]ed TV last night <u>t</u> [ ]. 3. He <u>app</u> [ ]lied for the position. 4. Fill in the <u>app</u> [ ]lication form. 5. Tell me about your bus[ ]ness. 6. I got <u>off</u> [ ] the <u>grass</u> [ ].	1. igeoseun <u>d</u> [ ]rinkeuimnida. (This is a drink.) 2. uri nait[ ]je gaja. (Let's go night club.) 3. nanun geu hoesae eop[ ]lai haetta. (I applied for the company.) 4. miguk daehage eop[ ]likeisyeoneul bonaeryeogo hamnida. (I want to send the application for auniversity in US.) 5. geuuy big[ ]niseu suwaneun jota. (His ability for business is good.) 6. na oneul geunmuneun <u>off</u> [ ]ya. (I'm off duty today.)
(b) Deletion	1. sta <u>r</u> e	1. The <u>r</u> e was no confirmation about the policy.	1. taieo[ ]je gumeongnasseo. (The tire is flat.)
(c) Insertion and Deletion	1. violent <u>t</u> [ ] 2. stamp[ ] 3. ax <u>e</u> [ ] 4. act[ ]	1. How did you paint[ ] this wall? 2. I played tennis last[ ] week.	1. paint[ ] chil da haenni? (Did you finish painting?) 2. ibuni last[ ] gaimiya. (This is the last game.)

Words in Table 1 are subject to insertion for the data in (a), and deletion for the data in (b) if used as a Korean loanword, while those data in (c) are subject to insertion or deletion. When eliciting the

speech of learners using the list in Table 1, we randomized the words among themselves, and the sentences among those with the similar syllable numbers.

The recording lists in the first and third columns were used to obtain L1 Korean speech materials of English loanwords in Korean context, while the lists in the first and second columns were to obtain L1 and L2 English speech materials of English words in English context.

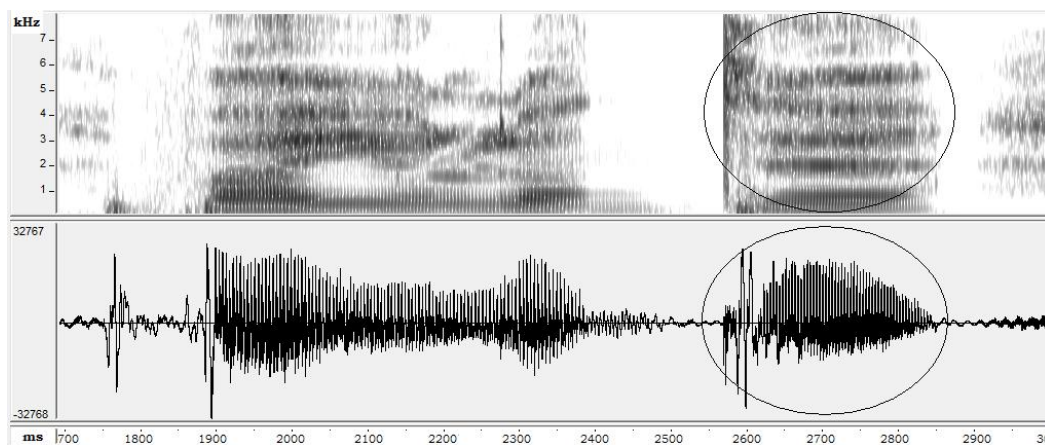


Figure 1: Illustration of an incorrectly inserted vowel [ɪ] at the end of the English word *violent*, spoken by a female Korean learner.

## 2.2 Participants

Research participants consisted of 11 Korean native speakers who read English loanwords in Korean context, 11 native speakers of American English and 124 Korean learners of English who read the English loanwords to Korean in English context.

All native speakers spoke the General American English, while learners the Standard South Korean as their native language. The learner's English proficiency level varied with respect to the scale of five ranks in accordance to TOEIC score. We then putatively used the scaling of score intervals in *TOEIC Can-Do Guide* (2000) as co-variance units for our analysis.<sup>1</sup> This follows the developmental effects that the more proficient the speaker is, the more target-like the perception of the target sound is for the listener.

## 2.3 Data acquisition procedure

The data acquisition procedure varied for all four sub-experiments. For the first sub-experiment, where we acquired Korean native speech of English loanwords, we used the following three stages of eliciting the reading speech. At the first stage, the Korean native speakers were given the English word list in English alphabet, and asked to transcribe the words into Korean alphabet. The transcription was done for the words in isolation. At the next stage, the speakers were given the Korean sentence list where the English loanwords are embedded hidden in empty blanks. The speakers

are then asked to fill in the blanks by copying down their own previously transcribed English loanwords in Korean alphabet. At the last stage, the speakers were asked to read the completed Korean scripts for recording the words in isolation and the words embedded in sentences.

For the second sub-experiment, where we acquired English native speech of the English loanwords in English context, the native speakers of American English underwent only one stage of the recording of the production stimuli.

For the third sub-experiment, where we acquired Korean learner speech of the English loanwords in English context by a cross-sectional study, the learners took listening and production tests after a brief listen and repeat practice of the recording stimuli, on a very first week of any class instruction to avoid experimental intervention.

For the fourth sub-experiment, where we acquired Korean learner speech of the English loanwords in English context by a longitudinal study, the procedure consisted of three stages: 1) pre-listening and production tests after a brief listen and repeat practice of the recording stimuli, 2) pronunciation instruction in class for 7 weeks that included many different aspects of pronunciation including, but not focused on, avoiding the Korean l-alternation and applying the English de-aspiration rules, and 3) post-listening and production tests on the same data after 7 weeks, and again on the different data after next 7 weeks.

## 2.4 Analysis

We measured the acoustic features that are relevant to insertion and deletion in English: the insertion in terms of the vowel formants as displayed in the spectrogram and the deletion in terms of the absence of burst and longer consonant duration.

In Figure 1, the vowel [ɪ] is inserted after the final consonant [n] in the word *violent* [naɪnʃiəntɪ].

<sup>1</sup> This edition of the *Can-do Guide* fits to our learner speech data, as we used the previous version of TOEIC listening test using the General American English. The next edition was published on 2009 to reflect the recent modification in TOEIC to contain different English accents around the world.

For analysis, we compared the learner speech with the native speech. For less clear cases of acoustic features, we used the additional means of the perceptual judgment by researchers and native speakers of English. We counted only the expected errors that are shown in the loanwords for the different L2 phones in question.

### 3 Results

A total of 3,256 word data were acquired from the recorded corpus of native and learner speech in word-level and sentence level production by 124 Korean learners of English and 11 Korean and 11 English native speakers.

The results from the cross-sectional study are shown in Figure 2 for word-level production and Figure 3 for sentence-level production. These figures represent the interference error rate in percentage with deletion (a), insertion (b), and insertion and deletion (c).

In both Figures 2 and 3, the relative frequency of insertion and deletion interference is higher in

insertion, as the values in (b) are generally higher than (a). The relative dominance within Figure 2(c) and 3(c) will soon be discussed in Figure 4, to be higher in insertion as well. The results so far in (a) and (b) support our Hypothesis 1 that insertion is preferred to deletion in learner speech.

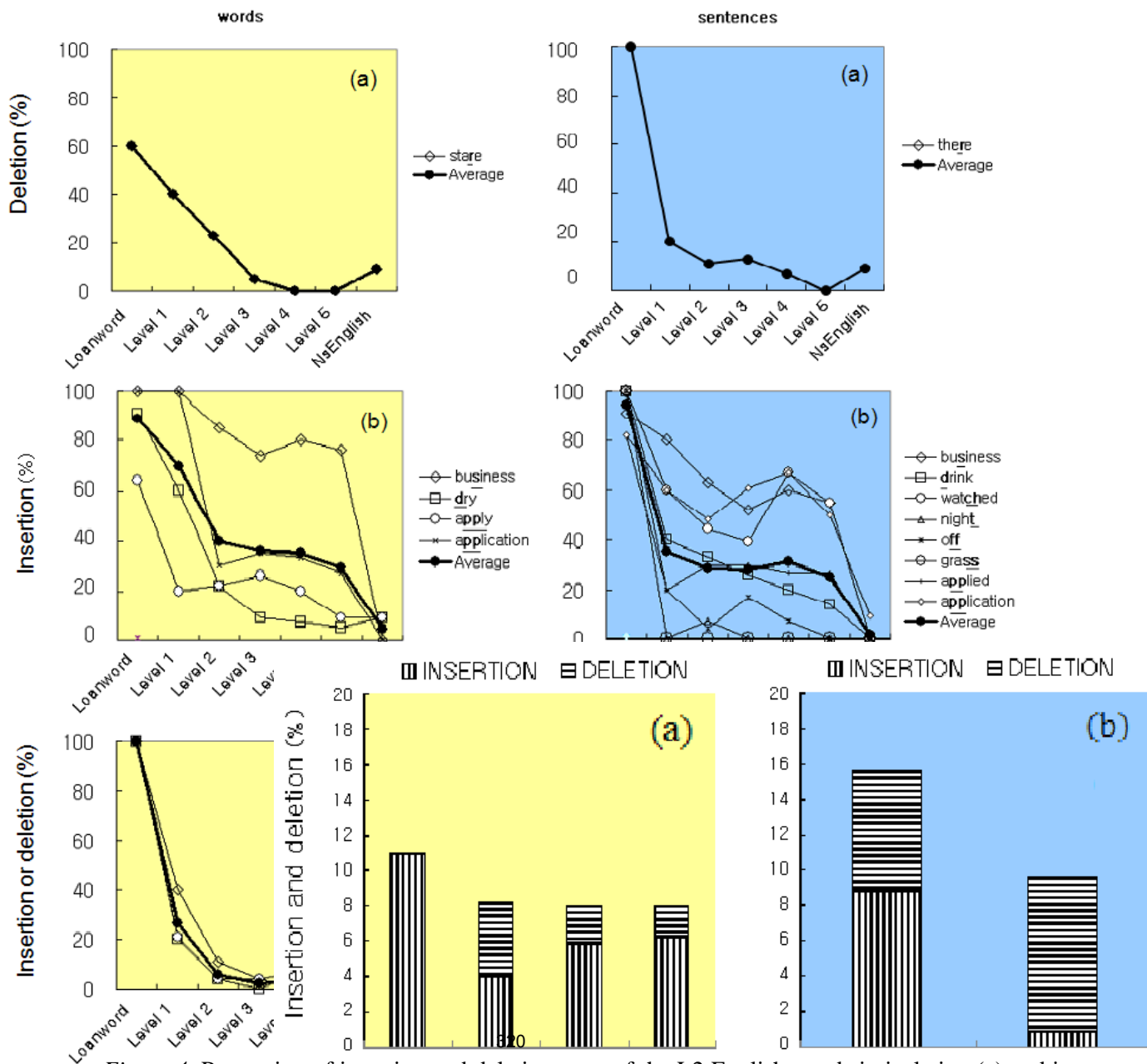


Figure 4: Proportion of insertion and deletion rates of the L2 English words in isolation (a) and in sentences (b), spoken by L1 Korean learners.

To move on to the developmental results in relation to Hypothesis 2, Figure 2(a) and 2(b) for word-level production supports the hypothesis that error rate decreases as the speaker proficiency increases. The sentence level production in 3(a) and 3(b) is less obvious.

The relative dominance between insertion and deletion within Figure 2(c) and 3(c) are shown in Figure 4. As graphically demonstrated in 4(a) for word-level production, and 4(b) for sentence-level production, insertion is more dominant over deletion except a single case of the sentence level production of the word *last*. The results thus support our Hypothesis 1 that insertion is preferred to deletion in learner speech.

The relative frequency and the developmental results from cross-sectional study in Figures 2 and 3 are compared to the results from the longitudinal study in Figure 5.

In Figure 5(a) for words in isolation, the average values before instruction shows significantly greater rate of insertion than deletion. The same is true in Figure 5(b) for words embedded in sentences. The higher rates of insertion over deletion are unexceptionally found in all stages of development. Our results in overall thus clearly support the Hypothesis 1 that insertion is preferred to deletion.

In the developmental results from a longitudinal study before and after instruction in Figure 5, the error rate in learner speech decreased over time. The post-test results in Figure 5 show the decrease regardless when the test is with the same data and the different data. The post-test was taken after seven weeks of interval; with the same data after the first seven weeks, and with the new data after the second seven weeks of instruction. The results are similar both for the words in isolation in 5(a) and the words embedded in sentences in 5(b). The developmental results thus support Hypothesis

2 that the rate of insertion and deletion decreases in learner speech according to the increase of the proficiency.

## 4 Discussions

The results indicate that the learner speech manifests significantly more insertion errors than deletion errors, in both word-level and sentence-level production. In addition, the error rate decreases according to the speech proficiency. Furthermore the error rate drops drastically after instruction to the extent that the difference between the insertion and deletion becomes less obvious. The results support that an explicit instruction is effective to phonological acquisition of the learner speech.

Additional findings in this study are the following. Firstly, learner speech is far more distant from loanwords, and much closer to target phonology. In other words, learner speech is far more faithful to target phonology than loanword phonology. Secondly, another difference between learner speech and loanword phonology is that learner speech develops toward target phonology, while loanword phonology is rather constant. Thirdly, needless to mention is an additional fact that loanwords conform to native phonotactics, whereas learner speech does not.

Currently, the discussions in this study are based on the average values of a large number of data. We plan to do a more sophisticated analysis with statistical significance.

## Acknowledgements

This work was supported in part by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (KRF-2006-321-A00933).

## References

ETS. (2000). *TOEIC Can-Do Guide: Linking*

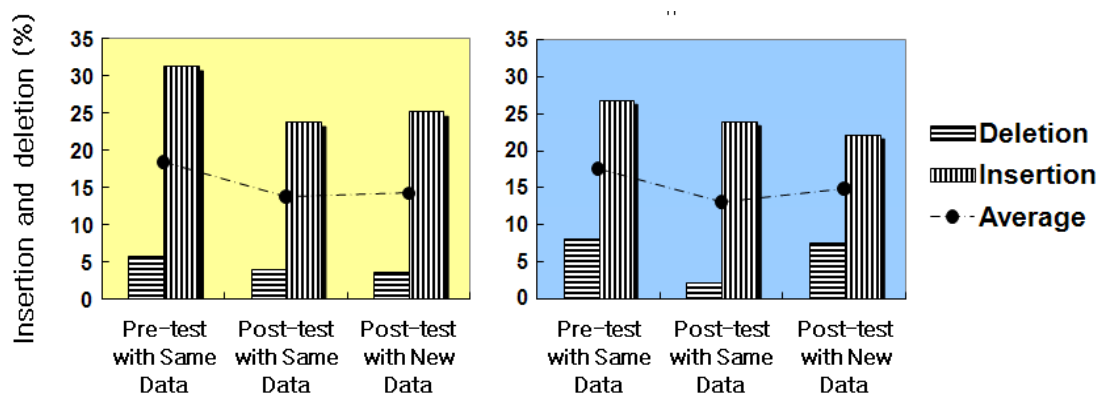


Figure 5: The post-test results show the decrease of insertion and deletion of segments in L2 English words spoken by L1 Korean learners, regardless whether the test is on the same data, on the different data, for the words spoken in isolation (left), or for the words spoken in sentences (right).

- TOEIC Scores to Activities Performed Using English*. (2000). Princeton, NJ: Author. [TOEIC\_Can\_Do.pdf, www.ets.org]
- Jenkins, J. (2000). *The Phonology of English as an International Language*. Oxford University Press.
- Official Loanword Transcription Regulations*. (1995). [The current edition in official use]. Ministry of Culture and Tourism, Republic of Korea. [<http://www.korean.go.kr>].
- Official Romanization system*. (2000). [The current edition in official use]. Ministry of Culture and Tourism, Republic of Korea. [<http://www.korean.go.kr>].
- Scovel, T. (1988). *A Time To Speak: A Psycholinguistic Inquiry Into The Critical Period For Human Speech*, New York: Newbury House.
- Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in foreign language learning. *Attention and awareness in foreign language learning* ed. by R. Schmidt, 1-63. Honolulu, HI: Univ. of Hawai'i. Second Language Teaching and Curriculum Center.

# The Effects of Dictation Practice in English Listening Classes

Myeong-Hee Seong<sup>1</sup> and Eun-Hee Nam<sup>2</sup>

<sup>1</sup>Eul-Ji University, <sup>2</sup>Chung-Ang University

seong@eulji.ac.kr, hieuni@hanmail.net

## Abstract

This study investigated the effects of dictation practice on Korean university students, in terms of listening proficiency and the use of listening strategy. The research was implemented for 15 weeks with 89 Korean university students. The subjects were divided into two groups. One of the groups functioned as the experimental group and the other as the control group. All conditions were the same in both groups, except that the experimental group had dictation practice on a regular basis in their listening classes through one semester. A sample TOEIC listening test was administered as a pretest and posttest. A questionnaire was used to find out the differences in the use of listening strategies between the two groups. On the basis of the results, the study suggested some guidelines for dictation practice in EFL listening classes, and called for more studies on its effects.

## Keywords

Dictation practice, Listening proficiency, Listening strategy

## Introduction

Teaching English is now one of the most important subjects in most Asian schools. This is because it is important for students to possess the English ability to communicate with other English users, with the growing of English as a lingua franca. Among 4 skills of English, listening is regarded as the language skill which learners usually find the most difficult, while it is regarded as a fundamental factor for English learners.

Listening has received increasing attention from EFL professionals in recent years. To find out the effective methods of enhancing English listening skill, a large number of researches have been attempted. One of them is dictation, and it has long been considered a useful method to improve learners' listening proficiency. Many methodologists have often made pedagogical claims for its value. As the ground of its effect on language learning, Lambert (1986) insists that a single input

(visual or auditory) is less effective than double input (visual and auditory) in a language comprehension. Despite their claims, it is not widely used in EFL classrooms, and there is also another voice that dictation is inefficient (Cartledge, 1968; Valette, 1964).

Under the circumstances, there have not been many studies concerning the effects of dictation, and it is necessary for English education professionals to examine the effects of dictation in EFL listening classes. For this reason, this study investigated the effects of dictation practice, in terms of listening proficiency and the use of listening strategy. For the purpose of the study, two research questions were set. They were as follows;

- 1) Does dictation practice improve listening proficiency?
- 2) What are the differences in the use of listening strategies between two groups? Does dictation practice make the participants use different listening strategies?

## 1 Theoretical Background

### 1.1 Dictation in Listening Classes

Dictation is an activity to write something that someone says or reads out as it is being said or immediately after it is said. Education literature suggests promising ways to use dictation in the language classroom, and dictation has been an exercise consistently employed throughout the history of language teaching. Davis and Rinvolucri (1988) write that decoding the sounds of English and recoding them in writing is a major learning task. Pappas (1977) considers dictation as a good means of developing learners' listening comprehension. Byrne (1977) advocates dictation as a useful pedagogical technique, and he believes that dictation involves listening and the ability to transform what is heard into its written form. Stanfield (1985) mentions that since in dictation the learner employs more than one faculty, s/he is more successful in internalizing the language. In addition, Frodesen (1991) suggests that dictation can be an effective way to address grammatical errors in

writing that may be the result of erroneous aural perception of English. She also says that if done systematically and regularly, dictation exercises improve students' ability to distinguish sounds in continuous speech as well as improving their spelling and their recognition of grammatically correct sentences and their production of them.

## 1.2 Listening Strategies as one of Language Learning Strategies

Oxford (1990) defines language learning strategies as specific actions taken by the learners to make learning easier, faster, more enjoyable, self-directed, more effective, and more transferable to new situation. Researches and theories in the field of ESL/EFL strongly imply that strategies obviously play an important role in learning English. Many researchers have discovered that good learners are strategic, meta-cognitively active in selecting strategies, and monitor their process in order to successfully accomplish given tasks (Macaro, 2001; O'Malley & Chamot, 1990; Oxford, 1990; Rubin, 1975). Phillips (1991) suggests that students' use of various learning strategies positively affect their command over new language skills.

Furthermore, researchers strongly implied that language learning strategies can be taught, and they have identified and classified their own learning strategies (Chamot, 1987; Oxford, 1990; Rubin, 1987). Oxford (1990) divided language learning strategies into direct and indirect strategies. She classified in detail them into six categories: memory, cognitive, compensation strategies were included in the direct strategies while metacognitive, affective, and social strategies were included in the indirect strategies. This study explores the students' use of listening strategies on the basis of Oxford's.

## 2 Method

### 2.1 Subjects

The subjects were 89 freshmen and sophomores attending a university in Korea. They were divided into two groups. These students were all non-English majors who enrolled in TOEIC classes as an elective; one of the groups functioned as the experimental group and the other as the control group.

### 2.2 Measures

#### 2.1.1 Listening proficiency

A sample TOEFL listening comprehension test (L/C) was used to measure participants listening proficiency. The test consists of 100 questions with four different parts. Even though a standard score of TOEFL listening comprehension ranges between 5

to 495, the score for the listening test of this study ranges from 1 to 100 for convenience in calculation.

#### 2.1.2 Listening strategies

A questionnaire was used to find out the differences in the use of listening strategies between the two groups. The questionnaire consisted of 6 categories according to Oxford's language learning strategies. It included 50 items modified to measure listening strategies for the purpose of this study. Each item presented on a 5 point Likert scale (1=never to 5=always) and it was written in Korean to help students understand clearly.

### 2.3 Procedures

The research has been implemented for 15 weeks. All conditions were the same for the experimental group and the control group, except that the experimental group had dictation practice on a regular basis in their listening classes throughout one semester.

At the beginning of the semester, a sample TOEIC listening test was administered as a pretest, and the same one was used as a posttest at the end of the semester. Right after the posttest, a listening strategy questionnaire was distributed to each student, and collected after 15 minutes.

### 2.4 Treatment

Both groups of students were met once a week in two consecutive 50-minute classes, with one hour for listening and one hour for reading section. In the listening section, the experimental group was asked to write what they hear from the recorded tapes. For this study, the researcher prepared a handout for dictation in each class. Since the study was conducted in TOEIC class, the participants were given different kinds of dictation practices according to the part of TOEIC listening test. The participants wrote down sentences for part 1, filled in the blanks for part 2 and 3, and used note taking for part 4.

## 3 Results

### 3.1 Listening Proficiency

The mean scores of control group and experimental group on the pre-test were compared to see whether they were same or different before the experiment started. Table 1 shows the result of T- test.

Table 1: Listening Proficiency on the Pre-test

Group	N	M	SD	t	sig
Experimental	40	44.3	11.85	2.27	.026
Control	40	50.4	12.24		

\* p<.05

Levene's test:  $F=0.03$ ,  $p=.867$

According to Table 1, there was no significant difference between the two groups on the pre-test. It can be concluded that both groups are nearly homogeneous, and the two groups started with the same proficiency level. Table 2 illustrates the results of T-test for the performance of the two groups on the post test.

Table 2: Listening Proficiency on the Post-test

Group	N	M	SD	t	sig
Experimental	41	54.1	8.80	1.44	.154
Control	45	57.2	11.12		

\*  $p<.05$

Levene's test:  $F=0.03$ ,  $p=.867$

As the above results indicate, there was no significant difference in the results of the post-test between the two groups. However, considering the mean scores of the two groups, the experimental group significantly performed better ( $N=41$ ;  $M=54.1$ ) than the control group ( $N=45$ ;  $m=57.2$ ) in the post-test. In other words, the experimental group who had regular practice with dictation made more improvement in their listening proficiency than the control group.

### 3.2 Listening Strategy Use

Table 3 presents the results from T-test of the overall listening strategy use between the two groups.

Table 3: Comparison of Overall Listening Strategy Use

Group	N	M	SD	t	p
Experimental	36	2.94	0.46	-1.2	0.73
Control	39	2.76	0.40		

\*  $p<.05$

Regarding the two groups' use of overall listening strategies, in terms of the frequency of use, there was no significant difference. Table 4 presents the results of T-test of direct listening strategies use: memory, cognitive, and compensation listening strategy.

Tables 4: Comparison of Direct Listening Strategies Use

Strategy	Group	N	M	SD	t	p
Memory	Experimental	38	2.87	0.50	-1.01	.316
	Control	44	2.75	0.56		
Cognitive	Experimental	39	3.05	0.46	-1.02	.313
	Control	45	2.94	0.46		
Compensation	Experimental	38	3.37	0.65	-0.93	.356
	Control	44	3.25	0.63		

\*  $p<.05$

Data about direct listening strategies use shows that the experimental group appeared to use them almost same or a little more than control group, and there was no statistical difference in the use of three direct listening strategies between the two groups. Table 5 presents the results of T-test of indirect listening strategies use: metacognitive, affective, and social listening strategy.

Tables 5: Comparison of Indirect Listening Strategies Use

Strategy	Group	N	M	SD	t	p
Metacognitive	Experimental	39	2.72	0.61	-2.11	.037*
	Control	45	2.44	0.56		
Affective	Experimental	38	3.01	0.56	-0.86	.390
	Control	44	2.97	0.53		
Social	Experimental	39	2.49	0.63	-1.33	.187
	Control	45	2.28	0.82		

\*  $p<.05$

Among the indirect listening strategies, there was a significant difference between the two groups. More precisely, the experimental group used metacognitive listening strategy more than control group ( $p=.037$ ) even though there was no significant difference.

On the whole, the most frequently used strategy was compensation strategy for both experimental and control group, and social strategy was the least used one for both of the two groups.

## 4 Conclusion

This study investigated the effect of dictation practice in English listening classes at university. For the purpose of this study, the experimental group took a dictation in every session of their listening class for a period of one semester.

In terms of the first research question, 1) Does dictation practice improve listening proficiency?, there was no statistically significant improvement. The results of the study support those of Jafarpur and Yamini (1993) in that they saw no improvement in the proficiency of the experimental group as compared with the control group. However, it is necessary to note that the present study showed the experimental group made more improvements in their listening proficiency than the control group. A comparison of the mean scores on the pre-test and post-test displayed that experimental group gained almost 10 points higher on the post-test than the pre-test, whereas the control group gained 6.7 points higher.

With regards to the improvement listening proficiency, this study has a couple of shortcomings. First, the time spent on the dictation practice was

just 50 minutes a week. Second, the post-test was given only 13 weeks after the pre-test. Besides, the 'incubation hypothesis' might have been partly disregarded in this study. Jafarpur and Yamini (1993) claim that since language learning is a process which improves over time and needs an incubation period before any learning can be seen in the performance of the learners. The present study lends support to their claim which implies that a dictation practice may show its effects in the long run, and the results of the study may hold for another impact of dictation practice when longer period of time is concerned.

The finding that explained the second research question, "2) What are the differences in the use of listening strategies between the two groups? Does dictation practice make the participants use different listening strategies?", was that the experimental group appeared to use more listening strategies than the control group. Additionally, the analysis on the use of 6 different listening strategies reveals metacognitive listening strategy had a significant difference between the two groups. However, there was no significant difference in use of overall listening strategy.

Considering that successful listeners use many strategies and various types of strategies when they learn and use a second language (Bacon, 1992; Chamot & Kupper, 1989; Vandegrift, 1997), English listening classes with dictation practice have to find learners' difficulties in listening and dictating and proceed towards processing their autonomy in listening strategies.

With regards to the results, this study suggests that the research to be undertaken should consider the followings; 1) Sufficient time of dictation practice in English listening classes should be allowed. 2) Research should be conducted in the long run for a better understanding of the effectiveness of dictation. 3) Qualitative methodology should be used to investigate whether the dictation practice contributes to improve English listening proficiency.

## References

- Bacon, S. M. (1992). Phases of listening to authentic input in Spanish: A descriptive study. *Foreign Language Annals*, 25, 317-334.
- Byrne, D. (1977). *Listening comprehension practice (Teacher's book)*. Longman.
- Cartledge, H. A. (1968). A defence of dictation. *ELT Journal*, 22, 226-231.
- Chamot, A. (1987). The learning strategies of ESL students. In A. Wenden. & J. Rubin (Eds.), *Learner strategies in language learning*(pp. 71-83). Englewood Cliffs, NJ: Prentice-Hall.
- Chamot, A. U., & Kupper, L. (1989). Learning strategies in foreign language instruction. *Foreign Language Annals*, 22, 13-24.
- Davis, P., & M. Rinvulcri. (1988). *Dictation: New methods, new possibilities*. Cambridge: Cambridge University Press.
- Frodesen, J. (1991). Grammar in writing. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language*(p.268). Boston: Heinle & Heinle Publishers.
- Jafarpur, A., & Yamini, M. (1993). Does practice with dictation improve language skills? *System*, 21, 359-369.
- Lambert, W. E. (1986). Fairing first-and second-language speech and writing in ways that aid language acquisition. In Vaid, J. (ed.), *Language processing in bilinguals: Psycholinguistic and neuropsychological perspective*,(pp.65-96). Hillsdale, NJ: Lawrence Erlbaum.
- Macaro, E. (2001). *Learning strategies in foreign and second language classrooms*. London: Continuum.
- O'Malley, J., & Chamot, A. (1990). *Learning Strategies in second Language Acquisition*. Cambridge University Press.
- Oxford. R. (1990). *Language learning strategies: What every teacher should know*. New York: Newbury House.
- Pappas, G. S. (1977). You mean you still give dictations? *Language Arts*, 54, 936-939.
- Phillips, V. (1991). A look at learner strategy use and ESL proficiency. *The CA TESOL Journal*, November, 57-67.
- Rubin, J. (1975). What the good learner can teach us. *TESOL Quarterly*, 9(1), 41-51.
- Rubin, J. (1987). Learner strategies: theoretical assumptions, research history and typology. In A. Wenden, & J. Rubin. (Eds.), *Learner strategies in language learning*(pp. 15-30). Englewood Cliffs, NJ: Prentice-Hall.
- Standfield, C. W. (1985). A history of dictation in foreign language teaching and testing. *Modern Language Journal*, 69, 121-128.
- Vandegrift, L. (1997). The comprehension strategies of second language (French) listeners: A descriptive study. *Foreign Language Annals*, 30(3), 337-409.

# **An Analysis of the Interrelationship among Learner Variables of Dictation Practice in English Listening Classes**

Myeong-Hee Seong<sup>1</sup> and Eun-Hee Nam<sup>2</sup>

<sup>1</sup>Eul-Ji University, <sup>2</sup>Chung-Ang University

seong@eulji.ac.kr, hieuni@hanmail.net

## **Abstract**

The purpose of this study is to investigate the relationships among students' attitudes and perceptions of dictation practice in English listening classes, listening proficiency, and listening strategies. The research was conducted for 15 weeks with 42 Korean university students. The subjects had dictation practice on a regular basis in their listening classes for one semester. A sample TOEIC listening test was conducted at the end of the semester. Two questionnaires were used: one was to see the use of listening strategies, and the other was to find out the participants' perceptions and attitudes towards dictation practice. Some implications of the findings were discussed, along with the suggestions for further research.

## **Keywords**

Dictation practice, Listening proficiency, Listening strategies, Learners' perception and attitudes

## **Introduction**

Listening is regarded as a crucial part of foreign language learning and listening comprehension in English as a prerequisite and necessary condition for communicative competence in English. Since

listening comprehension involves very dynamic process in which one must quickly catch the sounds and grasp their meaning, it requires special attention and training for acquisition of the skill. A variety of methods to promote the listening skill have been proposed and practiced, and dictation can be a useful method.

However, Harris (1969) and Takeuchi (1997) insist that dictation can be a useful pedagogical device but offers little empirical support. Davis and Rinvulcri (1988) also mention that dictation is an effective measure of teaching, and provide us with various methods of dictation, however, empirical support for their claim of effectiveness is slight. One example of empirical data was provided by Yoshida (1984). He found statistically significant positive correlations between dictation and the results of listening tests, and based on these findings, claims that dictation can be a good teaching device.

Some researchers insist that learners' attitudes and perceptions of language learning help language development (Gardner, 1994; Gardner & MacIntyre, 1993; MacIntyre & Noels, 1996). Others insist learning is affected in a positive manner when learning strategies are used (O'Malley & Chamot, 1990; Vandergrift, 1999).

However, there has been done little research concerning the effects of dictation in relation to other learner variables. Under the circumstances, it is necessary to explore the interrelationships among attitudes and perceptions and strategies of dictation practice in English listening class and listening proficiency. Based on the above claims, the purpose of this study was to examine students' attitudes and perceptions and strategies of dictation practice in English listening classes and their relationships with listening proficiency.

## 1. Literature Review

Dictation has been a kind of practice consistently employed throughout the history of language teaching (Mohammad, 1995). According to Paul Davis and Mario Rinvolucris (1988), the benefits of dictation practice can be described like the followings: The students are active during the practice; The students are all committed to produce a text in English; The students are active after exercising while correcting their dictation; Dictation leads to oral communication activities; dictation can be used as a lead-in to other activities; Dictation leads to oral communication activities; Dictation fosters unconscious thinking; Students are consciously working on the dictation but unconsciously working on the language at a deeper level; Dictation will often calm the students since in order to do a dictation properly, the students need to be quiet and focused on the task.

However, dictation is regarded by many teachers as somewhat old-fashioned, a relic of the grammar

translation method that dominated language teaching until the last couple of decades of the 20<sup>th</sup> century. Moreover, it is not clear which type of dictation has been used throughout the history of language teaching. Nor is it clear which method has been applied to administer dictation in language classes. In this aspect, we need to research teaching methods and techniques on dictation.

The term dictation refers to any activities in which students have the chance to write down what teachers say. Davis and Rinvolucris (1998) describe many variations in dictation format and present an array of classroom activities. These range, for example, from simple dictation exercises such as vocabulary dictation and finding out silent letters to more complex ones such as question sentence dictation combined with language games.

Even in the case of simple dictation, many variations are found. If students are required to write down everything that they hear, the smallest possible unit is a word. When teachers ask students to write down part of the input only, there is space for more variation and flexibility in the demands made on the students. The combination of sentence as input and word as a slot is possible: teachers read a sentence, and one part of the sentence is left blank for the students to complete. In this kind of exercise, students can use the written co-text as a clue to guess what the missing part is. At the same time, teachers can tacitly draw students' attention to what the focus is by making slots in target place. Variables of natural speed or controlled speed, and of degree of repetition, can be added as important planning factors. Top-down listening skills also helps them become proficient at 'chunking' what

they hear beyond word by word listening.

Focus on the meaning and to make as many memo-notes as possible is an examples of a good combination of listening and comprehending skills. Some students can get the rough meaning while listening to the sentences. To sum up, for dictation, the three skills of listening, comprehending and writing need to be employed closely connected with one another.

## 2. Methodology

### 2.1 Research Questions

- (1) What is the relationship between attitudes and perceptions of dictation practice and listening proficiency?
- (2) What is the relationship between the use of listening strategies and listening proficiency?

### 2.2 Subjects

Subjects were 42 first-year students of a university with a medical specialty in Korea. Their major was radiology. Students who had stayed in an English speaking country for more than one year were excluded from the subject group. Therefore, the total number of subject was 41. The female was 33 (80.5%) and male was 8 (19.5%). The subjects were divided into two groups (Lower Ability Group and Higher Ability Group), based on the results of the listening section of the TOEIC, conducted in May, at the end of the class.

### 2.3 Instrument

Two questionnaires were used: one was to see the use of the listening strategies, and the other was to

find out the participants' perceptions and attitudes toward dictation practice. The questionnaire of perception and attitudes on dictation practice in class consisted of 31 items: the items related with 'class participation' was 6, 'class interest' 7, 'usefulness of dictation class' 7, 'motivation-provoking' 6 and 'self evaluation on dictation class' 5. The questionnaire for listening strategies was designed based on the strategies format of Oxford (1990). The test for the study was a TOEIC test.

### 2.4. Design

For the purposes of this paper, we did not include dictating activities such as language games, but rather limited the scope to dictation as an activity focusing on listening to what is said and writing what is heard. Such writing includes not only precise realization of what is heard, but also note-taking type memos.

We let the students hear a sentence and asked them to write down the complete sentence. The source of dictation was either an audio-tape or the teacher's own voice. The sentence was said once or repeated. The dictation was at a natural speed. However, sometimes teachers cut the sentences into meaningful clauses or phrases, and read these slowly so that students can catch up as they write. When we wanted to train students to focus on specific grammatical categories and forms, such as verbs and verb forms, we created materials with appropriate slots for dictation practice.

All subjects were divided into two proficiency groups according to their performances on the TOEIC Test: Higher Ability Group and Lower

Ability Group. The Higher Ability Group included 21 subjects whose score on the test was more than 55. A total of 19 subjects were included in the Lower Ability Group, in which each subject scored less than 55 on the test.

During the 13-week period between April and July, subjects took a 45-minute lesson twice a week. For evaluation of the results, a test and a questionnaire were conducted.

### 3. Results and Discussion

#### 3.1. Research Question 1

What is the relationship between attitudes and perceptions of dictation practice and listening proficiency?

##### 3.1.1 Overall frequency of attitudes and perceptions of dictation practice between higher and lower level

Table 1: Descriptive statistics of attitudes and perceptions of dictation practice for two levels

	Level	N	Mean	S.D.	t value
Participation	L	19	3.68	.60	-1.20
	H	22	3.91	.60	
Usefulness	L	19	3.41	.49	-0.43
	H	22	3.48	.60	
Interest in Class	L	19	3.61	.69	0.23
	H	22	3.55	.83	
Motivation-Provoking	L	19	3.69	.54	0.90
	H	22	3.53	.63	
Confidence	L	18	2.94	.47	0.76
	H	21	2.79	.76	

Table 1 presents the frequency of attitudes and perceptions of dictation practice for higher and lower level.

#### 3.1.2 Correlation between attitudes and perceptions of dictation practice and listening proficiency

Table 2: Correlation between attitudes and perceptions of dictation practice and listening proficiency

		Test	P	U	I	M	C
T	Pearson Correlation	1	.120	-.096	-.129	-.126	-.090
	Sig. (2-tailed)		.454	.552	.420	.431	.585
P	Pearson Correlation	.120	1	.241	.320*	.143	.207
	Sig. (2-tailed)	.454		.129	.042	.372	.206
U	Pearson Correlation	-.096	.241	1	.620**	.636**	.613**
	Sig. (2-tailed)	.552	.129		.000	.000	.000
I	Pearson Correlation	-.129	.320*	.620**	1	.718**	.696**
	Sig. (2-tailed)	.420	.042	.000		.000	.000
M	Pearson Correlation	-.126	.143	.636**	.718**	1	.633**
	Sig. (2-tailed)	.431	.372	.000	.000		.000
C	Pearson Correlation	-.090	.207	.613**	.696**	.633**	1
	Sig. (2-tailed)	.585	.206	.000	.000	.000	

\*p<.05, \*\*p<.01

P: Participation, U: Usefulness, I: Interest in Class, M: Motivation-provoking, C: Confidence

As Table 2 shows, 14 of the correlations were significant.

### 3.2 Research Question 2

What is the relationship between the use of listening strategies and listening proficiency?

#### 3.2.1 Overall frequency of listening strategies for two levels

Table 3: Descriptive statistics of listening strategies

For two levels

	Level	N	Mean	S.D.	t value	P
Memory Strategies	L	17	2.76	.39	-1.19	.241
	H	21	2.96	.57		
Cognitive Strategies	L	18	3.04	.36	-0.03	.974
	H	21	3.05	.54		
Compensation Strategies	L	19	3.20	.63	-1.63	.112
	H	22	3.52	.63		
Metacognitive Strategies	L	18	2.59	.39	-1.18	.245
	H	22	2.81	.74		
Affective Strategies	L	19	2.89	.53	-2.00	.052
	H	22	3.23	.56		
Social Strategies	L	19	2.42	.62	-0.63	.536
	H	22	2.55	.64		
Total	L	16	2.83	.37	-1.30	.204
	H	20	3.03	.52		

L: Lower Ability, H: Higher Ability

Table 3 presents the frequency of listening strategies for two levels.

#### 3.2.2 Correlation between the use of listening strategies and listening proficiency

Table 4 presents these correlations.

Table 4: Correlation between the use of listening strategies and listening proficiency

Correlations								
		Test	M	C	Co	Me	A	S
T	Pearson Correlation	1	.164	.068	.410**	.230	.364*	.212
	Sig. (2-tailed)		.324	.682	.008	.153	.019	.184
M	Pearson Correlation	.164	1	.609**	.350*	.507**	.575**	.364*
	Sig. (2-tailed)	.324		.000	.031	.001	.000	.025
C	Pearson Correlation	.068	.609**	1	.480**	.752**	.461**	.606*
	Sig. (2-tailed)	.682	.000		.002	.000	.003	.000
Co	Pearson Correlation	.410**	.350*	.480**	1	.619**	.569**	.479*
	Sig. (2-tailed)	.008	.031	.002		.000	.000	.002
Me	Pearson Correlation	.230	.507**	.752**	.619**	1	.604**	.694*
	Sig. (2-tailed)	.153	.001	.000	.000		.000	.000
A	Pearson Correlation	.364*	.575**	.461**	.569**	.604**	1	.715*
	Sig. (2-tailed)	.019	.000	.003	.000	.000		.000

S	Pearson							
	Correlation	.212	.364*	.606**	.479**	.694**	.715**	1
	Sig.(2-tailed)	.184	.025	.000	.002	.000	.000	

\* p<.05, \*\*p<.01

M: Memory strategy, C: Cognitive strategy

Co: Compensation strategies, Me: Metacognitive strategy, A: Affective strategy, S; Social strategy

As Table 4 shows, 34 of the correlations were significant.

#### 4. Conclusion

In this study, no significant correlations were found with the strategies use between the English proficiency levels and the perceptions and attitudes between the levels, but a number of significant correlations were found.

Shortcoming in this experiment should be pointed out. The period of the experiment, 13 weeks was too short. This shortness may have had some influence upon the results gained. The number of the subjects, 41 was too small and this might have affected the analysis of the data.

#### References

- Davis, P. & M. Rinvolucri. (1988). *Dictation: New methods, new possibilities*. Cambridge: Cambridge Univ. Press.
- Gardner, R. C. (1994). On motivation, research agendas, and theoretical framework. *Modern Language Journal*, 78. 359-368.
- Gardner, R. C. & MacIntyre, P. D. (1993). A

student's contribution to second language;

Part II. *Language Teaching*, 26, 1-11.

Harris, D. P. (1969). *Testing English as a second language*. New York: McGraw-Hill

MacIntyre, P. D. & Noels, K. A. (1996). Using social-psychological variables to predict the use of language learning strategies. *Foreign Language Annals*, 29, 373-386.

Mohammad K. K. (1995). *The Effect of dictation practice on general language proficiency*. Master thesis. Teheran: University for Teacher Education.

O'Malley, J. M. & Chamot, A. U. (1990). *Learning Strategies in second language acquisition*. Cambridge: Cambridge University Press.

Oxford, R.L. (1990). *Language learning strategies: What every teacher should know*. Boston: Heinle and Heinle.

Takeuchi, O. Y. (1997). Dictation: Is it really effective for language teaching? *Audio-Visual Education* 20.

Vandergrift, L. (1999). Facilitating second language listening comprehension: acquiring successful strategies. *ELT journal* 53(3), 168-176.

Yoshida, K. (1984). *Eigo no risuningu (Listening in English Language Teaching)* Tokyo; Taishukan.

# **The Development of Namseoul-Waseda Cross-Cultural Distance Learning Project from 2005 to 2009**

Bok-Myung Chang

Department of English, Namseoul University

bmchang@nsu.ac.kr

## **Abstract**

This research focuses to survey on the development of Namseoul-Waseda Cross-Cultural Distance Learning Project (henceforth NWCCDLP) from the second semester of 2005 to the spring semester of 2009. NWCCDLP has been play a very important role in cultivating communicative competence and intercultural understanding among the students of Namseoul University since it was adopted in Namseoul University from the second semester of 2005. So this research aims to survey the development of NWCCDLP focusing on various forms and functions of this project, and the reactions from the students who participated in this project.

## **Keywords**

Distance Learning, Communicative competence, English as an International Language, NNS Interaction

## **1. Introduction**

These days, we are using English as a communicative tool in every aspects of life. Jenkins(2003) found that in the period between the end of the reign of Queen Elizabeth I in 1603 and the later years of the reign of Queen Elizabeth II at the start of the twenty-first century, the number of speakers of English increased from a mere five to seven million to somewhere between one-and-a half and two billion. The English language was spoken in the mid-sixteenth century only by a relatively small group of mother-tongue speakers born and bred within the shores of the British Isles, it is now spoken in almost every country of the world, with its majority speakers being those for whom it is not a first language.

Crystal (1997) found that currently, there are approximately seventy-five territories where English is spoken either as a first language (L1), or as an official (i.e. institutionalized) second language (L2) in fields such as government, law and education. From these data, we can infer that English plays not only a role as an agent of transcending the cultural values of England or

America, but also a role as a communicative tool for every aspect of life; economics, politics, culture, finance, travel, scientific research, military, etc.

Considering these factors in technological and cultural innovation, it is necessary to set up a new model of English as an international language. Until recent times, English was used as a communicative tool among native English speakers, but recently English was used as a tool among native English speakers(henceforth; NS) and nonnative English speakers(henceforth; NNS). Even more the communication between NNSs and NNSs is delivered through English, so it is urgent to set up a new model of English as an "International Language" or a "World English."

This work focused on facing up to the changing roles and status of English, and establishing the new notion of World English. The research process of this paper is based on the data collected during Namseoul University – Waseda University Cross-Cultural Distance Learning Project(henceforth; NWCCDLP) started from the second semester, 2005.

CCDL program aims to achieving the following purposes:

The first is to raise intercultural awareness through developing deeper understanding toward other cultures and appreciating diversity and multiple viewpoints. The second is to use English as a communication tool. The third is to be able to explain their own cultures in English.

## **2. The 21st century and cross-cultural communication**

In our view, the 21st century will be characterized by two concepts: highly advanced information technology and globalization. The advent of highly advanced information technology has created cyberspace and reduced the actual distance between states. Although we are in the incipient stage of the so-called borderless society, in the new century, it is not so hard to infer, such a borderless society will make a rapid development with further progress in highly advanced information technology. As a result,

people will be able to have more opportunities to meet and communicate with those who differ in, say, culture, value systems, religion, manners and customs, and so forth. The advent of globalization, on the other hand, has called on people and enterprises as well as states to get into step and cooperate with their overseas partners in order to attain common interests. Such movements toward globalization will intensify in the next century. Under such circumstances, it naturally follows that human interchange based on mutual understanding will play an important role in the 21st century.

### 3. Research Methodology

#### 3.1 Subjects

The students from WU and NSU participated in Namseoul University – Waseda University Cross-Cultural Distance Learning Project (henceforth NWCCDLP) started from the second semester, 2005.

#### 3.2 Data Collection and Analysis

The data has been collected from the fall semester of 2005 to the spring semester of 2009.

Data is analyzed by a qualitative method through analyzing the students' reflection report and a quantitative method through analyzing the results of the questionnaire implemented at the end of every semester.

### 4. The Development of NWCCDLP

From the start of second semester, 2005, Namseoul University participated in Namseoul – Waseda Universities Cross-Cultural Distance Learning Project; NWCCDLP). This project consists of two parts, the one is the video chatting activities, and the other is the video conferencing lecture(See 3. NWCCDLP). For the video chatting activities, Namseoul University students made partners with the students from Waseda University and agreed to meet in cyber space through a “Live On” program. During these activities, most students from both universities met for one hour once a week, and they had to record all the interactions through text chat and submit them as a report for the course.

The video conferencing lecture had the theme of “World Englishes and Miscommunications,” in which Namseoul Univ.(Korea), Waseda Univ. (Japan), Hudan Univ.(China), and Hannam Univ. (Korea) participated every Thursday from 14:40 to 16:10. Every student must study the lecture in advance through an on-demand lecture series on the web site. During the lecture, every student must

participate in the course through listening and speaking only in English.

#### 4.1 Contents of NWCCDLP from 2005 to 2009

2005	On-line Chatting, Video-Conferencing Lecture
2006	On-line Chatting, Video-Conferencing Lecture, New Model Lesson
2007	On-line Chatting, Video-Conferencing Lecture, New Model Lesson
2008	On-line Chatting, Video-Conferencing Lecture, New Model Lesson
2009	On-line Chatting, Video-Conferencing Lecture, New Model Lesson

#### 4.2 Technologies used in NWCCDLP from 2005 to 2009

The latest multimedia and internet technologies such as video conferencing, text and voice chat and e-mail systems are fully utilized for this project.

#### 4.3 Course Description

##### 4.3.1 Video chatting activities

The students from Namseoul and Waseda Universities were matched as partners and had to meet in cyberspace at the appointed time and date through the “Live On” system.

##### 4.3.2 Video conferencing lecture

English has been used as a common tool of communication in Asia. It has been reported that non-native users of English in the so-called outer circle and expanding circles outnumber the native speakers of English in the inner circle. English has thus been implanted in various regions of the world and appended some peculiarities in the respective regions. For this reason, this lecture's purpose is to make the students learn the mutual intelligibility concerning phonetic features, syntactic features, socio-cultural differences, and para-linguistic features which might cause some misunderstanding among native speakers and Asian interlocutors.

As a second part of NWCCDLP, Namseoul Univ. participated in this lecture along with Waseda Univ., Hudan Univ., and Hannam Univ. every Thursday. Most students from Namseoul participated actively in this lecture by expressing themselves and their opinions about World Englishes only in English. From the data analysis of the questionnaire collected from the students who participated in

NWCCDLP, it can be concluded that NWCCDLP was very helpful in activating the students' interests about world English and cultural understanding. So that this project resulted in the improvement of English communicative proficiency for Namseoul Students.

#### Syllabus for World English in NWCCDLP

Month	Date	Contents
9	1-29	Introduction for World English
10	6	Introduction for the participants
	13	Singaporean English
	20	Malay English / Indian English
	27	Philippine English
11	3	Break
	10	Hong Kong English
	17	Chinese English
	24	Korean English
12	1	Taiwan English

#### 4.3.3 New CCDL Model from 2007 to 2009

##### 1) Course Procedure of New CCDL Model

Time	Type of Group	Tool for Communication
Every Thursday	Project	
10:40-11:00	Group Presentation	Video Conferencing
11:00-11:20	Group Presentation	Video Conferencing
11:20-12:10	Small Group Chatting	Live-On System and MOODLE

##### 2) New CCDL model of the 2<sup>nd</sup> semester, 2008

##### Class Model

- ☐ Preparation Class – 1 week  
(Reading + Research +BBS)
- ☐ Joint Class – 1week  
(Voice Chatting + Text Chatting + BBS)
- ☐ Reflection Paper

##### Social and Global Issues

##### Topic 1 : Introduction

Facilitation Skills, Intercultural translation

##### Topic 2 : Happiness Factors

Individualism/Collectivism

##### Topic 3 : Family Roles

High/Low context communication styles

##### Topic 4 : Neighborhood

Relationship Development

##### Topic 5 : Climate Change : Agreement Styles

☐ Number of the students who participated in NWCCDLP from 2005 to 2009

Year	Spring Semester	Fall Semester
2005	-----	148
2006	175	98
2007	178	87
2008	142	78
2009	150	-----

## 5. Results and Conclusions

- 1) This class helped the students to have self-confidence in English.
- 2) This class helped the students to understand Japanese culture and people.
- 3) This class was very interesting.
- 4) This class helped the students to recognize the importance of English.
- 5) This class helped the students to improve their English abilities.

In conclusions, NWCCDLP can be proved as a good model for cultivating English proficiency of the students in EFL contexts. From the results of this research, the conclusions can be suggested as follows. The first thing is that most students could develop their English proficiency through the NNS-NNS interaction. So it is very helpful in English acquisition to communicate in English among non-native speakers. The second thing is that the international distance learning model can be suggested as English learning model in EFL context. The third thing is that one of the distance learning models, NWCCDLP was very impressed with the possibility of being able to get in touch with students and cultures in foreign countries. Most of the students who practiced English through NWCCDLP wished to participate in this CCDL model again if possible, and they think this project can help themselves to activate the interests in cultural understanding, get their self-confidence in using English, and improve their English skills.

## References

- Baek, N. S. (2002). A study in the teaching method of English literature through the Internet and its effect. *English Language & Literature Teaching*, 8(1), 1-13.
- Chang, B.M. (2004). A critical analysis of communicative strategies among non-native English speakers. *Modern English Education*, 5(2), 173-195.
- Chang, B.M. (2005). World English Acquisition through CCDL Project. *Journal of Pan-Pacific Association of Applied Linguistics*, 9(2),

241-261.

Chang, B.M. (2006). Exploring a New Model for Cross-Cultural Distance Learning. *Journal of Pan-Pacific Association of Applied Linguistics*, 10(2), 1-21.

Chang, B.M. (2007). Creating an Innovative Model for Distance Learning in EFL context. *Modern English Education*, 8(1), 4-32.

Chang, B.M. (2008). English Acquisition through NNS-NNS Interaction in EFL Context. *English Language & Literature Teaching*, 13(2), 1-23.

Crystal, D. (1997). *English as a Global Language*. Cambridge: Cambridge University Press.

Jenkins, J. (2003). *World Englishes*. New York: Routledge.

Kachru, B. (1992). *The Other Tongue: English Across Cultures* (2nd ed.). Urbana, IL: University of Illinois Press.

McArthur, A. (1987). The English languages?. *English Today*, 11, 9-13.

Modiano, M. (1999). Standard Englishes and educational practices for the world's lingua franca. *English Today*, 15 (4), 3-13.

Modiano, M. (2000). Rethinking ELT, *English Today*, 16(2), 30-34.

Park, K.J.(2001). The importance of CCDL in English education. *English Language & Literature Teaching*, 7(2), 77-101.

# A Method of Automatic Acquisition of Typed-Dependency Representation of Japanese Syntactic Structure

Masanori Oya

School of Education, Waseda University  
Masanorioya2008@gmail.com

## Abstract

This paper introduces a method of automatic acquisition of typed-dependency representation of Japanese syntactic structure, with its problems and possibilities.

## Keywords

Natural language processing, typed dependency, directed acyclic graphs.

## 1. Introduction

This paper introduces a typed-dependency representation (de Marneffe and Manning, 2008a) of Japanese syntactic structure, as part of automatic acquisition of grammatical information from a Japanese text corpus (Kyoto University Text Corpus ver.4, henceforth KTC4) (Kurohashi et al. 1997, Kurohashi et al. 1998).

## 2. Stanford Typed Dependency Parser

According to de Marneffe and Manning (2008a), typed-dependency representation has inherited a lot from Lexical-Functional Grammar (LFG) (Bresnan, 2001), and “aims to provide a simple, habitable design” of syntactic structure. In a typed-dependency representation, all syntactic information is represented as binary relations of two words which are connected by a typed dependency (such as *nsubj* “nominal subject” or *num* “numerical classifier”). This representation maps easily onto other common representations, such as Resource Description Framework (cf. Section 4), and researchers can implement it into various fields of NLP. De Marneffe and Manning (2008a) proposed the Stanford typed-dependency parser, which outputs both the context-free phrase structure grammar representation of the input sentence, and a typed-dependency acyclic directed graph, with each of the words as a node and each of the dependencies between two words as a labeled arc.

For example, Stanford typed dependency parser outputs the following dependency triples for a sentence “I have read the book”:

- (1)
  - nsubj*(read-3, I-1)
  - aux*(read-3, have-2)
  - det*(book-5, the-4)

*dobj*(read-3, book-5)

Each dependency triple represents the graph specifications between two words which are in dependency relationship. For example, the first triple states that the word “read” is the third in the sentence and “I” the first, and they are connected by a directed arc labeled with “*nsubj*”, meaning that the word “read” has “I” as its nominal subject.

These triples are equivalent to the following directed acyclic graph:

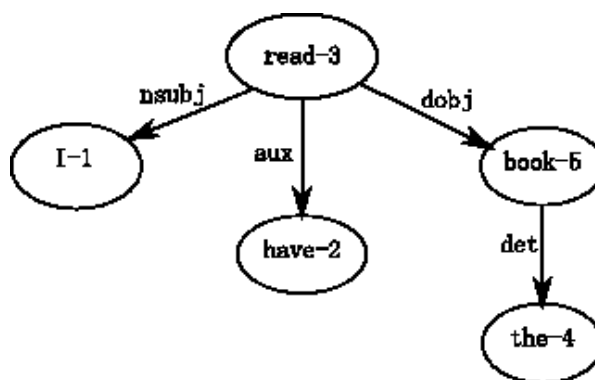


Figure 1: The directed acyclic graph for the sentence “I have read the book”.

When lexical information on each of the node is unified along with the directed arcs backward, then we have the f-structure representation for the sentence. For more details of Stanford typed-dependency parser, see de Marneffe and Manning (2008b).

## 3. Automatic transformation of Japanese text corpus representation and dependency parser into typed-dependency representation

I implemented the idea of de Marneffe and Manning (2008a, 2008b) into the parser output of KNP, a dependency parser for Japanese (Kurohashi et al. 2005), so that we can acquire typed-dependency representations of Japanese sentences, which can be further implemented in the field of NLP.

### 3.1 KTC4 and KNP

In KTC4, each sentence is divided into a number of syntactic units, and the dependencies among them are annotated on each of the units, along with other

morphological information. This corpus format is basically the same with the parser output of KNP, since both of them were developed along with each other (Kurohashi et al. 1998). Therefore, a text-processing program for KTC4 can also be applied to the output of KNP only with minor modifications in the program.

For example, the KNP output for a Japanese sentence “Watashi-wa sono hon-wo yonda” (I have read the book.) is as follows (Japanese words are Romanized and annotated with English words, and tags which are irrelevant for the present discussion are deleted):

(2)  
 # S-ID:1 KNP:2009/06/28  
 \* 3D  
 Watashi “I” noun normal  
 wa particle adverbial  
 \* 2D  
 sono “the” det adjectival  
 \* 3D  
 hon “book” noun normal  
 wo particle case  
 \* -1D  
 Yonda “read-past” verb  
 。 。 。 period  
 EOS

The asterisks are unit boundaries. The numbers next to asterisk specify the syntactic unit on which the current syntactic unit depends. For example, the first syntactic unit “watashi-wa” depends on the 3<sup>rd</sup> unit “yonda” (notice that the counting starts from zero; the first syntactic unit is the 0<sup>th</sup> unit, the last the 3<sup>rd</sup> unit). The last unit “yonda” is the root unit which depends on nothing, hence it is annotated with “-1”. The “D” next to the number specifies that the dependency is direct; there are two other types of dependency in KTC4/KNP representation: “A” means appositional dependency, and “P” means parallel (or coordination) dependency.

However, the dependency representation in KTC4 and KNP output is not typed, i.e., no grammatical function is annotated on the dependency. In this sense, we need to augment the corpus representation and KNP parser output in terms of which type of dependency must be annotated on a given dependency among syntactic units.

### 3.2 Automatic acquisition of typed dependencies among Japanese syntactic units from parser output

In Japanese, the type (or absence) of particles within a syntactic unit determines the type of

dependency between units. Therefore, the morphological information in KTC4 corpus annotations and a KNP parser output can be employed for automatic transformation of KTC4 corpus annotations and KNP parser output into LFG-style attribute-value pairs, and also into an acyclic directed graph which is similar to the output of the Stanford typed-dependency parser.

The grammatical function of a given syntactic unit is determined according to the morphological information which is provided as KTC4/KNP tags. Syntactic units are divided into the following categories (Oya, forthcoming):

- (3)
1. Particled inflective units: units which have at least one particle, and which have an inflecting element (verb, adjective, verbal suffix or copula) as their head.
  2. Particled non-inflective units: units which have at least one particle, and which do not have an inflecting element as their head.
  3. Non-particled, inflective units: units which have no particle, and which have an inflecting element as their head.
  4. Non-particled, non-inflective units: units which have no particle, and which do not have an inflecting element as their head.

Every syntactic unit belongs to one of these categories. The units in one category are further divided into subcategories according to the type of the particle they have, or according to the inflection form of the head. One grammatical function is assigned to each of the subcategories. The process is summarized in (4). The process proceeds automatically and is implemented with much use of regular expressions:

(4)

For each syntactic unit,

#### Step 1. Check the root.

1. If the current syntactic unit is the root of the sentence, go to Step 9.
2. Else, go to Step 2.

#### Step 2. Check the presence of particles.

1. If the current syntactic unit has a particle, go to Step 3.
2. Else, go to Step 4.

#### Step 3. (for particled units) Check the presence of inflecting elements.

1. If this syntactic unit has at least one inflecting element, go to Step 5.
2. Else, go to Step 6.

#### Step 4. (for non-particled units) Check the presence of inflecting elements.

1. If this syntactic unit has at least one inflecting

element, go to Step 7.

2. Else, go to Step 8.

**Step 5. (for particled, inflective units) Check the particle.**

1. If this unit has the case particle “-to”, then the grammatical function of this unit is a complement; annotate it with COMP.
2. Else, if this unit has a formal noun and either of the case particles “-ga”, “-wo”, “-ni”, then this unit has the grammatical function subject, object or oblique, respectively; annotate it with “subj”, “obj” or “obl”, respectively.
3. Else, if this unit has a formal noun and a case particle other than “-ga”, “-wo” and “-ni”, then this unit has the grammatical function particled adjunct; annotate it with “padj”. (Complement clause in this unit will be treated later).
4. Else, if this unit has a particle other than case particle, then this unit has the grammatical function sentential adjunct; annotate it with “sadj”.
5. Go to Step 9.

**Step 6. (for particled, non-inflective units) Check the particle.**

1. If the dependency type of this unit is coordination, then the grammatical function is “coord”.
2. Else, if the dependency type of this unit is apposition, then the grammatical function is APP. Else, if the particle is “-ga”, then the grammatical function is “subj”.
3. Else, if the particle is “-wo”, then the grammatical function is “obj”.
4. Else, if the particle is “-ni”, then the grammatical function is “obl”.
5. Else, if the particle is a case particle other than “-ga”, “-wo” and “-ni”, then the grammatical function is “padj”.
6. Else, if the particle is “-wa”, then the grammatical function is “topic”.
7. Else, if the particle is an adverbial particle other than “-wa”, then the grammatical function is “focus”.
8. Else, the grammatical function is “padj”.
9. Go to Step 9.

**Step 7. (Non-Particled, Inflective)**

1. If the inflection form of the last morpheme of this syntactic unit is the declarative base form, or the declarative ta form, then the grammatical function is “rel”.
2. Else, the grammatical function is “sadj”.
3. Go to Step 9.

**Step 8. (Non-Particled, Non-Inflective)**

1. If the dependency type of this unit is coordination, then the grammatical function is “coord”.

2. Else, if the dependency type of this unit is apposition, then the grammatical function is “app”.
3. Else, if the part of speech of the head of this unit is determiner, then the grammatical function is “det”. Else, the grammatical function is “adj”.
4. Go to Step 9.

**Step 9. (Root or not)**

1. If the current unit is the root of the sentence, then stop.
2. Else, move to the next unit and go to Step 1.

This process does not yield syntactic units whose grammatical function is ambiguous; this is based on an assumption that every syntactic unit has at most one grammatical function.

Through this process, we can automatically acquire the dependency triples from the KNP output for the sentence “Watashi-wa sono hon-wo yonda” as follows:

- (5)
- topic(yonda-4, watashi-1)
  - det(hon-3, sono-2)
  - obj(yonda-4, hon-3)

These triples are equivalent to the following directed acyclic graph:

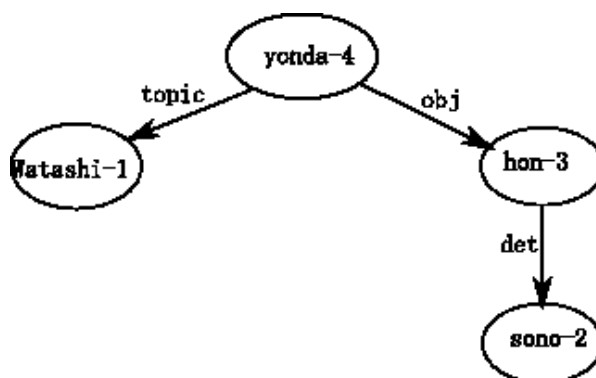


Figure 2: The directed acyclic graph for the sentence “Watashi-wa sono hon-wo yonda” (I have read the book).

As in the case of Stanford Parser, typed dependency representation can be transformed easily into other common representations, including RDF (cf. Section 4), and researchers can implement it into various fields of NLP.

### 3.3 Zero-pronoun identification

There are a number of problems in this method of automatic acquisition of typed dependencies among syntactic units in Japanese sentences. First of all, this method does not identify zero pronouns in a sentence. For example, the subject of the example

sentence is a zero pronoun which refers to the topic “watashi-wa”. This fact must be reflected in the dependency triples, yet the method describe above fails to do so. We need to have an additional method to address this issue, and Oya (forthcoming) proposed a simplistic method of adding appropriate subject zero pronouns in the parser output if a clause does not contain ga-marked noun phrase; this method is based on Subject Condition (Dalrymple 2001), which states that every verbal predicate must have a subject.

As for non-subject zero pronouns, the zero-pronoun identification poses a more difficult problem. Suppose a verbal predicate in an input sentence does not contain an object, or a wo-marked noun phrase. It is only after we can identify this verbal predicate as a transitive verb that we can put an object zero pronoun for this verbal predicate. If we put an object zero pronoun for an intransitive verbal predicate, then this analysis is just incorrect. Oya (forthcoming) has implemented a method for object zero-pronoun identification using morphological information of verbal predicates and the probability that a given verbal predicate has its object, yet the result of this method is the f-scores of 54 % for identifying object zero pronouns. Implementing more sophisticated methods for zero-pronoun is a question to be addressed in further research.

#### 4. Application of typed-dependency representations

Since directed acyclic graph representations of typed-dependency of sentences are simple and mathematically well-defined, it is easy to transform it in other common formats for different purposes. One of the possibilities is to transform it into Resource Description Framework (RDF).

RDF is a unified framework for describing resources on the Web. The purpose of RDF is to describe metadata (data for data), as one of the components of the Semantic Web (Berners-Lee et al. 2001).

The advantage of RDF is that it is standardized by World Wide Web Consortium (W3C), a non-profit organization for standardization of web technology; they provide the abstract syntax of RDF <http://www.w3.org/TR/rdf-concepts/>, and also provide the syntax of RDF/XML for actual applications using RDF <http://www.w3.org/TR/rdf-syntax-grammar/>. The validity of an RDF document can be validated by a Web page provided by W3C (<http://www.w3.org/RDF/Validator/>).

By formatting typed-dependency relations of sentences in a given text into RDF/XML, we can provide linguistic information to researchers who are working in various fields of text processing, but

less linguistics-oriented; as far as they know how to process RDF documents and the minimum knowledge of linguistics, they can utilize the RDF/XML typed-dependency corpora for their study.

#### 5. Conclusion

In this paper, I present a procedure of automatically determining the typed dependency among syntactic units of Japanese, with the remaining problem of zero-pronoun identification. I also discussed the possibilities of applying typed-dependency representations of English and Japanese sentences as a format of data for various purposes.

#### References

- Bresnan, J. (2001). *Lexical-Functional Syntax*. Blackwell, Oxford.
- Dalrymple, M. (2001). *Lexical Functional Grammar*. Academic Press.
- De Marneffe, M.C. & Manning, C. (2008a). The Stanford typed dependencies representation. In *COLING Workshop on Cross-framework and Cross-domain Parser Evaluation*. Retrieved June 27, 2009, from <http://nlp.stanford.edu/pubs/dependencies-coling08.pdf>
- De Marneffe, M.C. & Manning, C. (2008b). *Stanford Typed Dependencies Manual*. Retrieved June 27, 2009, from [http://nlp.stanford.edu/software/dependencies\\_manual.pdf](http://nlp.stanford.edu/software/dependencies_manual.pdf).
- Kurohashi, S. and Nagao, M. (1997). Kyoto University text corpus project. *Proceedings of the 3rd Annual Meeting of the Association for Natural Language Processing*, 115-118.
- Kurohashi, S. and Nagao, M. (1998). Building a Japanese parsed corpus while improving the parsing system. *Proceedings of the 1st International Conference on Language Resources and Evaluation*, 719-724.
- Berners-Lee, T., Hendler, J. and Lassila, O. (2001). The Semantic Web. *Scientific American Magazine*. Retrieved June 27, 2009, from <http://www.scientificamerican.com/article.cfm?id=the-semantic-web>
- Oya, M. (forthcoming). *Treebank-Based Automatic Acquisition of Wide Coverage, Deep Linguistic Resources for Japanese*. M.Sc. thesis, Dublin City University.

# How intelligible and acceptable is ‘Japanese English’ pronunciation?

Nozomi Kato

College of Applied International Studies, Tokiwa University

Nozomi.Kato@gmail.com

## Abstract

This study examines how Japanese accent in English pronunciation (Japanese English) sounds to American English native speakers (n=14) and Japanese native speakers (n=14) to make certain contributions to the investigation of the features of Japanese English functioning as an international language (EIL). Participants’ evaluation scores of both intelligibility and acceptability were compared in terms of mean ranks and the Mann Whitney Test. This study shows that CVC<sup>1</sup> could be one of Japanese English features as EIL, whereas a substitutive vowel /i/ for /I/, which is often regarded as one of the features of Japanese English, does not seem to work as EIL.

## Keywords

EIL, World Englishes, Asian Englishes, Japanese English, Pronunciation

## 1 Introduction

Now that English plays a significant role as an international language, according to Jenkins (2000), it is not appropriate to hold up as the standard only a few limited pronunciations—usually Received Pronunciation (RP) and General American (GA). She and many other scholars such as Kachru (1996) label these not as errors but varieties of English (the World Englishes) with a lot of various accents deviating from either RP or GA.

Although issues surrounding World Englishes have been often discussed, Japanese English has not been well established yet; however, more and more research on the World Englishes including Japanese English has been carried out in recent years (Nakano et al., 2007). This recent research demonstrates some of the features of Japanese English: the substitutions of /I/ (Nakano et al., 2007; Cross, 2002; Jenkins, 2000), /θ/ (Jolly, 2000; Jenkins, 2000), and CC<sup>2</sup> (Jenkins, 2000). The

current study investigates how intelligible and acceptable these raised features of Japanese English are. The research questions in this study are as follows:

- Q1. How intelligible and acceptable are the selected features of Japanese English (/i/ for /I/, /s/ for /θ/, and CVC for CC) for American English and Japanese speakers and participants with high and low proficiency of their respective foreign languages?
- Q2. Does the frequency of the word used in the corpus affect intelligibility and acceptability of Japanese English?

## 2 Method

### 2.1 Participants

The participants consisted of two main groups: (1) auditory material contributors and (2) the raters of the auditory text. The participants in Group (1) were a Japanese graduate student majoring in English language Teaching, an American undergraduate student from Utah, USA, and an American undergraduate student from California, USA. The participants in group (2) were (a) seven Japanese speakers with high proficiency English listening skill, (b) seven Japanese speakers with low proficiency English listening skill, (c) seven American English speakers with high proficiency Japanese listening skill, and (d) seven American English speakers with low proficiency Japanese listening skill.

The participants in (a), (b), (c), and (d) were selected according to their self-assessment in the following criteria (Table.1).

Table 1: Self-assessment for participants group 2

Participants	Criteria
American	The ratio of their understanding of Japanese films
Japanese	The ratio of their understanding of American films

<sup>1</sup> consonant + vowel + consonant (Japanese ‘imperfect consonant cluster’ having an additional vowel inserted between consonants)

<sup>2</sup> consonant + consonant (consonant cluster)

The participants (a) and (c) were those who marked 60% or more in their self-assessment. The

participants (b) and (d) were those who marked 40% or less in their self-assessment.

## 2.2 Instrument

Auditory material included the words and phrases with the following features: /i/, /I/, /s/, /θ/, and consonant clusters. To avoid any confusion caused by the difference between GA and RP, all the sounds of native speakers of English were collected only from General American speakers.

Those features were selected to investigate how Japanese accented English would be perceived as they are some of the features shown in many previous studies that Japanese learners of English often cannot pronounce exactly as the phonetic transcriptions indicate. They were also chosen because Jenkins (2000: 158-159) points out that some of them are not necessarily pronounced exactly to be intelligible and the other features need to be pronounced as the phonetic transcriptions indicate. According to her research, most substitutions of /θ/ are permissible; vowel sounds are intelligible as long as maintenance of their length contrasts is kept. However, consonant clusters need to be pronounced as the transcriptions indicate in order to be intelligible.

This auditory material was also chosen to investigate any relationship with frequency of the words and phrases. This frequency data was taken from concordance software called *VIEW*, which was developed by Mark Davis at Brigham Young University. *VIEW* is concordance software that allows you to investigate the frequency of the occurrence of given words or phrase. The system involves the major corpus data such as British National Corpus and Corpus of Contemporary American English.

It can be assumed that the words and phrases that have relatively high frequency rates may be more intelligible than the words and phrases that have relatively low frequency rate not because of deviation from phonetic transcription in their pronunciations, but because of familiarity with the words. Three types of minimal pairs were chosen to investigate such effect:

- (1) The type in which the words and phrases with Japanese accent occur less than those with American accent. For instance, as seen in Table 2, the word, 'think' occurs 88755 times; on the other hand, 'think' with Japanese accent, which sounds like 'sink,' occurs 1807 times.
- (2) The type in which the words or phrases with Japanese accent occur as often as those with American accent. For example, as shown in Table 2, the word, 'thick' occurs 4500 times

and its Japanese accented word sounding like 'sick' occurs 4333 times. They can be considered as a similar frequency pair.

- (3) The type in which the words or phrases with Japanese accent occur more often than those with American accent. For example, Japanese accented word for 'thumb' sounding like 'some' occurs as many as 167317 times, whereas the original sound of 'thumb' occurs only 1096 times.

With such criteria in the selection above, the following words and phrases as auditory material in this research were listed (Table 2).

Table 2: Auditory Material

GA	JE	JN	Words and Phrases
/θ/	/s/, /t/	○	think(88755)/sink(1807) thick(4500)/sick(4333) thumb(1096)/some(167317) thank you(9598)/sank you(0)
/I/	/i/	△	its(160568)/eats(390) live(16834)/leave(19820), fill(3941)/feel(25367) give
CC	CVC	×	smoke play

Notes:

JE: Japanese English

JN: Jenkins' notes toward JE taking from her discription of Lingua Franca Core (Jenkins, 2000: 158-159)

○: Permissible

△: Permissible as long as maintenance of vowel length contrast is kept

×: Not desirable to be intelligent

The number in the brackets: Occurrence rate from the Corpus Data (*VIEW*, Mark Davis at Brigham Young University)

## 2.3 Procedure

The data was collected in the following ways:

- (1) One Japanese speaker's pronunciations (Japanese English) and two American English speakers' pronunciations for all the items were recorded.
- (2) The participants in group 2 listened randomly to each of three different sounds that were recorded by the participants in group 1 so that the randomness would prevent the participants in group 2 from easily recognizing whether the sound of each item was American accented or Japanese accented. Each item was pronounced twice; one of the sounds was American

accented and the other was Japanese accented.

- (3) The participants were asked to evaluate each pronunciation item with a scale of one to six:

Q (a) How clearly can you recognize the listed word or phrase?

Q (b) How much would you like to accept this pronunciation in international conversation?

Notes:

As for (a), one is for the least clear and six for the clearest.

As for (b), one is for the least acceptable and six for the most acceptable.

Question (a) was used to investigate the intelligibility of Japanese English sound in each item. Question (b) was used to investigate the acceptability of Japanese English sound in each item. These are the definitions of intelligibility and acceptability in this study.

## 2.4 Analysis

Participants' evaluation scores of both intelligibility and acceptability were compared in terms of mean ranks and a non-parametric test, called Mann Whitney Test. This test was used to examine group difference of intelligibility and acceptability in perception of the accented phonemes (/i/ for /I/ and /s/ for /θ/) and the imperfect consonant cluster (CVC) between American and Japanese participants and the participants having high and low proficiency of their respective foreign languages. These two group contrasts (American vs. Japanese, high vs. low proficiency) were also examined in terms of corpus frequency of the item words and phrases.

## 3 Results

### 3.1 Analysis on intelligibility

This analysis answers a half of the research question 1 in this study.

- Q1. How intelligible and acceptable are the selected features of Japanese English (/i/ for /I/, /s/ for /θ/, and CVC for CC) for American English and Japanese speakers and participants with high and low proficiency of their respective foreign languages?

As seen in Table 3, American judges were stricter than Japanese judges as a whole. The Japanese English features: /i/ for /I/ and CVC for CC had an especially large deference between American and Japanese raters. This means that American evaluated those items much more strictly.

As for the judgments between the participants with low and high proficiency in their listening skill

for their foreign languages: Japanese and American English, those with high proficiency were relatively stricter, but the difference was smaller compared with the difference between American and Japanese judges according to Table 3.

Table 3: Mean Ranks on Intelligibility

	P	N	J/θ/	J/I/	JCC
	A	14	14.00	17.14	16.71
Mean	J	14	12.86	9.64	10.93
Rank	L	14	11.46	14.57	10.86
	H	14	15.13	16.43	16.79

Notes:

P: Participants

A: American

J: Japanese

L: The participants having low proficiency in Japanese or English listening skill as their foreign languages

H: The participants having high proficiency in Japanese or English listening skill as their foreign languages

J/θ/: An assumed Japanese English feature, /s/ functioning as a substitution of /θ/

J/I/: An assumed Japanese English feature, /i/ functioning as a substitution of /I/

JCC: An assumed Japanese English feature, an imperfect consonant cluster with a vowel inserted between two consonants

As seen in Tables 4 and 5, the results of the Mann-Whitney Test on intelligibility show that the difference between American and Japanese judgments regarding Japanese substitution of /i/ for /I/ is significant ( $p < .05$ ). This test also shows that the participants with high proficiency in the target languages (Japanese and English as foreign languages) were significantly stricter in their judgments regarding /i/ for /I/ than those with low proficiency ( $p < .05$ ).

An assumed feature of Japanese English, CVC, a substitution of CC was found to be not significant but close to significance in both of sets: American and Japanese participants, and the participants with low and high proficiency.

Table 4: Mann-Whitney Test on Intelligibility -A J

	J/θ/	J/I/	JCC
Mann-Whitney U	90.500	44.500	57.500
Z	-.347	-2.478	-1.868
Asymp. Sig. (2-tailed)	.729	.013	.062

None:

AJ: A comparison between the scores marked by the participants of native speakers of American English and Japanese

Table 5: Mann-Whitney Test on Intelligibility -HL

	J/θ/	J/I/	JCC
Mann-Whitney U	68.000	55.000	58.000
Z	-1.388	-1.991	-1.845
Asymp. Sig. (2-tailed)	.165	.046	.065

None:

HL: A comparison between the scores marked by the participants having high and low proficiency of Japanese or English listening skill as their foreign languages

### 3.2 Analysis on acceptability

This analysis answers the other half of the research question 1 shown below:

Q1. How intelligible and acceptable are the selected features of Japanese English (/i/ for /I/, /s/ for /θ/, and CVC for CC) for American English and Japanese speakers and participants with high and low proficiency of their respective foreign languages?

As seen in Table 6, the results of mean ranks were diverse. This results were different from the results of intelligibility, in which American group is always stricter than Japanese group. The same tendency is shown in the comparison between low and high proficiency groups.

Table 6: Mean Ranks on Acceptability

	P	N	J/θ/	J/I/	JCC
Mean	A	14	13.71	15.79	10.18
Rank	J	14	15.29	13.21	18.82
	L	14	13.75	15.11	16.32
	H	14	15.25	13.89	12.68

As seen Tables 7 and 8, we cannot see any significant difference in any combinations of groups in any items except the comparison between American and Japanese participants judging Japanese CVC for English CC. The difference is significant ( $P < .05$ ).

Table 7: Mann-Whitney Test on Intelligibility -A J

	J/θ/	J/I/	JCC
Mann-Whitney U	87.000	80.000	37.500
Z	-.508	-.833	-2.798
Asymp. Sig. (2-tailed)	.611	.405	.005

Table 8: Mann-Whitney Test on Intelligibility -HL

	J/θ/	J/I/	JCC
Mann-Whitney U	87.500	89.500	72.500
Z	-.485	-.394	-1.179
Asymp. Sig. (2-tailed)	.628	.694	.238

### 3.3 Analysis based on frequency of item use

This analysis answers the research question 2.

Q2. Does the frequency of the word used in the corpus affect intelligibility and acceptability of Japanese English?

Mean ranks and the Mann-Whitney test did not show any significant difference between American and Japanese groups or between high and low proficiency groups based on frequency of use of the words and phrases which contains the feature of /s/ for /θ/ on both intelligibility and acceptability measures. However, difference was observed between the items having the feature of /i/ for /I/. Therefore, the following analysis focuses on this feature.

#### 3.3.1 Intelligibility of /i/ for /I/

The hypothesis of this experiment was that participants might understand frequently used items better than relatively less frequently used items. However, Table 9 did not indicate such tendency. According to Table 2(p.3), the listed item, 'its' is much more often used than 'fill', but the difference in the mean ranks for 'its' and 'fill' was less than 3 in any comparison (see Table 9).

Table 9: Mean Ranks based on frequency of use the words and phrases having the feature of J/I/ regarding intelligibility

	P	N	give	its	live	fill
Mean	A	14	16.39	18.00	17.04	16.18
Rank	J	14	12.61	11.00	11.96	12.82
	L	14	16.54	16.46	18.00	14.18
	H	14	12.46	12.54	11.00	14.82

Notes:

The frequency of item use in comparison with minimal pair is shown below based on *VIEW* (see 2.2).

give/no minimal pair

its(160568)/eats(390)

live(16834)/leave(19820)

fill(3941)/feel(25367)

As seen in Table 10, the Mann-Whitney test demonstrates that the judgment of 'its' between American and Japanese participants was significantly different. American listeners were less able to understand this word with Japanese accent compared with Japanese listeners. The same tendency is found regarding 'live.' It is not significant, but the probability is close to significant.

On the other hand, Table 11 shows no significant difference in any items in the comparison between low and high proficiency groups.

Table 10: Mann-Whitney Test on Intelligibility -A J

		give	its	live	fill
Mann-Whitney		71.500	49.000	62.500	74.500
U					
Z		-1.245	-2.327	-1.668	-1.097
Asymp. Sig. (2-tailed)		.213	.020	.095	.272

Table 11: Mann-Whitney Test on Intelligibility -HL

		give	its	live	fill
Mann-Whitney		73.000	96.000	72.500	81.000
U					
Z		-1.175	-.095	-1.229	-.797
Asymp. Sig. (2-tailed)		.240	.925	.219	.425

### 3.3.2 Acceptability of /i/ for /I/

Mean ranks again did not indicate the effect of occurrence rate in perception of the items by American and Japanese groups and high and low proficiency groups. The largest difference in mean rank score between items and participants was only 2.08. It is difficult to say that frequency takes any role for the listeners' acceptability from this.

As seen in Table 12, no significant difference between American and Japanese groups can be found in the Mann-Whitney test. However, according to table 13, the Mann-Whitney test shows that there was a significant difference in acceptability of the word 'live' with Japanese accent between the participants with low and high proficiency of their target foreign languages. Those who had low proficiency judged more severely regarding 'live' compared with those who had high proficiency ( $p < .05$ ).

Table 12: Mann-Whitney Test on Acceptability -A J

		give	its	live	fill
Mann-Whitney		94.500	89.500	69.500	84.500
U					
Z		-.165	-.402	-1.372	-.633
Asymp. Sig. (2-tailed)		.869	.688	.170	.527

Table 13: Mann-Whitney Test on Acceptability -HL

		give	its	live	fill
Mann-Whitney		69.500	70.500	49.000	93.500
U					
Z		-1.339	-1.306	-2.303	-.210
Asymp. Sig. (2-tailed)		.180	.192	.021	.834

## 4 Discussion

It is interesting to compare the results in Schaiper (1992) with this study concerning raters' proficiency in the target language. Both studies show that those who have target language familiarity can understand the words and phrases

with the accents of the target languages.

It is also interesting to compare the results of perception between vowels and consonants in both studies. Schaiper (1992) shows that vowels are the most problematic features for listeners to perceive; however, this study shows that only the speakers of the target language prioritize the vowel /I/.

This study investigates the answers of the following questions:

- Q1. How intelligible and acceptable are the selected features of Japanese English (/i/ for /I/, /s/ for /θ/, and CVC for CC) for American English and Japanese speakers and participants with high and low proficiency of their respective foreign languages?
- Q2. Does the frequency of the word used in the corpus affect intelligibility and acceptability of Japanese English?

### 4.1 Intelligibility

#### 4.1.1 Comparison between Japanese and American

Regarding intelligibility, American participants seem to be stricter than Japanese participants as a whole. This is observed in the assessment of /i/ for /I/. Although it is not statistically significant, the similar tendency is found in CVC for CC. This tendency is understandable because it is natural that Japanese listeners can understand Japanese speakers' English with Japanese accent more easily than American listeners.

#### 4.1.2 Comparison between high and low proficiency

The difference is relatively smaller in this comparison than in the comparison between American and Japanese groups. However it seems that participants with high proficiency of languages judge more strictly than those with low proficiency. This is also natural because people with language training know how to assess the level of others in languages. This tendency is also found in Schaiper (1992) even though the target language is different (Spanish in his study).

### 4.2 Acceptability

The only thing that is significant is the difference in the feature of CVC for CC between American and Japanese participants. It seems that Japanese listeners are quite critical in imperfect consonant cluster. Japanese people *do* use imperfect consonant cluster, but in many cases, they may feel embarrassed. It may be useful to make Japanese

learners aware how intelligible and acceptable CVC is. This may help Japanese people feel more confident speaking Japanese English.

### 4.3 Frequency

Does the frequency of the word used in the corpus affect intelligibility and acceptability of Japanese English? For this question, the results of mean ranks and the Mann-Whitney tests showed which items in what comparison are influenced by frequency.

The results suggested there is no difference in the feature of /s/ for /θ/, whereas you can see some difference in the feature of /i/ for /I/.

#### 4.3.1 Intelligibility

American and Japanese participants perceived the word, 'its' significantly different. American raters were much more severe in rating this word. Although not statistically significant, similar tendency was found in 'live.'

This may explain that Japanese people, who do not have /I/ in their first language, do not much care or notice their mispronunciation of /I/; on the other hand American people knowing the distinction between /I/ and /i/ do notice and care the mispronunciation.

#### 4.3.2 Acceptability

As for acceptability, the difference comes from between a low and high proficiency comparison. The participants having high proficiency in the target languages seem to mark more critically in the sound of 'live' hearing /liv/ instead of /Iiv/. Each word in this minimal pair of live and leave occurs with almost the same frequency. This might be the reason why they would like to keep the distinction between them to avoid any confusion. Another speculation is that those who had high proficiency might know how to assess this vowel sound with some phonetic training.

## 5 Conclusion

This study discussed the intelligibility and acceptability of selected features of Japanese English which may relate to the selection of the features that could function as EIL.

As for intelligibility, the point should be raised that American listeners assess more severely than Japanese listeners regarding /i/ as a substitution of /I/ and CVC as a substitution of CC. It is also clear that people with high proficiency in the target language tend to have stricter rating for such Japanese accent.

Concerning acceptability, American listeners tend to be more generous than Japanese listeners

rating the sound of CVC for CC. This Phenomenon may tell us a potential of CVC to be a part of Japanese English.

Lastly, frequency of item use specifies that the most problematic and confusing feature was /i/ for /I/.

The recommendation for EIL from this study is:

1. Mastery of the vowel sound /I/ is necessary.
2. CVC for CC might be permissible.

For further study, it may be useful to investigate not just at the word and phrase level but also sentence and paragraph level so that one can apply the findings into more authentic situations. It will also be useful to have more diversity in participant's selection if one investigates EIL.

## 6 References

- Cross, J. (2002). A Comparison of Japanese and English Suprasegmental Pronunciation and an Aid to Raising Learner Awareness. *The Language Teacher* 26,4, 9-13.
- Davis, M. (n.d.). *VIEW (Variation In English Words and phrases)*. Retrieved December, 2006, from <http://view.byu.edu/>
- Jenkins, J. (2000). *The Phonology of English as an International Language*. Oxford: OUP.
- Jolly, Y.S. (2000). Articulatory Phonetics for In-service Teacher Training. *The Language Teacher* 24,8, 11-14.
- Kachru, B. (1992). *The Other Tongue*. Chicago: University of Illinois Press.
- Nakano et al. (2007). *On-Demand Internet Course Book: World Englishes and Miscommunications*. Tokyo: Waseda University International.
- Schaiper, K. (1992). Native Speaker Reaction to Non-Native Speech. *The Modern Language Journal* 76, 309-319.

## A Case Study on Developing a Vocabulary Testing

Norifumi Ueda<sup>1</sup>; Eiichiro Tsutsui<sup>2</sup>; Yusuke Kondo<sup>3</sup>; and Michiko Nakano<sup>4</sup>

<sup>1</sup>School of Foreign Language, Mejiro University, <sup>2</sup>Center for Foreign Language, Hiroshima International, <sup>3</sup>

Center for Foreign Language, Ritsumeikan University and <sup>4</sup>School of Education, Waseda University

<sup>1</sup>uedan@mejiro.ac.jp, <sup>2</sup>tsutsui@ic.hirokoku-u.ac.jp

<sup>3</sup>ykondo@fc.ritsumei.ac.jp, <sup>4</sup>nakanom@waseda.jp

### Abstract

The aim of this study is to develop a vocabulary test for testing vocabulary knowledge which is related to semantic and syntactic restrictions in choosing words in a context (the depth of vocabulary knowledge).

For the sake of this purpose, we chose basic verbs: *hear* and *listen* (Group 1); *see*, *look*, *watch*, *gaze*, and *stare* (Group 2); *say*, *talk*, *speak*, *tell* and *utter* (Group 3). The verbs in each group have similar meanings but occur in different syntactic or semantic circumstances. We made 8 items for Group 1, 20 items for Group 2 and 23 items for Group 3. In each test item, there is a blank to be filled in by the most appropriate verb(s). 280 university students participated in this test session. We selected the items by Winsteps and created a new vocabulary test with 37 test items.

### Keywords

Language testing, vocabulary test, vocabulary acquisition

### Introduction

Many types of vocabulary tests for second language learners have been developed. In the first section, we will review what characteristics are hypothesized as vocabulary knowledge, and what kinds of vocabulary tests have been developed. In the second section, we will report the developmental process in a lexical test. Then, we will discuss the problems found in this study.

## 1 Vocabulary knowledge and Vocabulary test

Many types of vocabulary tests for second language learners have been developed and these can be categorized into two types generally. One type is for testing the width of vocabulary knowledge like Nation's Vocabulary Levels Test (Nation, 1990) and the other, for testing the depth of vocabulary

knowledge such as Lex 30 (Meara & Fitzpatrick, 2000). In the followings, we review what vocabulary knowledge is hypothesized by researchers, and, then, what vocabulary tests estimate what kinds of vocabulary knowledge.

### 1.1 Vocabulary knowledge

Many researchers discuss the characteristics of vocabulary knowledge. Vocabulary knowledge is divided into two types: perceptive and productive vocabulary. (Nation, 1990).

Richards (1976) described vocabulary knowledge as follows:

1. The native speaker of a language continues to expand his vocabulary in adulthood, whereas there is comparatively little development of syntax in adult life.
2. Knowing a word means knowing the degree of probability of encountering that word in speech or print. For many words we also know the sort of words most likely to be found associated with the word.
3. Knowing a word implies knowing the limitations on the use of the word according to variations of function and situation.
4. Knowing a word means knowing the syntactic behaviour associated with the word.
5. Knowing a word entails knowledge of the underlying form of a word and the derivations that can be made from it.
6. Knowing a word entails knowledge of the network of associations between that word and other words in the language.
7. Knowing a word means knowing a semantic value of a word.
8. Knowing a word means knowing the different meanings associated with a word.

He claimed that learners should learn syntactic and semantic usage, derivational forms, associations, and word frequencies of the words.

Some researchers divide learners' vocabulary knowledge into two types of knowledge: depth and width of vocabulary knowledge. The learners' width of vocabulary knowledge means how much vocabulary learners know, while the learners' depth of vocabulary knowledge refers to how deeply or how detailed information the learners have about certain words.

## 1.2 Vocabulary knowledge from the perspective of psycholinguistics

There are many studies on vocabulary relations between concepts, and words both in L1 and L2, in L2 acquisition process in bilingual studies. In their study on relations between conceptual representations and two languages (L1 and L2) in the process of vocabulary acquisition by L2 learners, Ding et al. (2005) and Ijaz (1986) suggest that L2 learners should integrate the conceptual differences between two languages in the process of acquiring the other vocabulary.

Ueda (2007) studied the L2 learning process by Japanese learners of English, where English prepositions were used. The results of the study showed conceptual differences between two languages were integrated in the process of acquiring English vocabulary.

## 1.3 Vocabulary tests

Many vocabulary tests have been developed for evaluating learner's vocabulary size and knowledge. Different researchers developed different vocabulary tests to measure vocabulary knowledge, "depending on their view of vocabulary knowledge..., their preference for a particular dimension of knowledge, and their interest in either size or depth" (Laufer & Goldstein, 2004). Generally, however, these tests can be categorized into two types generally: One type is for testing the width of vocabulary knowledge like Nation's Vocabulary Levels Test (Nation, 1990) and the other, for testing the depth of vocabulary knowledge such as Lex 30 (Meara & Fitzpatrick, 2000). We will review how these two types of vocabulary tests have been made to evaluate width and depth of vocabulary knowledge in the next section.

### 1.3.1 Tests for width of vocabulary knowledge

Tests for width of vocabulary knowledge intend to evaluate how much the learners retain in their mental lexicon. The width of vocabulary knowledge is estimated by word frequencies.

In Vocabulary levels test (Nation, 1990), for example, items are randomly selected from each word frequency level: 2000-word level, 3000-word level, 5000-word level, the university

word level, and 10000-word level.

Each section of Vocabulary levels test consists of six words and three word definitions. The definitions in one section come from the words which are included in the higher level of the word frequencies: for example, the words from the 2000-word level use words in the first 1000 words for the definitions. In this test, the vocabulary levels of testees are estimated by the scores.

### 1.3.2 Tests for depth of vocabulary knowledge

Tests for depth of vocabulary knowledge intend to evaluate paradigmatic (synonyms), syntagmatic (collocations) and analytic knowledge (associations which represent one aspect or components of the meaning of the stimulus word and is likely to form part of its dictionary.) of vocabulary. (Read, 1993).

In Lex 30, for example, all test items are selected based on the following criteria:

1. All the stimulus words are highly frequent.
2. None of the stimulus words typically elicits a single, dominant primary response.
3. Each of the stimulus words typically generates responses which are not common words.

(Meara & Fitzpatrick, 2000)

In this test, testees were asked to write words using free word association. The collected data were analyzed according to the word frequencies. Scores were given by word frequency level. (For example, Words in Level 0 and 1 were given zero points, and those in other levels than Level 0 and 1, one point.)

## 1.4 Problems in Vocabulary tests for depth vocabulary knowledge

As we saw in the previous sections, the tests to evaluate width of knowledge use word frequencies to estimate learners' vocabulary size, whereas the tests for evaluation of depth of knowledge use word associations task.

Considering vocabulary knowledge from the perspective of L2 language acquisition, we should include syntactic and semantic aspects, as Richards (1976) pointed out (Vocabulary knowledge 3, 4, 6, 7, and 8), in vocabulary testing, especially in the tests for vocabulary depth knowledge in order to examine integration of conceptual differences between two languages in the acquisition process, which can show L2 learners' proficiency level to some extent. This is because associations are just one aspect of depth of vocabulary knowledge.

## 2 A case study of developing a test for depth of vocabulary knowledge

We aimed at developing a vocabulary test to evaluate depth of L2 vocabulary knowledge. In the test, synonyms were used, test items of which included syntactic and semantic information. The subjects were required to distinguish synonyms by using syntactic and semantic knowledge.

### 2.1 Subjects

280 university students participated as test subjects. They were from two different universities and they had various academic backgrounds: 8 subjects majors in robotics; 15, in information communication technology; 5, in architecture technology; 5, in socio-environmental design; 44, in Pharmaceutical Sciences; 22, nursing; 36, in psychology; 27, in social welfare; 86, in social information; and 32, in media.

### 2.2 Test items

As the test items, we chose basic verbs: hear and listen (Group 1); see, look, watch, gaze, and stare (Group 2); say, talk, speak, tell and utter (Group 3). The verbs in each group have similar meanings but occur in different syntactic or semantic circumstances. The subjects were required to know syntactic and semantic knowledge to use the synonyms. This vocabulary test aimed to estimate how much knowledge the learners have to use synonyms properly. We made 8 items for Group 1, 20 items for Group 2 and 23 items for Group 3. In each test item, there is a blank to be filled in by the most appropriate verb(s). The total number of the test items was 53. Here are some examples (in Table 1):

Table 1: Examples of the test items

[hear/ listen]	
Choose the proper word/words.	
1. Did you ( ) that noise? [hear/ listen]	
2. I ( ) to music	
3. I can ( ) somebody knocking the door.	
4. I told him not to go, but he wouldn't ( ).	
5. She doesn't ( )very well.	
6. You can ( ) that news again.	
7. ( ), I'm sure you can work this out.	
Do you ( ) me?	

The CGI on the internet was made for this test, and subjects accessed the internet and answer all the

items.

Table: The results of calculation by Winsteps

### 2.3 Method

Subjects were asked to answer the all the questions. There is no time limitation for answering the items.

All the items were analyzed to calculate infit and outfit by Winsteps 3.68.1., an IRT software, to select good items. All the test items, in fact, consisted of multiple choices (or the test items developed here were not simple one answer to one test item), because this type of test items could not calculated by Winsteps. Hence we counted one choice as one test item: For example, in a question (Did you ( ) that noise? [hear/ listen]), we regarded this as two test items.

### 2.4 Results

We used Rasch Modeling to evaluate item fit for each test item (in Table 2). We found a lot of misfit items<sup>1</sup> and 41 items were left. (See Appendix.)

The good items using verbs in Group 1 were as follows: for *hear*, Item No.1, 2, 3, 5, and 7; and for *listen*, item No. 4, 6 and 8.

The proper items using verbs in Group 2 were: for *tell*, item number 9, 10, 11, 18 and 21; for *say*, item No. 13, 14, 15, 19 and 20; for *speak*, item No. 9, 11, 12, 13, and 17; for *talk*, item No. 9, 13, and 16; and for *utter*, item No. 13, 14 and 20

The good fit items for the verbs in Group 3 were: for *see*, item No. 22, 26, 30 and 35; for *watch*, item No. 23, 24, 27, 32, 35 and 37; for *look*, item No. 24, 29 and 30; for *view*, item No. 24 and 33; for *stare*, item No. 33; for *gaze*, item No. 28, 34 and 35; and for *browse*, item No. 24, 25, 27 and 36.

Table 2: The results of the calculation by Winsteps

	RAW Score	Model Measure	INFIT		OUTFIT	
			MNSQ	ZSTD	MNSQ	ZSTD
Mean	64.8	0	1	0.1	1	0.1
S.D.	45.9	1.05	0.03	0.5	0.7	0.7
MAX.	272	2.67	1.09	3.1	1.43	3.3
MIN	5	-5.67	0.93	-1.4	0.78	-1.4

<sup>1</sup> In McNamara (1996), mean square (MNSQ) values greater than 1.3 show significant misfit, and the values below 0.75, significant overfit. But Winsteps adopts the criterion that MNSQ values between 0.5 and 1.5 are productive.

### 3 Conclusion

In this case study, we report the new type of tests dealing with the depth of lexical knowledge. This kind of test is very rare and very important to estimate L2 learner's correct vocabulary knowledge. This study can contribute very much to teaching setting. The multiple choices we adopted in this vocabulary test are difficult to calculate by Winsteps. However, to consider the guess ratio by three parameter model, we have to gather more than 1000 subjects. This point is still left to evaluate, so we will explore the better items in further study.

### References

- Dong, Y., Gui, S. and MacWhinney, B. (2005). Shared and Separate meanings in the bilingual mental lexicon. *Bilingualism: Language and Cognition* 8, 3. 221-238.
- Ijaz, H. (1986). Linguistic and cognitive determinants of lexical acquisition in a second language. *Language Learning*, 36, 401- 451.
- Laufer, B. and Goldstein, Z. (2004). Testing Vocabulary Knowledge: Size, Strength, and Computer Adaptiveness. *Language Learning* 54:3, 399-436.
- Reed, J (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing*, 10, 355-371.
- Richards, J. (1976). The role of vocabulary teaching. *TESOL Quarterly*. 10-1, 77-89.
- McNamara, T. (1996). *Measuring Second Language Performance*. Addison Wesley Longman.
- Meara, P. and Fitzpatrick, T. (2000). Lex 30: an improved methods of assessing productive vocabulary in an L2. *System* 28, 19-30.
- Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. Heinle and Heinle Publishers: Boston.
- Ueda, N (2007). Developmental Processes in the L2 Mental Lexicon. Unpublished M.Sc. Dissertation: The University of Edinburg.
7. ( ), I'm sure you can work this out.
8. Do you ( ) me?
- [Tell, say, speak, talk, utter.]
9. You should ( ) me your name and address.
10. I wouldn't ( ) so
11. Who was that you were ( ) to at the party?
12. He refused to ( ) his name.
13. Do you ( ) English
14. He never ( ) a word.
15. We need to ( ) before meeting.
16. She ( ) a sigh.
17. The red light ( ) you when the machine is ready to use.
18. I would like to ( ) about our culture.
19. The stress of work ( ) on their marriage.
20. The clock ( ) it is five now.
21. Most babies start to ( ) by 18 months.
- [see, watch, look at, view, stare, gaze, browse]
22. I can( ) that you're not very happy with the situation.
23. Jack ( ) them slowly crime up the wall.
24. Sorry, I wasn't ( ).
25. Tourists came to ( ) the gardens every year.
26. She doesn't like being ( ) at.
27. She ( ) at me in disbelief when I told her the news.
28. She ( ) through the magazine.
29. I can ( ) why.
30. I ( ) the baseball game.
31. ( ) at me.
32. The building is more beautiful when you ( ) it from the river.
33. It is rude to ( ).
34. He ( ) in a library.
35. Can I( ) your ticket?
36. Can you ( ) my bag?
37. She was ( ) vacantly into the space.

### Appendix: Test items after the selection by Winsteps.

[hear/ listen]

1. Did you ( ) that noise?
2. I ( ) to music
3. I can ( ) somebody knocking the door.
4. I told him not to go, but he wouldn't ( ).
5. She doesn't ( ) very well.
6. You can ( ) that news again.

# Range Analysis of Cross-Cultural Distance Learning (CCDL) Reflection Work Sheet: A Pilot Study

Satoshi Yoshida, Hideyuki Kumaki and Akiko Watanabe

Waseda University

satoshi-tko.jpn@akane.waseda.jp, t-kumaki@mua.biglobe.ne.jp,

aki49@fuji.waseda.jp

## Abstract

This study analyzes the reflection work sheets written by Cross-Cultural Distance Learning (CCDL) participants and examines if there are significant differences in vocabulary use among three different nationalities: Chinese, Taiwanese and Japanese (Waseda) students. We employed range analysis and adopted the resulting outputs called Lexical Frequency Profiles (LFP: Laufer and Nation, 1995) as the parameter of lexical richness. All the range analyses were run with 16 base word lists which were developed by Paul Nation on the basis of the vocabulary frequency levels observed in British National Corpus (BNC), and the concept of word families (Bauer and Nation, 1993). The participants in this study were 66 students who participated in CCDL courses in the fall semester of 2008, and submitted the reflection work sheets as a class assignment. We grouped these students into three groups: Chinese students (N=22), Taiwanese students (N=22) and Japanese students (N=22), and compared the resulting percentages of word types calculated for each frequency level according to the base word lists. The result showed that Taiwanese students used frequent words in their reflection sheets more than Chinese and Japanese students. The result also implied that Japanese students tended to use relatively infrequent words more than Taiwanese students.

## Keywords

Range Analysis, Lexical Richness, Vocabulary Use, Lexical Frequency Profiles, Word Types, Word Families, Cross-Cultural Distance Learning

## 1 Introduction

### 1.1 CCDL at Waseda University

Cross-Cultural Distance Learning (CCDL) courses at Waseda University in 2008 were joint cyber interactive courses with East Asian Universities. These courses provided theme-based discussion using on-line chat system called LiveOn. CCDL

courses were designed to offer the participating students not only the opportunities to communicate in English with Asian students, but also the opportunities to develop their “transferable skills” such as facilitation skills, which are adaptable generic skills for life and work (Nakano et al, 2008). In order to enhance these skills, for example, every student was expected to be a facilitator in their joint sessions. Table 1 summarizes the schedule of CCDL classes dealing with Social and Global Issues as the main discussion topic.

Table 1: Schedule Overview (as cited in Nakano, Yokota and Karseras, 2009: 1)

Social&Global Issues Schedule Overview				
WEEK	Preparation/Join	TOPIC	SKILLS	HOMEWORK
W1	Preparation	Introduction	Facilitations	None
W2			Intercultural Translation	HW1: Reading facilitation skills
W3			Live On	HW2: Reading & Video Dr. Nakano's lecture & conversation
W4	Preparation	Happiness Factors	Individualism/	HW3: Reading Images of
W5	Joint Class 2		Collectivism	HW4: Reading Value Orientations: Individualism and Collectivism
W6	Preparation	Family Roles	High/Low context	HW5: Reflection Paper Topic 1&2
W7	Joint Class 3		Communication	HW6: Research Family Roles
W8	Preparation	International Students	Relationship	HW7: Reading Support network for International Students
W9	Joint Class 4		development	HW8: Research international Students at Your University
W10	Preparation	Climate Change	Agreement	HW9: Reflection Paper Topic 3&4
W11	Joint Class 5		Styles	HW10: Research Climate Change
W12	Post Joint Class	Presentation	Multiple Intelligences	HW11: Power Point
W13			PowerPoint	HW12: Rehearsal
W14			Presentation	
W15			Delivery	

As in Table 1, CCDL courses consist of three types of activities: Preparation sessions, Joint lessons and Post Joint lessons. In 2008, these CCDL courses dealt with four topics: Happiness Factors, Family Roles, International Students, and Climate Change; as current social and global issues. All the participating students were supposed to study the same topics along with the schedule by studying the same textbook.

## 1.2 Reflection work sheet

The reflection work sheets collected in this study were one of the class assignments, in which the students were supposed to summarize what they learned from joint lessons as well as the preparation activities they joined. The main purpose of this reflective writing is to make students more conscious about their own subjective experiences, specifically reflecting back on the discussion and internalizing their learning by writing, and noticing changes in their own perceptual awareness (Nakano et al, 2009). That is, these reflection papers could be one of the indicators which would reflect to what extent the students could grasp the complicated issues in each class, as well as, to what extent the students could come to function the academic words (e.g., key words for each topic) as their productive vocabulary. Therefore, it would be valuable to put their papers into analysis, in order to evaluate their lexical development in terms of lexical richness.

For data collection, we sought out the classes in which both Waseda and foreign students submitted their reflection work sheets. We should carefully choose the classes because all the Waseda students were supposed to submit their reflection work sheets after each joint session whereas foreign students did not have to do so. Because it was optional for them, in some classes, we could not collect the reflection work sheets from the foreign students. In other classes, on the other hand, all the foreign students opted to write their reflective sheets based on the class schedule shown in Table 1. Because we could expect all the participating students in the latter classes to be highly motivated to study the topics in CCDL classes, we collected the reflection work sheets from these classes and compare the participants' vocabulary use among CCDL participants with different nationalities.

## 2 Background

### 2.1 Lexical Richness and L2 Learner's written productions

According to Laufer and Nation (1995), there are several factors which would affect the quality of L2 written composition, such as lexical richness, familiarity with the topic and writing skills. Among these affecting factors, the concept of lexical richness has drawn the attentions from L2 researchers, because, if a certain measure could reliably estimate the richness, it would not only be the indicator of the quality of some L2 production, but also provide some useful information about the vocabulary size and L2 proficiency level of the learners. In fact, the concept of lexical richness has been frequently adopted or adapted in the

studies of L2 vocabulary use (e.g., Laufer, 1991).

As for the measure of lexical richness, it can be estimated by quantifying the varieties of active lexes observed in L2 productions. According to Laufer and Nation (1995), there are various formulas to estimate the richness, -e.g., lexical originality, lexical density, lexical sophistication and lexical variation (i.e., type/token ratio). Laufer and Nation (ibid), however, pointed out that these frequently used measures would possess some shortcomings in their calculations. In order to overcome these shortcomings, they introduced a new measure of lexical richness called Lexical Frequency Profile (LFP: Laufer and Nation, 1995), in which lexical richness were to be measured based on some frequency levels of vocabulary. They claimed that LFP would be a reliable measure of lexical richness, showing its considerable correlations with direct measures of vocabulary size, the feasibility to differentiate the learners with different language proficiency, and the consistency of the calculation. Recently, therefore, LFP has been frequently employed in the studies of productive L2 vocabulary (e.g., Sugimori, 2008).

In this study, therefore, we empirically adopted LFP as a measure of lexical richness and compare the differences in vocabulary use among CCDL participants with different nationalities. In so doing, for data analysis, we employed Range Program and the attached base word lists, both of which enable us to run range analysis for producing LFP, as an analytic tool. The following section shows how to use Range Program and how to read LFP.

### 2.2 Range Analysis and LFP

In this section, we briefly introduce how Range Program works for estimating lexical richness and what the LFP shows us, according to the instruction papers offered by Paul Nation. Range analysis can be implemented by installing the written compositions to be analyzed and some Base Word Lists (BWL) into Range Program. The word lists should be edited and separated on the basis of some frequency levels of the words contained in it, because LFP is to be calculated for frequency level. Now, two set of useful BWL are available on the website of Paul Nation. One of the BWL set is called General Service List, which is separated into 3 types of BWL: most frequent 1000 word families-list, second frequent 1000 families-list and University level Word list (for details, see Laufer and Nation, 1995). Another set of BWL is developed by Paul Nation, on the basis of frequency levels observed in British National Corpus (BNC). This set contains 16 lists (for details, see next section).

In range analysis, the program divides the sorts of vocabulary, used in the written composition, into different frequency levels, according to the levels defined by BWL. Therefore, the LFP, an output of range analysis, shows the summary of the resulting proportions of words at different frequency levels. Table 2 shows an example of LFP, which was calculated by analyzing the abstract of this paper with installing BNC 16 word lists.

Table2: LFP of ABSTRACT of this study (example)

WORD LIST	TOKENS/	%	TYPES/	%	FAMILIES
1	138/	70.77	70/	62.5	59
2	25/	12.82	18/	16.1	16
3	5/	2.56	3/	2.68	1
4	9/	4.62	7/	6.25	7
5	2/	1.03	2/	1.79	2
6	0/	0	0/	0	0
7	1/	0.51	1/	0.89	1
8	3/	1.54	1/	0.89	1
9	1/	0.51	1/	0.89	1
10	0/	0	0/	0	0
11	0/	0	0/	0	0
12	0/	0	0/	0	0
13	0/	0	0/	0	0
14	2/	1.03	1/	0.89	1
15	3/	1.54	3/	2.68	3
16	0/	0	0/	0	0
NiL	6/	3.08	5/	4.46	?????
Total	195		112		92

Note: NiL stands for “Not in any list”.

As Table 2 shows, LFP is calculated for Tokens, (Word) Types and (Word) Families, in which the number of words and resulting percentages are given. Along with the results of 16 word lists, we can find the number of the word types in the result of NiL, which means “word types not found in any list”. That is, they are a set of words which are not found in these 16 lists. In this analysis, we can find that 5 words, “LFP”, “Waseda”, “CCDL”, “BNC” and “Cross-Cultural” are not found in any list of BNC 16 word lists. Regarding the case of “Cross-Cultural”, it is categorized in NiL because of a hyphen occurred in the two words. If we separate the words into “Cross” and “Cultural”, the program categorizes them into the appropriate frequency levels. As for the question mark which appears in the column of FAMILIES, it indicates the possibility that the word in NiL may or may not be categorized into word families.

The important values in LFP are resulting percentages of each measure. The most frequently reported values in the previous studies were those

of (Word) Families, which would be calculated with some treatments for the words in NiL (e.g., Laufer and Nation, 1995). In this study, however, we adopted the resulting percentages of Word Types as the parameter of lexical richness and used the mean percentages for group comparison. We believed that our decision would be reasonable because the concept of Word Families seemed to be developed, in particular, for reducing the vocabulary load in READING, focusing on base and affixes of the words (Bauer and Nation, 1993). Therefore, several criteria that were referred to in grouping the bands of words for receptive use would not be applicable to categorizing L2 productive vocabulary (e.g., predictability of the meaning of affix). Hence, for group comparison, we employed the resulting percentages of Word Types calculated for each frequency level.

### 3 Method

#### 3.1 The Participants

The participants in this study were 66 students who were all enrolled in CCDL courses dealing with *Social and Global Issues* as the main discussion topic. All the participants submitted their first reflection work sheet as a class assignment. These students were divided into three groups according to his/her nationality: Chinese students group (N=22), Taiwanese students group (N=22) and Japanese (i.e., Waseda) students group (N=22).

#### 3.2 Reflection Work Sheet

The reflection work sheets analyzed in this study were written on topics 1&2 -i.e., “introduction” and “happiness factors”. Because all the participants took some lectures or joined the preparation classes, we expected them to have some familiarity with the topics. The students were assigned to write the paper in 400-450 words. The mean length of the papers written by Chinese students, Taiwanese students and Japanese students were 442.5, 459.7 and 386.6 respectively.

#### 3.3 Analysis

##### 3.3.1 Range Analysis with BNC 16 word lists

We employed range analysis for analyzing the learners’ reflection work sheets and adopted the resulting LFP as the parameter of lexical richness.

All the range analyses were run with BNC 16 word lists. Therefore, resulting LFP showed the proportions of the vocabulary used in the composition, by 17 categories (BWL1-16 plus NiL), as in Table 2. According to user-guide attached to Range Program, among 16 BWL, the lists 1-14 contains each 1000 word family which were categorized on the basis of frequency levels,

whereas lists 15 and 16 are not frequency-based, but the lists of pronouns and of interjections respectively. Therefore, it should be noted that BWL 1-14 precisely indicates the frequency levels of vocabulary, that is, base word list 1 stands for the list which contains most frequent 1000 word families, whereas basic word list 14 contains most infrequent 1000 word families.

### 3.3.2 Data Processing

Prior to the analysis, we edited all the reflection sheets to make the resulting LFP as stable. That is, we discussed how to deal with errors (e.g., miss-spelled words), proper nouns, hyphenated words and words or sentences which were not produced by CCDL participants, which frequently occurred in the reflection work sheets. As the tentative solutions, we followed four procedures as follows.

#### (1) Errors

If any, we remained all the words with some mistakes in spelling or usage, intact, so that we could observe them in NiL list.

#### (2) Proper Nouns

If any, we DID NOT exclude the proper nouns that appeared in body text of the reflection sheets, so that if they were not included in base word list 15, we could observe them in NiL list.

#### (3) Hyphenated words

If any, we remained all the hyphenated words intact (e.g., “self-confidence” and “self-efficacy”), so that we could observe them in NiL list.

#### (4) Private Information etc.

We excluded all the private information about the writer (i.e., Name, student ID), and information about the courses and the assignment (e.g., the name of the courses, the title of the paper), and sentences that were not produced by the writer (e.g., questions and tasks offered by the instructor) from the text with the help of “ignore <>”function installed in Range Program.

We edited all the reflection sheets based on the four procedures above, and run range analysis on each reflection paper. Although our procedures were different from those in Laufer and Nation (1995) or those recommended by instruction papers by Paul Nation, we believed that our procedures would enable us to see the detailed contents of the students’ reflection work sheets, in terms of the words in NiL. We will discuss the features of these words in section 3.3.

## 4 Results and Discussions

### 4.1 Range Analysis

Table 3 shows the mean percentages and standard deviations of Word Types observed in each frequency level.

Table 3: Mean percentages and standard deviations of Word Types for each frequency level

Word List	Chinese		Taiwanese		Japanese	
	M	SD	M	SD	M	SD
1	80.84	3.67	83.92	4.80	80.88	4.14
2	11.53	2.07	9.14	2.66	10.84	2.35
3	2.10	1.32	2.13	1.22	2.59	1.21
4	1.17	0.75	0.85	0.83	1.26	0.77
5	0.51	0.50	0.34	0.49	0.49	0.52
6	0.37	0.51	0.35	0.41	0.68	0.54
7	0.09	0.19	0.06	0.16	0.20	0.40
8	0.21	0.28	0.49	0.55	0.33	0.42
9	0.19	0.26	0.04	0.19	0.18	0.26
10	0.05	0.17	0.02	0.09	0.04	0.13
11	0.15	0.27	0.16	0.34	0.04	0.14
12	0.00	0.00	0.02	0.09	0.02	0.10
13	0.04	0.14	0.00	0.00	0.08	0.18
14	0.02	0.09	0.00	0.00	0.04	0.14
15	0.04	0.12	0.04	0.19	0.19	0.37
16	0.00	0.00	0.03	0.14	0.02	0.10
NiL	2.67	2.66	2.40	1.92	2.10	0.91

Note: NiL stands for “Not in any list”. (N=22) for three groups.

As in Table 3, the types of words in BWL 1, 2 and 3 could cover about 95 percent of the productive vocabulary observed in the reflection sheets, written by three groups of CCDL participants.

### 4.2 ANOVA on BWL 1, 2, 3 and above 3000

We run One-way independent ANOVA (analysis of variance) to compare the vocabulary use among three groups, in terms of the resulting percentages of Word Types calculated for each frequency level. In so doing, we combined the percentages observed in the BWL 4-14 and regarded the cumulative percentage as the proportion of “above 3000” frequency level; in which the percentages of BWL 15 and 16 were also combined because of their low percentages. As shown in Table 3, a large amount of words that appeared in the reflection work sheets could be covered by BWL 1, 2 and 3, so that it would be meaningless to compare the percentages of vocabulary in BWL 4-16 separately. We believed that the inclusion of these percentages would be reasonable because Laufer and Nation (1995), in which GSL word lists were employed,

also carried out this sort of procedure. Hence, we put the resulting percentages of Word Types calculated for BWL 1, 2, 3 and for above 3000 in ANOVA. As for resulting percentage of NiL, we did not statistically analyze because each of them contained some errors, proper nouns or hyphenated words (see 2.3.1). Instead we will discuss the features of these words in section 3.3.

Table 4 summarizes the results of ANOVA with mean percentages and standard deviations of Word Types observed in BWL 1, 2, 3 and above 3000.

Table 4: Mean percentages and standard deviations of Word Types for each BWL and result of ANOVA

	BWL 1	BWL 2	BWL 3	above 3000
Chinese	80.84	11.53	2.10	2.84
(SD)	3.67	2.07	1.32	0.99
Taiwanese	83.92	9.14	2.13	2.41
(SD)	4.80	2.66	1.22	1.80
Japanese	80.88	10.84	2.59	3.59
(SD)	4.14	2.35	1.21	1.85
F-test	3.856*	5.933**	1.04	3.10
p-value	0.026	0.004	0.361	0.052

Note: \* indicates  $p < 0.05$ . \*\* indicates  $p < 0.01$ .

As Table 4 indicates, we could find significant differences in the percentages of BWL 1 among three groups. In *post hoc analysis* with Duncan procedure, we found that Taiwanese students used more vocabulary in BWL1 than Chinese and Japanese students. Between Chinese and Japanese, we could not find significant differences. In the comparison of the percentages of BWL2, we also found significant differences among three groups. In *host hoc analysis*, it was revealed that Taiwanese students used significantly less lexes in BWL2 than the other groups. Between Chinese students and Japanese students, we could not find significant differences. With regard to BWL 3, we could not find significant differences among three groups. Lastly, in the comparison of the percentages calculated for above 3000 vocabulary level, we could not find significant differences, but found significant tendency among three groups. That is, *host hoc analysis* implied that Japanese students would have used more types of words in the list of above 3000 than Taiwanese students.

To sum up, we can say that Taiwanese students relied on more frequent words (i.e., those in BWL1) in writing their reflection work sheets whereas the other two groups used almost same amount of the words in BWL 1 and 2. The result also implied the possibility that Japanese students might use more infrequent words (i.e., those in BWL 4-14) than Taiwanese students. Regarding the mean

length of the reflection sheets by Taiwanese students, however, we should carefully interpret these results.

### 4.3 Word Types in NiL

The following lists (Tables 5-8) show the words which were categorized into NiL. We can summarize the features of these words as follows: (1) Words which were spelled correctly but not categorized into BNC 16 word lists; (2) Hyphenated words; (3) Proper nouns; and (4) Errors.

Table 5: Words which were spelled correctly but not categorized into BNC 16 word lists <22 words>

CHINGLISH	INTERCULTURAL
COLLECTIVISTIC	INTERNET
CONVERSATIONISTS	MIC
CONFUCIAN	OVERUSE
CONFUCIANISM	OVERUSED
EASYGOING	ROOMMATE
EXAGGERATEDLY	SATISFACTIONS
FRIENDNESS	SCHOOLWORK
FUNDAMENTALITY	SOPHOMORE
HOMETOWN	WEBSITE
INDIVIDUALLISTIC	WHITEBOARD

As in Table 5, there were 22 words which were properly spelled, but categorized into NiL. Therefore, we can assume that, in general, words in this category can be considered as “high level”. Furthermore, regarding the relationship with the topic of the reflection sheets, we can find that several words that are closely related to the topics appear in this list. This result might imply that the participants could come to activate the complicated words that taught in CCDL classes as their productive vocabulary. In order to examine this fact, we will be required to make a new word list which includes these key words for following study. On the other hand, we believe that seemingly frequent words, which can be used in daily conversation (e.g., “WEBSITE” and “INTERNET”), should be included in some lists according to their frequency levels.

Table 6: Hyphenated words <13 words>

ALL-ROUND	SELF-CONCERNED
CHINESE-ONLY	SELF-CONFIDENCE
FACE-TO-FACE	SELF-EGOIST
FACE-CONSCIOUSNESS	SELF-RELIANCE
HIGH-TECHNOLOGY	WELL-BEING
LOW-MIDDLE	WELL-PREPARED
SELF-CENTERED	

As in Table 6, we found 13 hyphenated words in NiL as we intended. We believed that, if they were correctly used in the reflection sheets, it should be considered as high-level academic terms.

Table7: Proper nouns <10 words>

DAIREN	NATTO
GYOZA	OSECHI
KANJI	OSETHI
LIVEON	TAKOYAKI
LIVE-ON	WASEDA

In Table 7, we can find 10 proper nouns. Because these data were collected in CCDL courses, we could find a word, “Live-on” or “Liveon”, appeared more than 120 times in their compositions. A word “Waseda” can be also predicted to frequently appear in the compositions. Therefore these words should be included in BWL 15; otherwise we should make a new BWL for proper nouns relevant to CCDL context.

Table 8: Errors <13words>

ADVICES	ESPERIENCE
AFFAIRES	FAMILES
AVARIABLE	FRIENDNESS
BIGGIST	JAPANSES
CONCERATE	MOBLEPHONES
DISSCUSSION	PSYSICAL
EQUIPMENTS	

Table 8 shows 13 types of “errors”, that is, words with some mistakes in spelling or usage. Though we could have corrected their mistakes to be counted as a Word Types as in Laufer and Nation (1995), we did not correct them to see if there would be common errors among CCDL participants. The remarkable features among these 13 are errors which are collective nouns in Native Speaker Englishes (therefore they should be singular), but which are acceptable in Asian Englishes as a Lingua Franca, as in “ADVICES” and “EQUIPMENTS”. Because all the participants in this study were non-native speakers (NNS) of English, these errors should be saved in the word list for “NNS’ common errors”, and utilized with some BWL to estimate the error rates.

## 5 Summary

To sum up the findings in this study, we found that a large amount of vocabulary used in the reflection work sheets can be covered by the BWL 1, 2 and 3 of BNC word lists. Therefore, we should discuss if the lists would match the vocabulary size of CCDL participants.

In the comparison of vocabulary use, we found that Taiwanese students used significantly more words in BWL 1 than the other group. With regard to BWL 2, Taiwanese students used significantly less words than the other two groups. The result also implied the possibility that Waseda students tended to use more infrequent words than Taiwanese students. It should be noted again that these results might be caused by the differences in each mean length of the reflection work sheets.

Through the examination of the words in NiL, we believed that our procedures mentioned in 3.3.2 worked well for deeper investigation of seemingly problematic words appeared in the reflection sheets. At the same time, we recognized that we should analyze the data once again with some necessary treatments for the words in NiL to see more stable LFP as a reliable measure of lexical richness.

## References

- Bauer, L. and Nation, I.S.P. (1993). Word Families. *International Journal of Lexicography*. 6., 3. 235-79.
- Laufer, B. and Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*. 16, 3: 307-322.
- Laufer, B. (1991). The development of L2 lexis in the expression of the advanced language speakers. *Modern Language Journal*. 75, 4. 440-448.
- Nakano, M., Murao, R., Yokota, M., Karseras, A. S., Ito, S., and McDermott, D. (2008). Developing transferable skills and social intelligence through theme-based cross cultural distance learning 2008 spring semester. In JACET select committee (eds.), *Information Communication Technology Practice & Research 2007* (pp.191-199). Waseda University International.
- Nakano, M., Yokota, M., and Karseras, S. A. (2009). *Cross cultural distance learning social and global issues: Teacher’s note*. Waseda University International.
- Sugimori, N. (2008). Eibun-raiting-ni-okeru-Lexical Frequency Profile –no-teijyousei-ni-kan-suru-kousatsu. [A study of the stability of lexical frequency profile for writing composition]. *The proceedings of 34<sup>th</sup> conference of the Japan Society of English Language Education*. 120-121.

## Website

- Nation, P. *The range programme and the BNC lists*, retrieved <http://www.victoria.ac.nz/lals/staff/paul-nation/nation.aspx>.

# Cross-Cultural Distance Learning (CCDL) and the Learner's Motivation toward the CCDL CMC Activities: A Survey on Three Types of CCDL Classes

Satoshi Yoshida and Michiko Nakano

Waseda University

satoshi-tko.jpn@akane.waseda.jp, nakanom@waseda.jp

## Abstract

The main purpose of this study is to compare the learners' motivation toward CCDL Computer Mediated Communication (CMC) chat activities by focusing on the features of class types: autonomy supportive student-centered elective English courses and teacher-oriented compulsory English classes. On evaluating the motivational differences among CCDL participants, we adopted theoretical framework of Self-Determination Theory (SDT: Deci and Ryan, 1985, Ryan and Deci, 2000 and 2002). We employed a questionnaire based on Language Learning Orientation Scale (LLOS: Noels et al., 2000) which was designed to assess learner's motivation based on SDT. The participants in this study were 204 Waseda university students who participated in CCDL classes in the fall semester of 2008. We factor-analyzed the data and compared the motivational differences by each factor score. The result showed that the students in autonomy supportive student-centered classes showed higher degree of self-determination than those in teacher-oriented English classes. This finding demonstrated a strong relationship between autonomy supportive class environment and motivational enhancement. This result also supports the notion that a learner's motivation would be less if the class was compulsory while it would be more desirable if one chose the class based on his/ her decision (Nakano, 2006).

## Keywords

Cross-Cultural Distance Learning, Motivation, Self-Determination Theory, Autonomy Supportive Class, Student-Centered Class, Teacher-Oriented Class

## 1 Introduction

### 1.1 CCDL and Motivation

Nakano (2006) categorized the types of CCDL

classes into two types: elective CCDL courses (i.e., CCDL courses especially designed as CCDL) and compulsory CCDL classes (i.e., regular English classes with extra curricular CMC chat sessions), and explained the predictable motivational differences between the learners who participated in the former type of CCDL and those in the latter one, referring to the Self-Determination continuum (see Figure 1) posited by SDT (Park, 2006; Ryan and Deci, 2002). In her explanation, she claimed that the learner's motivation would be less if the class was compulsory while it would be more desirable if one chose the class based on his/ her decision.

Types of Motivation	Amotivation	Extrinsic Motivation				Intrinsic Motivation
Types of Regulation	Non-Regulation	External Regulation	Introjected Regulation	Identified Regulation	Integrated Regulation	Intrinsic Regulation
Quality of Behavior	Nonself-determined					Self-determined

Figure 1: The Self-Determination Continuum (as cited in Ryan and Deci, 2002: 16)

### 1.2 Our Studies of Motivation

In order to test Nakano's (2006) assumption, we conducted two pilot studies among Waseda University students (Nakano and Yoshida, 2008; Nakano, Yoshida and Owada, 2008), employing a questionnaire based on Language Learning Orientation Scale (LLOS: Noels et al., 2000). In these studies, we verified the applicability of questionnaire items, and, through the examination of motivational comparison, we found the possibility that autonomy supportive class environment which can be seen in elective CCDL courses (e.g., student-centered approach) would be closely related to motivational enhancement.

In this study, therefore, we categorized the participants into more practical groups, specifying the features of elective CCDL classes as example of autonomy supportive student-centered classes and those of compulsory CCDL classes as teacher-oriented classes, and compare the motivational differences among the participants.

## 2 Survey

### 2.1 Participants

The participants in this study were 204 Waseda university students who were all enrolled in CCDL activities. 67 out of 204 were the students participated in elective CCDL classes -e.g., Global Literacy Course and Theme-based CCDL course (Elective CCDL group); 85 were English majors participated in compulsory CCDL activities (extra chat group, English Majors); and 52 were non-English majors in compulsory CCDL activities (extra chat group, Non-English Majors).

### 2.2 Questionnaire

The questionnaire used in this study was based on 21 items in LLOS (Noels et al., 2000) and 3 items in Park (2006), which can be categorized into 7 subscales: Intrinsic Motivation for Knowledge (IMK), Intrinsic Motivation for Accomplishment (IMA), Intrinsic Motivation for Stimulation (IMS), Identified Regulations within Extrinsic Motivation (EMID), Introjected Regulations within EM (EMINTRO), External Regulations within EM (EMEX), and Amotivation (AMOT). We modified them for CCDL context. In total, 24 items were distributed among the participants. All the items were ordered randomly, and the participants rated each item along 7-point Likert scale.

### 2.3 Procedure

We collected the data during fall semester of 2008. Along with the paper-based survey, we distributed the questionnaire online. The web-site for online questionnaire was available during October, 2008. On the first page, the instructions about how to answer the questions were given. It was also mentioned that their answers would never affect their course grade, and that their private information would never be revealed. The students could access the site anytime they would like to answer the questionnaire.

### 2.4 Analytic Procedure

In this study, we analyzed the data in two ways. In Analysis 1, we followed the same procedures as in our Pilot Experiments (see Field, 2005; Nakano and Yoshida, 2008). That is, if any, items that had high correlation or seemingly no correlation with other items were excluded from the analysis, referring to correlation matrix as well as its determinant.

As for Analysis 2, we factor-analyzed all the items by specifying the number of factors to be extracted as predicted. We did not refer to the determinant of correlation matrix to see construct

measured by overall items and to discuss if the high correlation(s) as well as relatively low correlation(s) would cause a serious problem in the result of factor analysis. This method would be well worth considering in the sense of supporting the results obtained in Analysis 1, which sometimes cause excessive item reduction. This process would be also important for developing the questionnaire for future study.

We compared the students' motivational differences among three types of CCDL classes by mean factor scores. Both in Analysis 1 and 2, we will discuss the result of motivational comparison.

## 3 Results and Discussions

### 3.1 Analysis 1

#### 3.1.1 Exploratory Factor Analysis

In this section, we analyzed the data following the same procedures as in our Pilot Experiments 1 and 2. We employed Maximum Likelihood Technique followed by promax rotation to obtain the oblique solution. The alpha decision level set for *test of goodness of fit* was 0.01.

Prior to the analysis, we checked each item and determinant of correlation matrix, and found that items 13, 14 and 21 had floor effects; items 6, 16 and 18 had high correlation with other item; items 8, 14 and 20 had relatively low correlations with other items. Accordingly the determinant of the correlation matrix became smaller than required value, suggesting that we should exclude those problematic items until the determinant became qualified. Finally, we excluded all the problematic items from the analysis (i.e., 8 items mentioned above). Therefore, we factor- analyzed remaining 16 items.

Because several items were excluded before the analysis, first factor extraction was based on Kaiser's criterion. As a result, Heywood case happened in the analysis. As a solution for this problem, we examined whether we can exclude more items, retaining the computational adequacy. This procedure would be reasonable because we have already excluded some seemingly problematic items from the analysis in order to improve the determinant of value. Among the remaining 16 items, we found that excluding item 2 led to an adequate solution. Therefore we continued the analysis without item 2.

In the reanalysis, KMO Statistics of Sampling Adequacy was 0.914. Bartlett's Test of sphericity yielded statistical significance. All the items had adequate values in KMO statistics for individual items, ranging from 0.768 to 0.949. As a result, test of goodness of fit yielded statistical significance for three-factor model ( $\chi^2=116.914$ , *df*

= 88,  $p < 0.001$ ). In the pattern matrix, there were no items which were cross-loaded onto the extracted factors. Because there were no items to be excluded from the analysis, we can reject the null hypothesis, as the result of test of goodness of fit indicates. Accordingly, we discussed whether three factors were sufficient for the data set by increasing the number of factors to be extracted.

In second analysis, we specified the number of factors to be extracted as 4 factors. All the criteria we referred to in the analysis were verified as in first analysis. As a result, test of goodness of fit yielded statistical non-significance for four-factor model ( $\chi^2=73.736$ ,  $df=51$ ,  $p=0.02$ ). Table 1 shows the pattern matrix obtained from second analysis.

Table1: The Pattern Matrix of Second Analysis

Label	Item	Factor			
		1	2	3	4
IMA2	item11	<b>0.93</b>	0.07	0.00	-0.14
EMEX2	item10	<b>0.75</b>	-0.25	0.05	0.21
IMA3	item1	<b>0.72</b>	0.12	-0.10	-0.08
EMINTRO1	item24	<b>0.68</b>	-0.02	0.11	-0.03
IMA1	item15	<b>0.43</b>	0.16	0.04	0.21
IMS3	item3	<b>0.40</b>	0.14	-0.08	0.33
IMK1	item5	-0.07	<b>0.90</b>	-0.01	-0.05
IMK2	item9	0.06	<b>0.85</b>	-0.01	-0.06
IMK3	item22	-0.07	<b>0.79</b>	0.06	0.10
EMEX4	item4	0.06	0.08	<b>0.95</b>	-0.10
EMINTRO4	item7	0.10	0.03	<b>0.89</b>	0.13
EMEX1	item23	-0.23	-0.05	<b>0.61</b>	0.04
AMOT3	item12	0.06	-0.14	<b>0.42</b>	-0.11
EMID1	item19	0.18	0.05	0.03	<b>0.75</b>
EMID2	item17	0.08	0.31	-0.07	<b>0.54</b>
% of variance		45.85	9.48	6.08	1.87
eigenvalue		7.29	1.74	1.26	0.71
Correlation		1	0.64	-0.47	0.72
			1	-0.5	0.72
				1	-0.59
					1

Note: Each item label stands for the abbreviations of subcategories.

We renamed each extracted factor according to the items contained in them. The parenthesized value for each item indicates Cronbach's  $\alpha$ .

**Factor1: '(Transition from EM to Intrinsic) Motivation for Accomplishment'**  
(6 items,  $\alpha = 0.870$ )

**Factor2: 'Intrinsic Motivation for Knowledge'**  
(3 items,  $\alpha = 0.860$ )

**Factor3: '(Transition from Amotivation to) Extrinsic Regulations by External Pressures':** (4 items,  $\alpha = 0.823$ )

**Factor4: 'Identified Regulations'**  
(2 items,  $\alpha = 0.879$ )

Factor 1 stands for motivation for accomplishment. This factor would reflect the specific relation between motivation for pursuing accomplishment in CCDL activities and the students' future directions as well as concerns for their partners in the class. In this sense, however, this factor no longer reflects the participants' intrinsic predispositions for leaning English. Therefore we tentatively labeled this factor "Motivation for Accomplishment". We will discuss the details of this factor structure below (see Analysis 2). Factor 3 refers to extrinsic regulations for external pressures. EMEX1, 4 and EMINTRO 4 reflects the similar features: *external pressures mainly caused by teachers*. As for AMOT3, it can be one of the components of this factor because the students who feel external pressures to do the given tasks are expected to be amotivated in the long run (Nakano, 2006).

### 3.1.2 Reliability Analysis

To validate the internal consistency, we conducted reliability analysis by *alpha if item deleted*. As a result, we found that there were no items which would dramatically improve overall Cronbach's  $\alpha$  if they were deleted. In other words, we verified the structures of four factors extracted in this analysis as stable for the data set.

### 3.1.3 Mean Factor Scores

We calculated the factor scores for four factors by regression method. Figure 2 shows a comparison among three types of CCDL classes.

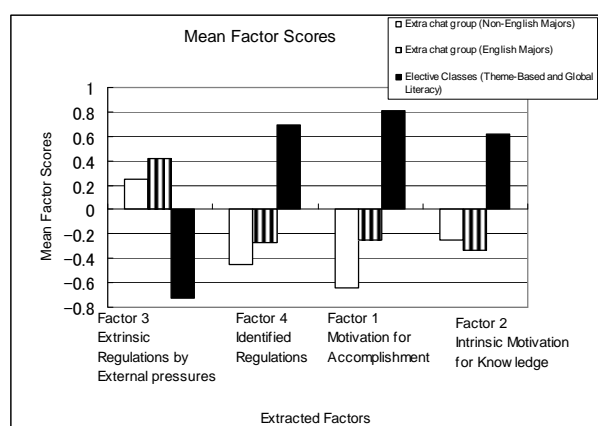


Figure 2: Mean Factor Scores

As Figure 2 indicates, the students in elective classes show high motivational propensities with regard to highly self-determined motivations and self-regulations. The degree of extrinsic regulations by external pressures is the lowest in the

comparison. The result implies that the students in elective classes (i.e., theme-based CCDL courses and Global Literacy course) opt to join CCDL activities autonomously, without feeling external pressures. Moreover, their mean factor score of motivation for accomplishment is obviously higher than the others.

On the other hand, both English majors and non-English majors in extra chat sessions show low scores on highly self-determined motivation, Factors 1, 2, and 4, whereas their scores of extrinsic regulations by external pressures are shown to be higher than those in elective classes. Because Factor 3 reflects teacher's pressures, it is suggested that these students are supposed to join extra chat sessions as a part of compulsory subjects. Especially, English majors tend to perceive their teacher's pressures more than non-English majors. On the other hand, non-English majors show the lowest level in comparison of Factor 3: motivation for accomplishment. The result would reflect the fact that non-English majors might not recognize the purpose of activities or specific relation between learning English in CCDL activities and their academic accomplishment, reflecting the fact that they are majoring in the a subject other than English.

Through the examination of motivational comparison, the students participated in elective classes showed higher motivational propensities whereas those in regular English classes showed low level of motivation as we assumed. However, it would be reasonable to further examine the factor structure obtained in this analysis because Factors 1 and 3 showed somewhat debatable aspects in term of theoretical relationship among types of motivation postulated by SDT. That is, the factor structure might result from the item reduction, following the recommendation by Field (2005) and Oshio (2004). Therefore we employed second analytic procedure to discuss if the factor structure observed in this analysis would be valid or not.

## 3.2 Analysis 2

### 3.2.1 Exploratory Factor Analysis

In this analysis, we specified the number of factors to be extracted as we predicted and discussed the remarkable factor structure observed in Analysis 1 by comparing each resulting factor model. We also aimed to discuss the methodological issues recommended by Field (2005) and Oshio (2004). In so doing, we put all the items into the analysis even though there were high correlations, ceiling effects or floor effects among 24 items. As for factor extraction technique, we employed principle factor analysis followed by promax rotation to

obtain the oblique solution.

To determine the number of factors to be extracted, we made two assumptions based on the intention we had in making the questionnaire and the findings in this study. First assumption was that because we set seven subscales in the questionnaire, seven-factor model would be suitable for the data. Second assumption was that the relation among componential items within Factor 1 we observed in Analysis 1 would be also observed in this analysis. In this case, because the factor included 6 items that were from different four subscales, the number of factors to be extracted would be less than seven factors. We tested this assumption by specifying the number as *6 factors*.

In testing the first assumption, we found that seven-factor model would not be suitable in terms of factor structure; although the first factor indicated almost same factor structure with Factor 1 we observed in Analysis 1. It implies that the relations among these items might reflect the reality of motivational propensities among CCDL participants. In order to closely examine the specific relations, we tested second assumption by specifying the number of factors to be extracted as *6 factors*.

In first factor analysis for six-factor model, KMO Sampling Adequacy was 0.910. Bartlett's Test of sphericity yielded statistical significance. All the items had adequate values in KMO statistics for individual items, ranging from 0.719 to 0.966. Because item 15 did not have appreciable factor loadings onto any extracted factor, we excluded it from second analysis. In second analysis, KMO Sampling Adequacy was 0.910. Bartlett's Test of sphericity yielded statistical significance. All the items had adequate values in KMO statistics for individual items, ranging from 0.711 to 0.966.

As a result, the resulting factor structure seemed to be suitable although items 14 and 23 cross-loaded onto two factors. To examine the reliability of this factor solution, we run reliability analysis with calculating *alpha if item deleted*. As a result, we found that deleting items 8, 14 and 20 would increase the overall alpha of Factors 4, 6, and 5 respectively. Therefore we excluded them and reanalyzed the data.

In third analysis, KMO Sampling Adequacy was 0.917. Bartlett's Test of sphericity yielded statistical significance. All the items had adequate values in KMO statistics for individual items, ranging from 0.754 to 0.966. The pattern matrix obtained from this analysis was shown in Table 2 (see next page). Though Factor 5 had only two componential items, it would be a suitable solution for the data set because the factor structure improved more than second analysis. To examine

the suitability, reliability analysis with calculating *alpha if item deleted* was done for each extracted factor.

Table 2: The Pattern Matrix (6 factor model)

Label	item	Factor					
		1	2	3	4	5	6
IMA2	item11	<b>0.83</b>	0.10	0.04	0.00	0.03	-0.06
EMID4	item2	<b>0.75</b>	-0.01	0.09	-0.04	-0.08	0.17
IMA3	item1	<b>0.74</b>	0.18	0.01	-0.12	0.00	-0.14
EMEX2	item10	<b>0.72</b>	-0.23	-0.14	0.19	-0.05	0.25
EM INTRO1	item24	<b>0.68</b>	-0.06	0.04	0.06	0.28	-0.21
IMS3	item3	<b>0.41</b>	0.21	-0.03	-0.04	-0.04	0.32
IMK1	item5	0.02	<b>0.90</b>	-0.01	0.02	-0.14	0.00
IMK2	item9	-0.03	<b>0.79</b>	-0.02	0.02	0.11	0.02
IMK3	item22	-0.06	<b>0.75</b>	-0.01	0.05	0.05	0.09
AMOT1	item13	-0.05	-0.09	<b>0.95</b>	-0.04	0.05	0.10
AMOT3	item12	0.16	-0.05	<b>0.83</b>	0.02	-0.15	0.11
AMOT2	item21	-0.06	0.12	<b>0.69</b>	-0.03	0.07	-0.22
EMEX1	item23	-0.21	0.03	<b>0.38</b>	<b>0.36</b>	0.01	-0.02
EM INTRO4	item7	0.06	0.01	-0.03	<b>0.92</b>	0.04	0.10
EMEX4	item4	0.06	0.07	0.03	<b>0.85</b>	-0.07	-0.13
IMS2	item18	0.05	0.05	-0.01	-0.01	<b>0.73</b>	0.17
IMS1	item16	0.26	0.05	-0.04	0.00	<b>0.63</b>	0.08
EMID1	item19	0.13	0.07	0.09	0.03	0.19	<b>0.69</b>
EMID3	item6	0.15	0.24	0.00	-0.08	-0.04	<b>0.59</b>
EMID2	item17	0.00	0.31	-0.05	0.01	0.10	<b>0.57</b>
% of Variance		48.77	9.13	5.16	3.48	2.37	1.63
eigenvalues		10.02	2.14	1.38	0.98	0.74	0.63
correlation		1	0.6	-0.49	-0.43	0.6	0.73
			1	-0.48	-0.47	0.63	0.67
				1	0.64	-0.37	-0.58
					1	-0.36	-0.58
						1	0.57

*Note:* The number of factors to be extracted was specified as 4 factors. Each item label stands for the abbreviations of subcategories.

As a result, there were no items which would dramatically improve overall Cronbach's  $\alpha$  if they were deleted. In other words, we verified the structure of six factors extracted in this analysis as stable for the data set. As for item 23, we regarded it as a component of Factors 3 and 4.

On the basis of componential items, we renamed each factor. Factor 1 was found to be similar to Factor 1 we extracted in Analysis 1 that was renamed 'Motivation for accomplishment'. 5 items in this analysis were in correspondence with those in Analysis 1. Besides, in this analysis, item 2, asking if the students internalize the goal of the class with regard to future works, loaded on the factor. That is, two items (items 2 and 10) related to the students' future directions are associated with

intrinsic motivation for accomplishment. This finding might imply the students' needs to acquire English ability in CCDL activities so as to make use of it in their future. In this sense, the motivation is considered to be *instrumental* (Deci and Ryan 1985; Gardner, 1985; Ryan and Deci, 2000 and 2002) per se. That is, this factor represents not only innate tendencies for pursuing accomplishment but also instrumental aspects subsumed in the students' highly self-determined predispositions toward CCDL activities. Therefore we labeled the factor as representing 'motivation for accomplishment'.

Factors 2, 5 and 6 were extracted as we intended. Accordingly, we renamed intrinsic motivation for knowledge, intrinsic motivation for stimulation and identified regulations respectively. Factor 3 was renamed as amotivation although it included an item in EMEX subscale. The relation between extrinsic regulations and amotivation was expected by Nakano (2006) as mentioned earlier. Factor 4 was regarded as extrinsic regulations for external pressures. All the items reflect external pressures mainly caused by teachers as we discussed in Analysis 1. We found the same 3 items clustered within one factor.

### 3.2.2 Mean Factor Scores

We calculated factor scores for six factors by regression method. Figure 3 shows a comparison among three groups. The extracted factors were ordered along the self-determination continuum postulated by SDT.

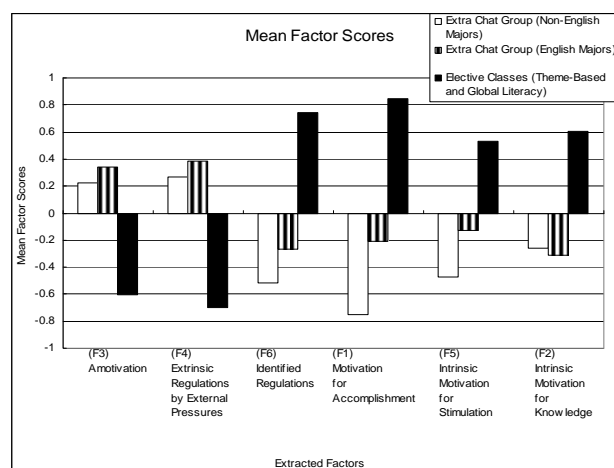


Figure 3: Mean Factor Scores

As shown, the students in elective CCDL classes show the high mean scores on types of autonomous motivations (Factors 1, 2, 5 and 6) while their mean factor scores on types of controlled motivations (Factors 3 and 4) are the lowest. That is, they are highly motivated to join CCDL activities. In particular, as in Analysis 1, the result indicates that the students in elective

CCDL classes would pursue accomplishment in the class activities. Furthermore, because the specific relation between accomplishment in the class and future contingencies was verified in Analysis 2, it might be said that the students in elective CCDL courses are highly motivated to pursue sense of accomplishment in CCDL activities for their futures. The causality is implied in their high mean score on both intrinsic motivation for stimulation that represents the students' innate interest in learning English in CCDL classes, and identified regulations that represent how the students internalize the goals and purposes of activities. These motivational aspects should be investigated in future research.

On the other hand, both English majors and non-English majors in extra chat sessions show high mean scores on types of controlled motivation whereas their scores on types of autonomous motivations are lower than average. It implies that the students in extra chat session are supposed to join CCDL activities as a part of compulsory subject. Comparing the students in extra chat sessions, English majors show the higher scores on Amotivation and Extrinsic Regulations for External Pressures than non-English majors. The result would reflect the controlling aspects of the class design. That is, the students are demotivated by external pressures that force them to participate in the activities. This aspect should be investigated with regard to three psychological needs for competence, autonomy and relatedness posited SDT (Ryan and Deci, 2002) to develop more suitable environment to enhance the students' internalization of non-self-determined regulations.

As for non-English majors, they show lowest mean scores on Identified Regulations, Motivation for Accomplishment and Intrinsic Motivation for Stimulation. Considering the fact that these students are majoring in a subject other than English, the result would be reasonable. In other words, the students would not internalize the importance as well as the values of CCDL activities. It would be because that they do not find specific relation between CCDL activities and their academic development and accomplishment.

#### 4 Summary

The remarkable finding both in Analysis 1 and 2 was that learner's motivation for pursuing accomplishment was found to be associated with the tangible rewards (i.e. future jobs) and their partners (i.e., partner in the class). It implied that the students' motivation for accomplishment included some instrumental aspect. Another finding was that one of the self-regulatory types reflects the extrinsic pressures caused by teachers as observed both in Analysis 1 and 2. That is, this

factor delineated the controlling aspect of compulsory English classes.

Through the examination of factor structure and motivational differences among three CCDL classes we investigated, we verified that the findings in this study support the notion that learner's motivation would be less if the class was compulsory while it would be more desirable if one chose the class based on his/her decision (Nakano, 2006).

Lastly, through Analyses 1 and 2, we found that high correlations ( $r > 0.8$ ) among items in questionnaire would not cause a serious problems but show a suitable solution for the data set if they occurred in same subscale. Rather, items that had relatively low correlations with others would be inaccurate to be included in the analysis.

#### References

- Deci, E.L., Ryan, R.M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. New York, Plenum Press.
- Field, A., (2005). *Discovering Statistics using SPSS*. 2nd Edition, Sage.
- Nakano, M., and CCDL Research Center. (2006). *CCDL Teacher's manual, Vol.1*. Waseda University.
- Nakano, M. and Yoshida, S. (2008). A Pilot Study: Exploring a Relationship between Four Kinds of Motivation and Self-regulations for Second Language Learning among Japanese University Students in Cross-Cultural Distance Learning Contexts. *Proceedings of the 13<sup>th</sup> Conference of PAAL*, 91-94.
- Nakano, M., Yoshida, S. and Owada, K. (2008). Cross-Cultural Distance Learning Programs and Program Assessment. *Proceedings of 9<sup>th</sup> APRU Distance Learning and the Internet Conference 2008*, 189-198.
- Noels, K.A., Pelletier, L.G., Clement, R., and Vallerand, R.J., (2000). Why Are You Learning a Second Language? Motivational Orientations and Self-Determination Theory. *Language Learning*, 50, 57-85.
- Park, H. (2006). Motivation in a foreign language learning context: The effect of gender issues in a Korean social and educational context. Unpublished doctoral dissertation, Korea University.
- Ryan, R. M., and Deci, E. L. (2002). Overview of self-determination theory: An organismic dialectical perspectives. In E.L. Deci and R.M. Ryan (Eds.), *Handbook of self-determination research* (pp.3-33). Rochester, NY: University of Rochester Press.

# Assessing Japanese EFL Learner's Social Skills in Cross-Cultural Distance Learning (CCDL) Context: a Longitudinal Study among Waseda University Students

Satoshi Yoshida and Michiko Nakano

Waseda University

satoshi-tko.jpn@akane.waseda.jp, nakanom@waseda.jp

## Abstract

The main purpose of this study is to describe the internal structure of social skills which would be required in cross-cultural communication, employing a psychological scale called Students Skillstreaming Checklist (Goldstein et al., 1997). Also, this study aims to compare the perceived frequency of skill use among CCDL participants so as to examine the specific relationship between each skill and CCDL activities. In doing so, first, we referred to I-T correlations observed in our two pilot studies and selected the relevant skill items for deeper investigation. In this process, 34 skill items were selected as relevant to CCDL activities. In order to verify the relevance, we distributed this 34-item questionnaire among 200 Waseda University students. The data were factor-analyzed so as to explore the latent variables among them. The result showed that the extracted 4 factors were closely related to CCDL activities as well as a series of skills required in cross-cultural communications. The result also indicated that the resulting factor structures in this study were almost the same as those in our pilot studies.

## Keywords

Social Skills, Theme-based CCDL courses, Cross-Cultural Distance Learning context

## 1 Introduction

### 1.1 CCDL activities and Social Skills

CCDL is one of the most adequate and currently challenging methods of foreign language teaching introduced by Waseda University among Asian partner universities in 1999. The program was renewed in 2008 and new classes called theme-based CCDL courses were officially offered by Open Education Center, Waseda University (Nakano, Yoshida, and Owada, 2008). Although theme-based CCDL courses can be classified into three types according to the main theme that are to

be dealt with in the class-*Global & Social Issues*, *Media Issues and International Career Path Issues*, three courses have been implemented by the common educational purpose (see Nakano, Yoshida and Owada, 2008). That is, in theme-based CCDL courses, a series of skills necessary for cross-cultural communications such as *Facilitation Skills*, *Emotional Intelligence*, *Social Intelligence and Researching Skill*, has been taught. Because these skills seem to be a series of required skills to maintain a good relationship with the interlocutors in cross-cultural communications, we regarded these skills as 'Social Skills' and conducted evaluative studies among Waseda University students (Yoshida and Nakano, 2008; Nakano, Yoshida and Owada, 2008) as pilot studies for the present study.

### 1.2 Our Pilot Studies

In our pilot studies, Yoshida and Nakano (2008) and, Nakano, Yoshida and Owada (2008) examined the applicability of a psychological scale called Student Skillstreaming checklist (Goldstein et al., 1997), which contains 50 items categorized into 6 subscale: Basic (beginning) Social Skills (8 items), Advanced Social Skills (6 items), Skills for Dealing with Feelings (7 items), Skills Alternatives to Aggression (9 items), Skills for Dealing with Stress (12 items), and Planning Skills (8 items).

We followed the statistical procedures used in Kikuchi (1988 and 2007), who optimized Goldstein et al.'s questionnaire and reported the remaining 18 items as Kikuchi's Scale of Social Skills-18 (KiSS-18). In his optimizing procedure, Kikuchi utilized Item-Total correlation analysis (I-T correlation analysis) and reduced the total number of items from 50 to 18, by selecting each three item (i.e., highly correlated items) from 6 subcategories. His procedures seemed appropriate because, in our pilot study, we found that it was impossible to factor-analyze 50 items.

We also adopted Kikuchi's definition of social

skills, i.e., “the skills to convey interpersonal relationship smoothly (Kikuchi, 1988: p.187)” for our discussions, and found that remaining 18 items in our study can be regarded as important to the CCDL participants. In this process, we verified the applicability of Goldstein et al.’s scale for our research purpose.

In pilot studies, however, we confronted a methodological issue in the process of optimizing the questionnaire. That is, some skill items considered to be closely related to CCDL activities, in terms of I-T correlation coefficient, were not included in our remaining 18 items. In his optimizing process, Kikuchi (1988 and 2007) also reported this sort of problem. Moreover, Kikuchi did not mention the logical or statistical reasons for remaining 18 items from the original 50 items. Therefore, it would be necessary for us to reconsider the optimizing procedures offered by Kikuchi (1988) for deeper investigation.

In this study, therefore, we review I-T correlation coefficients observed in our pilot studies 1 and 2 (Yoshida and Nakano, 2008; Nakano, Yoshida and Owada, 2008) and select the relevant skill items once again. Then we will discuss the significance of those selected items by conducting a following survey among CCDL participants. In this process, we will also compare the participants’ skill use, focusing on the types of CCDL classes.

## 2 Selecting Skill Items

Table 1 summarizes I-T correlation coefficients observed in our pilot study 1 (Yoshida and Nakano, 2008) and pilot study 2 (Nakano, Yoshida and Owada, 2008).

Table 1: I-T correlation coefficient in our pilot studies

Label	item	Pilot 2	Pilot 1
Basic	item1	0.50	0.48
Basic	item2	0.48	0.54
Basic	item3	0.55	0.49
Basic	item4	0.45	0.47
Basic	item5	0.47	0.46
Basic	item6	0.54	0.61
Basic	item7	0.51	0.59
Basic	item8	0.52	0.49
Advanced	item9	0.36	0.22
Advanced	item10	0.50	0.43
Advanced	item11	0.50	0.48
Advanced	item12	0.40	0.32
Advanced	item13	0.45	0.44
Advanced	item14	0.55	0.56

Label	item	Pilot 2	Pilot 1
Feeling	item15	0.41	0.35
Feeling	item16	0.54	0.46
Feeling	item17	0.52	0.50
Feeling	item18	0.48	0.42
Feeling	item19	0.36	0.41
Feeling	item20	0.55	0.48
Feeling	item21	0.33	0.24
Aggression	item22	0.56	0.47
Aggression	item23	0.37	0.36
Aggression	item24	0.53	0.33
Aggression	item25	0.55	0.48
Aggression	item26	0.43	0.40
Aggression	item27	0.39	0.38
Aggression	item28	0.47	0.55
Aggression	item29	0.22	0.24
Aggression	item30	0.36	0.35
Stress	item31	0.40	0.35
Stress	item32	0.59	0.56
Stress	item33	0.41	0.38
Stress	item34	0.47	0.50
Stress	item35	0.34	0.37
Stress	item36	0.46	0.36
Stress	item37	0.41	0.41
Stress	item38	0.50	0.34
Stress	item39	0.55	0.45
Stress	item40	0.53	0.43
Stress	item41	0.45	0.41
Stress	item42	0.42	0.42
Planning	item43	0.56	0.55
Planning	item44	0.55	0.52
Planning	item45	0.57	0.61
Planning	item46	0.44	0.47
Planning	item47	0.55	0.51
Planning	item48	0.47	0.48
Planning	item49	0.48	0.43
Planning	item50	0.46	0.55

*Note:* Each label stands for subcategories in Goldstein et al. (1997). The items with highlighted values in the columns, Pilot 1 and Pilot 2 are those retained as 18 items in each pilot study.

In selecting skill items, we set criteria as follows: (1) I-T correlation in the pilot study 2 should be higher than or almost equal to 0.45; and, (2) I-T correlation in the pilot study 1 should be higher than or almost equal to 0.4. On the basis of the criteria, we decided to retain 34 items as relevant to CCDL activities, in which items 24 and

36 were tentatively retained as potential items (i.e.,  $r < 0.4$  in pilot study 1). In order to verify the relevance between these skill items and CCDL activities, we distributed the 34-item questionnaire to Waseda University students and conducted a following survey (see next section).

### 3 Method

#### 3.1 Class Types and Participants

We divided the participants into three groups according to the types of CCDL class so as to make it easy to compare the skill use. They were (1) theme-based CCDL courses group<sup>1</sup>, (2) extra chat CCDL group and (3) cyber session CCDL group. The features of each class can be summarized as in following three sections.

##### 3.1.1 Theme-based CCDL

Theme-based CCDL courses were designed to be specializing in cross-cultural communication through CMC chat activities performed in regular class hours. These classes emphasize the importance of various kinds of skills required in cross-cultural communications, as mentioned earlier. Because the participating students are supposed to acquire these skills in the class, we expect them to frequently utilize various kinds of skills.

##### 3.1.2 Extra chat group

Extra chat group included all the regular English classes in which extra curricular chat sessions were adopted as a part of the class activities. In these classes, a series of skills emphasized in theme-based CCDL courses were not taught. Therefore we can regard this group as non-theme based.

##### 3.1.3 Cyber Session CCDL group

The features of these classes were different from the other two types of CCDL classes, because these classes were designed to be specializing in cross cultural communication via *video-conferencing system*. In their cyber sessions, moreover, the participating students were expected to be a moderator of their discussions, so that they may or may not acquire the required skills in the class. We regarded these classes as non-theme based and differentiated from theme-based CCDL courses.

#### 3.2 Participants

The participants in this study were 200 Waseda

University students who were all enrolled in CCDL classes. 70 out of 200 were in theme-based CCDL courses group; 101 in extra chat group and 29 in cyber session CCDL group.

#### 3.3 Questionnaire

The questionnaire used in this study was the brief questionnaire we selected based on Goldstein et al.'s (1997) scale (see, section 2). It contained 34 items. We employed both printed forms and online questionnaires in this survey.

#### 3.4 Procedure

The data collections were conducted during the fall semester of 2008. The students participated in this survey were asked to answer all the questions based on the situations in which they would use English. We told that their answers would never affect their course grade, and that their private information would never be revealed. These instructions were given before they answered the questionnaire. And they were also printed on the first page of the questionnaire.

The online questionnaire was available during October, 2008. The students could answer the questionnaire anytime they would like to do. The instructions about the questionnaire were given on the first page of the web-site.

#### 3.5 Analytic Procedure

We run exploratory factor analysis on 34 items. Though we failed to obtain the desirable solution in the analysis on overall 50 items in Pilot Study1, at this stage it would be reasonable to examine the latent variables among the 34 items that were carefully selected as the skills relevant to CCDL context. At the same time, I-T correlation analysis was done in order to verify the suitability of those items. Then, the items which had enough correlation (i.e.,  $r > 0.4$ ) were factor-analyzed. Also, to examine the relationship between each skill and CCDL context, we compared the skill use among the three types of CCDL classes by mean factor scores.

#### 3.6 Research Hypothesis

(1) Because the 34 items were selected as relevant to CCDL activities, the considerable I-T correlation (i.e.,  $r > 0.4$ ) will be a piece of evidence for its importance.

(2) Mean factor scores in theme-based CCDL courses (group) and cyber session classes (group) will be greater than those of extra chat group because the former two groups are designed to be specializing in cross-cultural distance learning.

<sup>1</sup> Because the number of students in Global Literacy course was 4, we regarded them as the members of theme-based CCDL courses.

## 4 Results and Discussions

### 4.1 I-T Correlation Analysis

We calculated I-T correlations of remaining 34 items. Table 2 shows the result of I-T correlation analysis.

Table 2: Result of I-T correlation Analysis

Label	Item	I-T Correlation Coefficient
Basic	item1	0.604
Basic	item2	0.645
Basic	item3	0.620
Basic	item4	0.552
Basic	item5	0.620
Basic	item6	0.607
Basic	item7	0.634
Basic	item8	0.576
Advanced	item10	0.561
Advanced	item11	0.703
Advanced	item13	0.664
Advanced	item14	0.682
Feeling	item16	0.673
Feeling	item17	0.696
Feeling	item18	0.639
Feeling	item20	0.644
Aggression	item22	0.713
Aggression	item24	0.693
Aggression	item25	0.693
Aggression	item28	0.618
Stress	item32	0.712
Stress	item34	0.540
Stress	item36	0.648
Stress	item38	0.637
Stress	item39	0.639
Stress	item40	0.701
Stress	item41	0.556
Planning	item43	0.683
Planning	item44	0.620
Planning	item45	0.582
Planning	item47	0.653
Planning	item48	0.568
Planning	item49	0.599
Planning	item50	0.508

*Note:* Each label stands for each subcategory in Goldstein et al. (1997). Each item number refers to the number in the original questionnaire. (N=200)

As shown in Table 2, all the items had appropriate correlation with the total score in the questionnaire ( $r > 0.4$ ). That is, all the skills measured by those 34 items could be regarded as important for the participating students. Therefore we put all the 34 items in factor analysis.

### 4.2 Factor Analysis

In factor analysis, we employed principle factor analysis, followed by promax rotation to obtain the oblique solution. All the procedures we followed were the same as in our pilot studies 1 and 2 (for details, see Field, 2005; or Yoshida and Nakano, 2008). Prior to the analysis, we checked all the variables and found that items 5 and 11 had ceiling effects. However, we tentatively included these items in the analysis. It was because that item 5, 'skill for saying thank you for others', and item 11, 'skill for apologizing to others', would be frequently used by Japanese EFL learners. Therefore it would be valuable to see the relationships they would hold with other variables.

In the correlation matrix for 34 items, there were no items which had high correlations or no correlations with others. Therefore, we continued the analysis without eliminating any item at this stage.

In factor analysis, we continued to run the analysis until the simplest pattern was obtained. We employed Kaiser's criterion to decide the number of factors to be extracted. In first analysis, KMO Measure of Sampling Adequacy was 0.943. Bartlett's Test of sphericity yielded statistical significance. All the KMO statistics for individual items were greater than 0.50. Therefore we continued to run the analysis and excluded the variable which did not have enough factor loadings or which cross-loaded onto the extracted factors in each analysis. In this process, items 1, 17, 18, 22, 24, 25, 36 and 39 were excluded in each analysis because of their low factor loadings; items 3, 14, 44 and 50 were excluded because of their cross-loadings. In final analysis, KMO Measure of Sampling Adequacy was 0.925. Bartlett's Test of sphericity yielded statistical significance. All the KMO statistics for individual items were greater than 0.50, ranging from 0.873 to 0.957. As a result, four factors were extracted. Table 3 shows the pattern matrix obtained from the final analysis (see next page). We renamed each extracted factor according to the componential items. The parenthesized number indicates the Cronbach's  $\alpha$ .

**Factor1: 'Basic Skills for L2 Communication'**  
(7 items,  $\alpha=0.885$ )

**Factor2: 'Planning Skills'** (6 items,  $\alpha=0.845$ )

**Factor3: 'Strategic Skills for Keeping Good Relationship with Others'** (4 items,  $\alpha=0.799$ )

**Factor4: 'Skills for Dealing with Affective Issues'**  
(5 items,  $\alpha=0.839$ )

Table 3: The pattern matrix of final factor analysis

Label	Item	Factor			
		1	2	3	4
Basic	item2	<b>0.81</b>	0.02	-0.05	0.00
Feeling	item16	<b>0.73</b>	0.11	-0.08	0.04
Basic	item6	<b>0.66</b>	-0.17	0.34	-0.04
Advanced	item10	<b>0.64</b>	-0.02	-0.02	0.11
Basic	item4	<b>0.63</b>	0.11	0.15	-0.20
Basic	item7	<b>0.59</b>	-0.08	0.13	0.15
Advanced	item11	<b>0.46</b>	0.19	-0.09	0.30
Planning	item47	0.06	<b>0.75</b>	0.13	-0.17
Planning	item43	0.18	<b>0.73</b>	-0.29	0.16
Stress	item41	0.03	<b>0.70</b>	-0.05	-0.02
Planning	item48	-0.13	<b>0.67</b>	0.21	-0.05
Planning	item45	0.07	<b>0.62</b>	0.03	-0.03
Stress	item38	-0.04	<b>0.41</b>	0.27	0.12
Basic	item5	0.12	0.00	<b>0.79</b>	-0.12
Advanced	item13	0.00	-0.01	<b>0.65</b>	0.22
Planning	item49	0.03	0.15	<b>0.53</b>	0.05
Basic	item8	0.33	0.07	<b>0.36</b>	-0.05
Aggression	item28	0.11	-0.10	-0.13	<b>0.88</b>
Stress	item32	0.11	-0.10	0.10	<b>0.77</b>
Stress	item34	-0.19	0.23	0.13	<b>0.47</b>
Feeling	item20	-0.09	0.26	0.19	<b>0.42</b>
Stress	item40	-0.02	0.25	0.24	<b>0.39</b>
% of variance		41.44	5.88	3.77	3.02
eigenvalue		9.57	1.76	1.24	1.09
Correlation		1	0.57	0.54	0.63
			1	0.62	0.69
				1	0.57
					1

Note: Each item label stands for the subcategories in Goldstein et al. (1997). Each item number refers to those in the original questionnaire.

Factor 1 stands for ‘Beginning Skills for L2 communication’. The items contained in this factor are considered to be fundamental in L2 communication. It is remarkable that items 2, 6, 7 and 10 are in correspondence with those in the factor renamed as ‘Basic Social Skills for L2 communication’ in our pilot study 1. As compared with the factor in Pilot Study 2 named ‘Fundamental Skills for L2 Communication’, 4 items (items 6, 10, 11 and 16) were in correspondence with those in the present experiment. Therefore it can be said that the skills contained in this factor are essential for L2 communication in CCDL context.

Factor 2 indicates ‘Planning Skills’. We also found this sort of factor continuously throughout our pilot studies. Especially, items 43 and 45 were found to be important components of this factor. Considering the componential items in Factor 2, it

obviously indicates the skills for preparing for achieving any task. These skills were necessarily required in cross cultural communication in CCDL context. Interestingly, item 38 defined as a skill for “responding to failure (Goldstein et al., 1997)” is also loaded on this factor. It implies that the students would learn from mistakes in the past and they try not repeat similar mistakes. This finding would reflect the fact that the students begin, through trials and errors, to attain good learners’ well-prepared attitudes for cross-cultural communication tasks.

Factor 3 represents ‘Strategic Skills for Keeping Good Relationship with Others’. Though the componential items, items 5 and 13 were found to have ceiling effect, they made up the comprehensible factor in the analysis. Because skills for “saying thank you to others” and “apologizing others” are expected to be frequently used by Japanese EFL learners as mentioned earlier, the factor consisting of these skills should be regarded as essential in L2 communication. It is because that these skills are considered to be one of the communication strategies to maintain good relationships with other people. Regarding item 8, skill for ‘giving a complement to others’ is conceptually relevant skill to the former two skills, so that it should be also regarded as one of the strategies for keeping good relationships with others. That is, this factor is concerned with L2 communication strategies. Item 49 was to measure the skill for “making a decision (Goldstein et al., 1997).” In the light of the relation of this skill with other three skills, it would be possible to interpret this skill as that for making a best decision by grasping the situation surrounding the person in question. That is, this factor might imply that the students adjust their communication strategies, depending on the situation to situation.

Factor 4 stands for ‘Skills for Affective Issues’. As all the items in this factor indicate, the skills represented by this factor are required to deal with problems that people would face in their inter-personal communication. Referring to Goldstein et al.’s (1997) classification, this factor consists of items belonging to Skills for Dealing with Feeling, Aggression and Stress. Thus, the skill covers wide area of human’s affective issues. Because it is expected that students in CCDL activities would face the problems such as misunderstandings, disagreement of opinion and accusation in cross-cultural communication, these skills are definitely important. That is, this factor reflects the necessity to understand one’s own feeling in order to facilitate the cross-cultural communication. This finding might suggest that the notion of Emotional Intelligence that has been

emphasized in theme-based CCDL courses conforms to CCDL context.

To sum up, all the factors extracted in the present study were interpretable with regard to the factor structure and the specific relationships between each item and CCDL context. To verify the specific relationships, we compare skill use among three types of CCDL classes.

### 4.3 Mean Factor Scores

The factor scores for each extracted factor were calculated by regression method. We compared frequency of skill use in terms of mean factor scores for each group. Figure 1 indicates the comparison among the three class types.

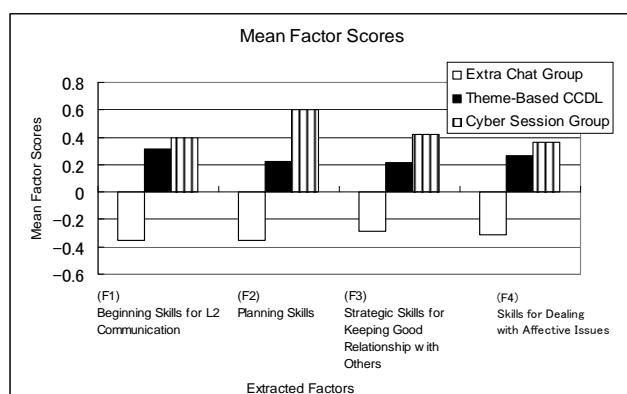


Figure 1: Mean factor scores

As Figure 1 indicates, the students who participated in Cyber session classes show the highest skill use in 4 factors. As for those in theme-based CCDL courses, their mean scores for 4 factors are shown to be relatively lower but almost similar patterns to Cyber session classes. On the other hand, the students in Extra chat group show the lowest skill use in 4 factors. This result implies that these 4 skills observed in the present study are relevant to CCDL context because two classes designed to specialize in CCDL activities show higher skill use in each comparison as we assumed. In this sense, the applicability of 34 questionnaire items selected in this study was verified in terms of specific relationships between each item and CCDL activities.

## 5 Summary

First, we verified the suitability of 34 items selected by I-T correlation analysis. As we hypothesized, all the items had appropriate correlation with the total score (i.e.,  $r > 0.4$ ). Drawing on Kikuchi's (1988 and 2007) explanation, these 34 items should be regarded as important skills to the participating students for facilitating cross-cultural communication. Therefore, research hypothesis 1 was verified.

We could see the internal structure of social skills that would be required in CCDL context. The four factors extracted in exploratory factor analysis were comprehensible with regard to the relationships between each item and CCDL activities. In comparison of skill use among three types of CCDL classes, it was demonstrated that the students in two classes designed to specialize in cross-cultural communication (i.e., theme-based CCDL courses and Cyber Session classes) would frequently utilize these four skills in CCDL activities. It implied that the skills represented by 4 factors were closely related to CCDL context. In this sense, research hypothesis 2 was verified.

On the basis of these findings, therefore, it can be said that the brief questionnaire we developed would be appreciable to assess and evaluate the skill use among CCDL classes. To validate the findings in this study, future research will be required.

### Acknowledgment

We would like to express our sincerest thanks to staff members of Waseda University International and CCDL Support Center, who helped us to collect the data for this study as well as those for our motivational study cited in this booklet. We also thank all the professors who allowed us to distribute our questionnaire in their class hours.

### References

- Field, A., (2005). *Discovering Statistics using SPSS*. 2nd Edition, Sage.
- Goldstein, A. P., McGinnis, E., Sparfkin, R. P., Gershaw, N. J. and Klein, P. (1997). *Skillstreaming the adolescent: new strategies and perspectives for teaching prosocial skills*. Revised Edition. Research Press.
- Kikuchi, A. (1988). *Omoiyari -wo- kagaku -suru*. Kawashima Shoten.
- Kikuchi, A., (2007). *Shakaiteki-Sukiru-wo-hakaru-KiSS-18 Handbook*. Kawashima Shoten.
- Nakano, M., Yoshida, S. and Owada, K. (2008). Cross-Cultural Distance Learning (CCDL) Programs and Program Assessment. *Proceedings of 9<sup>th</sup> APRU Distance Learning and the Internet Conference 2008*, 189-198.
- Yoshida, S. and Nakano, M. (2008). Assessing Social Skills among Japanese EFL Learners: A Pilot Study. *Proceedings of the 13<sup>th</sup> Conference of Pan-Pacific Association of Applied Linguistics*, 205-208.

## Lesson Review Tests and CEFR Can-do Statements

Michiko Nakano<sup>1</sup> Hikaru Sugiyama<sup>2</sup>, Manabu Itoh<sup>2</sup>, Yusuke Kondo<sup>3</sup> and Hajime Tsubaki<sup>4</sup>

<sup>1</sup>School of Education, Waseda University, <sup>2</sup>Waseda University International Co., <sup>3</sup>Ritsumeikan University,

<sup>4</sup>GITS, Waseda University

nakanom@waseda.jp, sugiyama@w-int.jp, yusukekondo@asagi.waseda.jp

### Abstract

English Tutorials give Lesson Review Tests (LRT) every seventh lesson, totaling three LRTs per semester. We have assigned a can-do statement for each test item. LRT has four sections: 30 vocabulary items, 18 idioms, 18 Listening comprehension and 18 partial dictations. English Tutorials have six proficiency levels which approximately correspond to A1, A2, B1, B2, C1 and C2 in the Common European Framework of Reference (CEFR). Thus, the total number of test items amounts to 540 vocabulary items, 324 idioms, 324 listening comprehension items and 324 partial dictations. However, these test items are not homogeneous, since the textbooks contain some mixed items; for instance, Beginners (A1) includes A1 and A2 items. This is because English at University Level requires these mixed items. This tendency can be observed at all the six levels of English Tutorials. In this presentation, we will clarify whether learners' performance indicate some differentiations along the ordinal scale, by running Exametrika program developed by Shojima (2009). Exametrika based on neural test theory was adopted here, since it can yield relevant can-do statements which are ranked along the six levels. According to Shojima (2008), neural test theory uses the mechanism of a self-organizing map. The item reference profile represents statistical characteristics of each test item. It assumes that the latent scale is ordinal and that each examinee is located on the latent rank scale. Thus, the framework fits our testing conditions highly well.

### Keywords

Neural Test Theory, Exametrika, Can-do statements, CEFR

### Introduction

The English Tutorial Lessons at Waseda University

have been offered by the Open Education Center since 1997, and they have been popular among students in the university. In the past four years, about 10,000 students took these Tutorial lessons per year. These courses aim to promote practical English skills so that the students can communicate functionally well in English. In order to achieve this objective, one tutor teaches a small group of four students. This small group training is effective to reduce students' speech anxiety in English and to provide social contexts of speech situations. Since the tutorial lessons create a context for socialization, it can promote acquisition of communicative competence and it is effective to let learners use their passive knowledge of vocabulary and grammar automatically and stably. The English Tutorials are designed based on the Common European Framework of References (CEFR) and there are six levels: beginners, basic, pre-intermediate, intermediate, pre-advanced and advanced. These levels roughly correspond to the six levels in the CEFR: A1, A2, B1, B2, C1 and C2 respectively. Each unit of the textbooks contains 2 can-do functions. Lesson Review Test examines a learner's mastery of each function in terms of vocabulary, idioms, listening comprehension and partial dictations. Each test item is mapped onto the list of CEFR descriptors. It is hoped that Lesson Review Tests can output the degree of mastery of can-do functions.

The CEFR was publicized in 2001 at the International Conference to Commemorate European Year of Languages held at the Free University, Germany, after a long-term empirical investigation in order to set up common standards in language teaching and learning. Since then, many books and articles about the CEFR have been published. It is now known as the framework providing the most comprehensive descriptors of language learning. The CEFR managed to define L2 proficiency in functional terms so that the same descriptors can be used to define learning goals, to develop learning materials and activities, and to judge learning achievement. At Waseda University,

Michiko Nakano, the Director of English Language Division, Open Education Center, introduced CEFR into the framework of language teaching at Open Education Center in 2001. Nakano and her associates have modified teaching materials and tasks every year so that all the language learning activities in English Tutorials follow the CEFR descriptors and guidelines. In 2006 we started to validate the linkage of the tutorials to the CEFR. The present report is one of our efforts in this direction.

## 1 Method

In order to examine the relationship between the Lesson Review Test scores and the CEFR can-do statements, we conducted test sessions for beginners, basic, pre-intermediate, intermediate, pre-advanced and advanced level course participants three times per semester. This paper examines the score results, autumn term, 2008. Since the data is massive, we present the result on the first session on Basic course for illustration here.

### 1.1 Materials

The test items are constructed on the basis of can-do functions in each unit. Section 1 provides 30 vocabulary items, Section 2, 18 idioms, Section 3, 18 Listen Comprehension items and Section 4, partial dictations. Two examples in Unit 1 are given below.

Basic 1-1 Can-do A1

"If you study linguistics or literature you are in the ( ).

- (1) School of Political Science and Economics
- (2) School of Letters Arts and Sciences
- (3) School of Human Sciences
- (4) School of Commerce"

1-2 Can-do A1

"If you study accounting you are in the ( ).

- (1) School of Political Science and Economics
- (2) School of Letters Arts and Sciences
- (3) School of Human Sciences
- (4) School of Commerce"

2-1 Can-do A1

"When you introduce yourself to people you don't know you would say ( ).

- (1) "Nice to meet you."
- (2) "What's new?"
- (3) "Wow!"
- (4) "I'm sorry, I don't follow you."

2-2 Can-do A1

"When you greet people you know you would say ( ).

- (1) "Nice to meet you."
- (2) "How have you been?"

(3) "Not bad."

(4) "What do you mean?"

3-1 Can-do A1

"Q. What year is Li in?"

- (1) freshman
- (2) sophomore
- (3) junior
- (4) senior

3-2 Can-do A1

"Q. What is Li studying?"

- (1) Law
- (2) Education
- (3) Mathematics
- (4) English

4-1 Can-do A1

"Kei : Hi nice to meet you. My name's Kei Yamamoto. Please call me Kei."

Li : ( ) ( ) ( )  
( ) ( ) Kei. I'm Li. I'm majoring in ( ).

4-2 Can-do A1

"Li: What are you studying?

Kei: I'm studying ( ). I'm a ( ).  
What year are you in?

Li: I'm a ( ).

## 1.2 Correspondence list of Can-do topics and CEFR descriptors

Appendix 1 gives a correspondence list of can-do functions and CEFR descriptors.

### 1.3 1<sup>st</sup> test session

The first test session took place at the 7<sup>th</sup> lesson and 135 students who attended the basic course volunteered to take a test

## 2 Result and Discussion

As Table1 indicate, Basic course contains a mixed level items form A1, A2 lower, A2 and A2 upper. For this reason we set the number of ranks as 4.

Table 1

A1l	A1	A1u	A2l	A2	A2u
.	1	-	.	.	.
.	1	-	.	.	.
.	1	-	.	.	.
.	1	-	.	.	.
.	1	-	.	.	.
.	.	-	.	4	2
.	.	-	.	4	2
.	.	-	.	1	1
.	.	-	.	1	1
.	.	-	.	1	1
.	.	-	1	.	2
.	.	-	.	.	2
.	.	-	1	.	2
.	.	-	.	.	2
.	.	-	1	.	.
.	.	-	.	.	1
.	.	-	1	.	.
.	.	-	.	.	1
.	.	-	1	.	.
.	.	-	.	.	1
.	.	-	1	.	.
.	.	-	1	.	.
.	.	-	1	.	.
.	.	-	1	.	.

Table 2 gives the summary of statistics.  
Table 2 Basic Unit 1-6 Test Summary

基本統計量				
受験者数(N)	135			
項目数(n)	84			
最小値(Mn)	39			
最大値(Max)	79			
中央値(Median)	69			
平均値(Mean)	68.185			
分散(Var.)	53.719			
標準偏差(SD)	7.329			
アルファ係数	0.816			
	Rank 1	Rank 2	Rank 3	Rank 4
テスト参照プロファイル	53.691	64.005	69.469	75.009
潜在ランク分布(LRD)	9	38	48	40
ランク・メンバーシップ分	10.152	37.470	45.999	41.379
相対IRP	0.639	0.762	0.827	0.893
相対LRD	0.067	0.281	0.356	0.296
相対RMD	0.075	0.278	0.341	0.307
テスト適合度 RMPIに基づくテスト適合度				
カイ2乗値	18.867	7.227		
有効自由度 effective df	6758.822	6758.822		
P値	1.000	1.000		
NFI	0.916	0.958		
RFI	0.916	0.958		
IFI	1.000	1.000		
TLI	1.000	1.000		
CFI	1.000	1.000		
RMSEA	0.000	0.000		
AIC	-13498.777	-13510.416		
CAIC	-39893.832	-39905.472		
BIC	-33135.011	-33146.650		

Table 3 gives a partial list of Item Reference Profiles. It indicates that Item 1-25 can not be ranked along the four point ordinal scale; the item does not have a discriminatory power. Item 1-9 shows a similar tendency. Item 1-30 cannot differentiate rank2 and rank3 examinees. This is endorsed by Item graphs shown in Fig 1 ~ Fig 4.

項目参照プロファイル(IRP)			
Rank 1	Rank 2	Rank 3	Rank 4
0.599	0.749	0.838	0.979
0.514	0.743	0.794	0.867
0.740	0.894	0.925	0.936
0.759	0.937	0.941	0.949
0.912	0.951	0.996	1.000
0.807	0.933	0.955	0.973
0.911	0.962	0.971	0.996
0.950	0.973	0.983	0.994
0.383	0.526	0.675	0.907
0.532	0.796	0.881	0.938
0.734	0.820	0.832	0.914
0.539	0.764	0.796	0.888
0.748	0.920	0.961	0.968
0.922	0.937	0.949	0.958
0.571	0.590	0.686	0.691
0.539	0.775	0.867	0.873
0.949	0.975	0.983	0.995
0.612	0.702	0.718	0.740
0.919	0.975	0.975	0.976
0.651	0.891	0.955	0.980
0.775	0.818	0.880	0.921
0.613	0.716	0.742	0.875
0.572	0.823	0.826	0.890
0.401	0.817	0.844	0.974
0.290	0.307	0.310	0.316
0.795	0.855	0.861	0.984
0.771	0.977	0.978	0.990
0.766	0.918	0.955	0.981
0.673	0.881	0.900	0.975
0.307	0.465	0.480	0.670

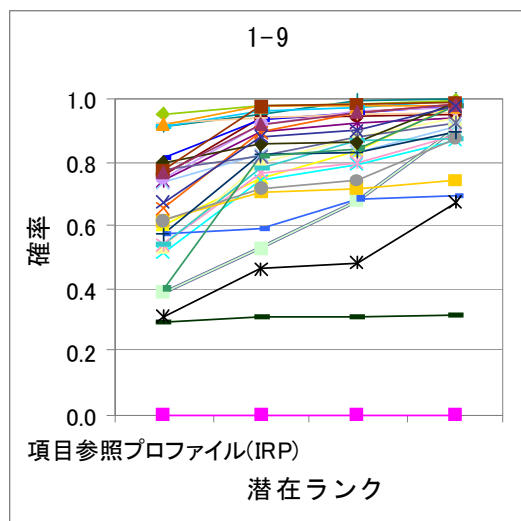


Fig 1 Item 1-9

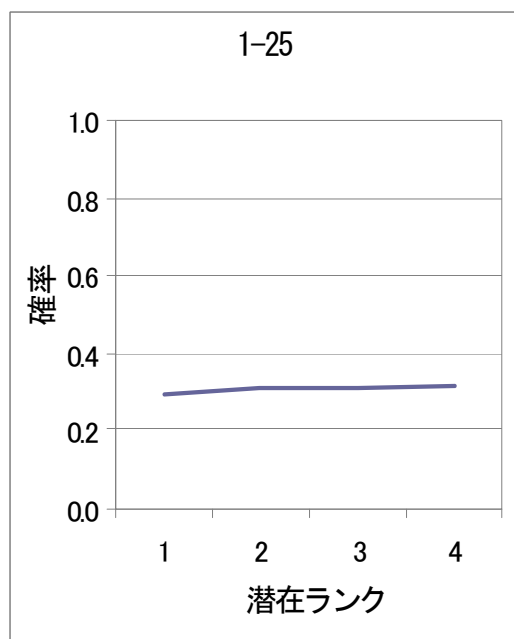


Fig 1-25

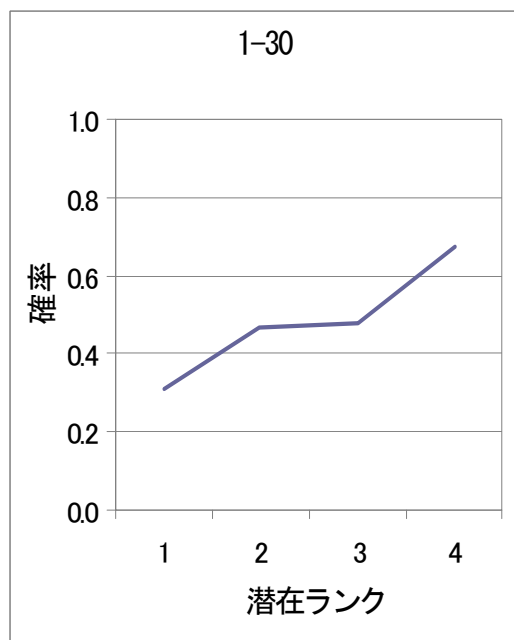


Fig 1-30

Table 4 gives a partial list of Examinees' Ranking Estimated. Since examinees' ranking are estimated, we can output the following can-do statements, depending on the rank estimated.

Feedback to Rank1 examinees:

- Can make an introduction and use basic greeting and leave-taking expressions.
- Can ask and answer simple questions, initiate and respond to simple statements in areas of immediate need or on very familiar topics.

Feedback to Rank2 examinees (their proficiency subsumes Rank 1 can-do mastery):

- Can communicate in simple and routine

tasks using simple phrases to ask for and provide things, to get simple information and discuss what to do next.

- Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters to do with work and free time.
- Can say what he/she likes and dislikes.
- Can describe people, places, and possessions in simple terms.
- Can describe people, places, and possessions in simple terms.

Feedback to Rank 3 examinees (their proficiency subsumes Rank 1 and Rank 2 can-do mastery):

- Can participate in short conversations in routine contexts on topics of interest.

Feedback to Rank 4 examinees (their proficiency subsumes Rank 1, Rank 2 and Rank3 can-do mastery):

- Can participate in short conversations in routine contexts on topics of interest.
- Can generally understand clear, standard speech on familiar matters directed at him/her, provided he/she can ask for repetition or reformation from time to time.
- Can describe plans and arrangements, habits and routines, past activities and personal experiences.
- Can discuss what to do in the evening, at the weekend.
- Can ask and answer questions about habits and routines.
- Can explain what he/she likes and dislikes.
- Can explain what he/she likes and dislikes.
- Can express how he/she feels in simple terms, and express thanks.
- Can describe plans and arrangements, habits and routines, past activities and personal experiences.
- Can ask and answer questions about pastimes and past activities.
- Can describe plans and arrangements, habits and routines, past activities and personal experiences.
- Can briefly reasons and explanations for opinions, plans and activities.

### 3.0 Summary

Exametrika developed by Shojima (2009) based on Neural Net Theory proves to be useful in four ways; (1) the examinees can be ranked appropriately; (2) it estimate the probability of correct responses, depending on their ranks; (3) by looking at Item Reference Profiles, we can

determine which item should be removed and altered; (4) in conjunction with the test items with can-do statements, one can device the appropriate feedbacks specific to their ranked proficiency.

Table A partial list of Examinees' Ranking Estimated

潜在ランク 推定値	Rank 1	Rank 2	Rank 3	Rank 4
2	0.003	0.982	0.015	0.000
2	0.118	0.868	0.014	0.000
3	0.000	0.210	0.733	0.057
3	0.000	0.090	0.909	0.000
2	0.089	0.911	0.000	0.000
4	0.000	0.001	0.066	0.933
3	0.001	0.191	0.806	0.002
2	0.001	0.866	0.133	0.000
4	0.000	0.000	0.022	0.978
2	0.003	0.995	0.002	0.000
2	0.005	0.969	0.026	0.000
4	0.000	0.012	0.064	0.924
4	0.000	0.001	0.005	0.994
2	0.000	0.758	0.242	0.000
2	0.241	0.757	0.002	0.000
4	0.000	0.000	0.006	0.994
2	0.002	0.903	0.095	0.000
2	0.000	0.537	0.297	0.165
2	0.000	0.888	0.104	0.008
2	0.002	0.818	0.169	0.011
2	0.005	0.993	0.002	0.000
3	0.000	0.448	0.546	0.007
2	0.004	0.994	0.002	0.000
1	1.000	0.000	0.000	0.000
4	0.000	0.001	0.017	0.983
3	0.001	0.402	0.596	0.001
4	0.000	0.000	0.032	0.968
3	0.000	0.033	0.961	0.006
4	0.000	0.000	0.045	0.955
2	0.016	0.950	0.034	0.000

### References

- Shojima, K. (2008). Neural Test Theory: Latent Rank Theory for Analyzing Test Data. In K. Shigematsu, A. Okada, T. Imaizumi, and T. Hoshino (eds.) *New Trends in Psychometrics*. Universal Academy Press Inc. 1-37.

Shojima, K. (2007). Latent Rank Theory:  
Estimation of Item Reference Profile by  
Marginal Maximum Likelihood Method  
with EM Algorithm. *DNC Research Note*.  
07-12.

## Appendix

Can Do List: Reach Out Basic			
Unit	Can-do	Level	Level in ELP
1	Introduce oneself and greet people	A1	<b>Conversation A1</b> Can make an introduction and use basic greeting and leave-taking expressions.
	Use conversation management English	A2 lower	<b>Goal-oriented co-operation A2 (lower)</b> Can communicate in simple and routine tasks using simple phrases to ask for and provide things, to get simple information and discuss what to do next.
	<b>Extension Activity</b>	A2 lower	<b>Sustained Monologue Describing Experience A2 (lower)</b> Can describe his/her family, living conditions, educational background, present or most recent job.
2	Be a good listener.	A1	<b>Overall Spoken Interaction A1</b> Can ask and answer simple questions, initiate and respond to simple statements in areas of immediate need or on very familiar topics.
	Continue conversations	A2 Upper	<b>Conversation A2 (upper)</b> Can participate in short conversations in routine contexts on topics of interest.
	<b>Extension Activity</b>	A2 Upper	<b>Conversation A2 (upper)</b> Can generally understand clear, standard speech on familiar matters directed at him/her, provided he/she can ask for repetition or reformation from time to time.
3	Talk about daily routines	A2 upper	<b>Sustained Monologue Describing Experience A2 (upper)</b> Can describe plans and arrangements, habits and routines, past activities and personal experiences.
		A2 upper	<b>Informal Discussion with Friends A2 (upper)</b> Can discuss what to do in the evening, at the weekend.
	Talk about free time activities	A1 lower	<b>Overall Spoken Interaction A1(lower)</b> Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters to do with work and free time.
	<b>Extension Activity</b>	A2 Upper	<b>Information Exchange A2 (upper)</b> Can ask and answer questions about habits and routines.
4	Express likes and dislikes	A2 lower	<b>Conversation A2 (lower)</b> Can say what he/she likes and dislikes.
	Explain likes and dislikes	A2 upper	<b>Sustained Monologue Describing Experience A2 (upper)</b> Can explain what he/she likes and dislikes.
	<b>Extension Activity</b>	A2 upper	<b>Sustained Monologue Describing Experience A2 (upper)</b> Can explain what he/she likes and dislikes.
5	Describe my feelings.	A2 upper	<b>Conversation A2 (upper)</b> Can express how he/she feels in simple terms, and express thanks.
	Discuss recent events	A2 upper	<b>Sustained Monologue Describing Experience A2 (upper)</b> Can describe plans and arrangements, habits and routines, past activities and personal experiences.
		A2 upper	<b>Information Exchange A2 (upper)</b> Can ask and answer questions about pastimes and past activities.
	<b>Extension Activity</b>	A2 upper	<b>Sustained Monologue Describing Experience A2 (upper)</b> Can describe plans and arrangements, habits and routines, past activities and personal experiences.
6	Describe and compare places	A2 lower	<b>Sustained Monologue Describing Experience A2 (lower)</b> Can describe people, places, and possessions in simple terms.
	Talk about the features of places	A2 lower	<b>Sustained Monologue Describing Experience A2 (lower)</b> Can describe people, places, and possessions in simple terms.
	<b>Extension Activity</b>	B1 lower	<b>Sustained Monologue Putting a Case B1 (lower)</b> Can briefly reasons and explanations for opinions, plans and activities.

## English Tutorials, CEFR and ACPA

Michiko Nakano<sup>1</sup> Kazuharu Owada<sup>2</sup> Eiichiro Tsutsui<sup>3</sup> and Yusuke Kondo<sup>4</sup>

<sup>1</sup>School of Education, Waseda University, <sup>2</sup>Tokyo College of Music, <sup>3</sup>International Hiroshima University and

<sup>4</sup>Ritsumeikan University

nakanom@waseda.jp, yusukekondo@asagi.waseda.jp

### Abstract

Accreditation Council for Practical Abilities (ACPA) has developed the Standard Skill Descriptions in the domains of Information Technology and Business Management since 2003. Our descriptions have been publicized and ACPA has been conducting the accreditation services for IT or Business courses and certification tests for practical abilities in these domains.

In 2008 ACPA developed the Standard Skill Descriptions in language learning (particularly in English), using the same framework of the above ACPA Standard Skill Descriptions based on Common European Framework of Reference for language (CEFR). This is due to the specific request from Open Education Center, Waseda University. At Open Education Center, English has been taught based on the CEFR since 2001. Over the years they have achieved the fairly cohesive syllabus by incorporating most of the requirements the CEFR specifies into their curriculum as a whole. Their efforts to improve curriculum and methods of teaching have been very impressive, although the development is gradual and still continuing. Since they managed to provide some experimental evidence for validation, we decided to consider their request seriously. In 2009 we hope to issue accreditation service for Open Education Center a part of CEFR-based curriculum, Tutorial English.

### Keywords

Accreditation, CEFR, English Tutorials

### Introduction

Accreditation Council for Practical Abilities (ACPA) has developed the Standard Skill Descriptions in the domains of Information Technology and Business Management since 2003. Our descriptions have been publicized and ACPA has been conducting the accreditation services for IT or Business courses and certification tests for practical

abilities in these domains.

In 2008 ACPA developed the Standard Skill Descriptions in language learning (particularly in English), using the same framework of the above ACPA Standard Skill Descriptions based on Common European Framework of Reference for language (CEFR). This is due to the specific request from Open Education Center, Waseda University. At Open Education Center, English has been taught based on the CEFR since 2001. Over the years they have achieved the fairly cohesive syllabus by incorporating most of the requirements the CEFR specifies into their curriculum as a whole. Their efforts to improve curriculum and methods of teaching have been very impressive, although the development is gradual and still continuing. Since they managed to provide some experimental evidence for validation, we decided to consider their request seriously. In 2009 we hope to issue accreditation service for Open Education Center a part of CEFR-based curriculum, Tutorial English.

The CEFR has been well recognized all over the world. It is hoped that ACPA can assist Japanese learners of English at University levels to improve practical abilities, based on the International Standards. ACPA Initiatives will be useful and helpful for Educational Institutes in Japan, Japanese industries and individual learners and eventually contribute to the development of practical language (English) skills which Japanese Businessmen will urgently require in order to act as competent global citizens.

This document explains the outline of the Standard Skills Descriptions in the domain of language (English).

### 1. Accreditation Council of Practical Abilities (ACPA)

#### 1.1 What is ACPA

The Accreditation Council for Practical Abilities (ACPA) is a non-profit organization established in

2003 with the support of the government, corporations and higher educational institutions. Its objective is to establish an online network system to assure educational quality in practical abilities and to issue certificates to individuals, and to conduct activities to nurture human resources needed by the present-day society.

## 1.2 ACPA Standard Skills Descriptions

The ACPA formulates standards in practical skills which are required in practical works in the form of standard skills descriptions. The ACPA certifies the courses and tests used in institutions, and issues certificates to individuals.

The ACPA Standard Skills Descriptions provides a skill standard matrix which depicts a relationship between a profession (or a job) and a specific set of practical skills required by a given profession or a job. The skill standard matrix indicates a relationship between occupational categories and skills in the form of a table showing which skills are required for each occupational category. It helps institutions and individuals to understand a required level of a set of skills. The IT domain Standard Skills Guide has been used for the certification service since 2006. The Business Management domain Standard Skills Guide has been used for the certification service since 2007.

## 2. Standard Skills Guide for English

### 2.1 Background of Standard Skills Guide for English

The global activities of business enterprises have rapidly spread in the present-day society. The importance of international standards of English skills have been recognized not only in Asia but also all over the world. Although many major enterprises in Asia adopt TOEIC test for evaluating English competence in business, it is not always true that high score in TOEIC test reflects actual content of practical abilities in daily business negotiations and interactions. The common European framework of reference (CEFR), on the other hand, describes actual content of practical abilities in language use. Using the matrix of ACPA practical ability descriptions which has been introduced in IT domain and Business Management domain, we have rearranged and revised the CEFR illustrative descriptors to suit English Language Education as a foreign language in Japan. It is hoped that this revision is appropriate in Asia as well.

The CEFR aims to provide a common basis for the elaboration of syllabuses, curriculum guidelines, examinations, teaching materials, textbooks, and teaching methods in language learning. It is widely used as authoritative international standards.

In developing the ACPA Standard Skills Guide, the CEFR illustrative descriptors are translated into Japanese, independently from the translation by S. Yoshijima, R. Ohashi, et. al. (2004). The proficiency levels of ACPA Standard Skills Guide adopt six levels, following the CEFR.

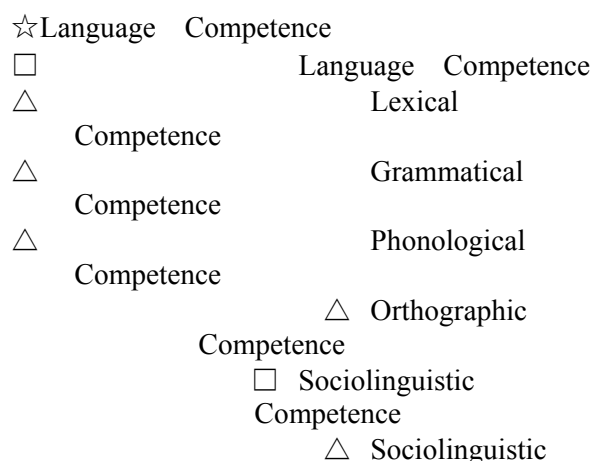
### 2.2 Purpose of Standard Skills Guide for English

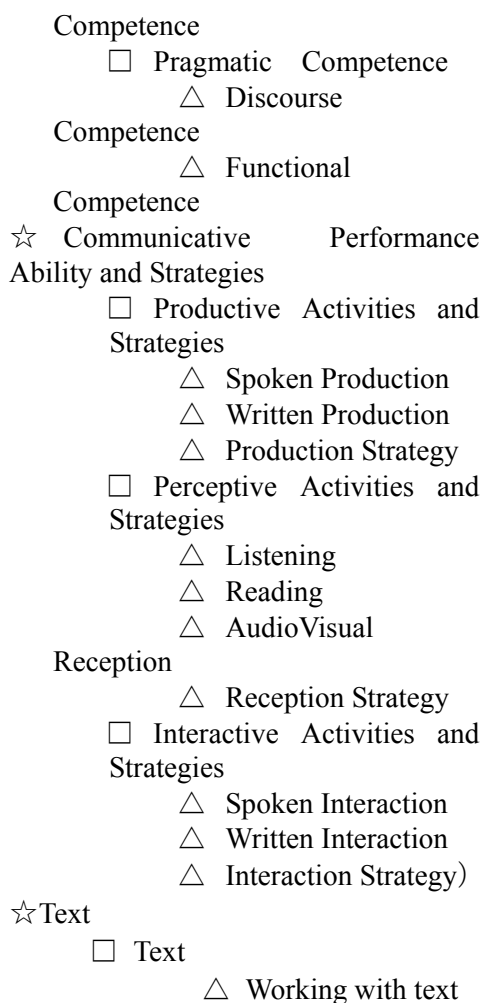
The ACPA aims to accredit courses of practical abilities in English by examining the syllabus, teaching materials, and methods of teaching first. After a learner completes an accredited course, his or her practical skills learned can be described in practical terms and a learner can obtain certificate which describes his skill content. We can also suggest to a learner about his or her further studies to improve his or her practical skills required in her or his future occupation.

The ACPA standard descriptors clarify the objective and goal of the course and the level of learning. It also provides guidelines to develop systematic curriculum, and learning materials so that one can prevent a mismatch between a learner's learning objectives and the course objectives.

### 2.3 ACPA Format

The ACPA format has six sections: Level, Category, Sub-category, Skill Set, Items (elements) and Descriptors. Level in heads represents six proficiency levels: Beginners (A1), Basic (A2), Pre-Intermediate (B1), Intermediate (B2), Pre-advanced and Advanced (C1), Advanced and Advanced + (C2). Categories represent Language Competence, Communicative Ability and Strategies and Text. Each category has subcategories; for instance, Language Competence has four subcategories: lexical competence, grammatical competence, phonological competence and orthographic competence. The CEFR descriptors are arranged hierarchy in terms of the language use domains. The ACPA format can be outlined as follows:





Note : ☆ : Categories, □ : subcategories △ : Skill Sets

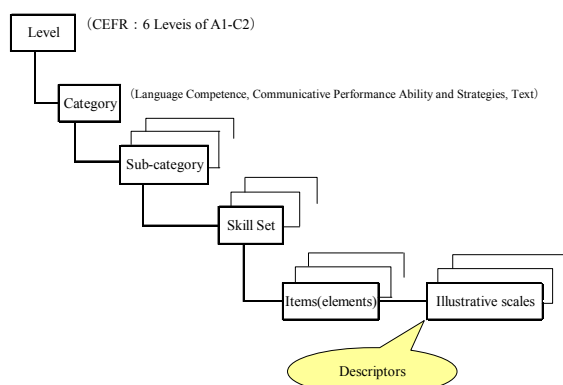


Fig.1 Structure of Skill Standard Guide of English

## 2.4 Levels and Courses

Six levels are set up in relation to the CEFR.

### Beginners Level 【A1】

ACPA certificate does not include A1.

### Basic Level 【A2】

Candidates who complete English Tutorial Basic Level have to take WeTEC Test whose test scores need to be more than 573 which is approximately

525 TOEIC score.

### Pre-Intermediate Level 【B1】

Candidates who complete English Tutorial Pre-Intermediate Level have to take WeTEC Test whose test scores need to be more than 608 which is approximately 565 TOEIC score.

### Intermediate Level 【B2】

Candidates who complete English Tutorial Intermediate Level have to take WeTEC Test whose test scores need to be more than 733 which is approximately 705 TOEIC score.

### Pre-advanced Level 【C1】

Candidates who complete English Tutorial Pre-advanced Level as well as Discussion Tutorial Intermediate and whose course grade must be A+ or A have more than 800 WeTEC score which is approximately 780 TOEIC score.

### Advanced Level 【C2】

Candidates who complete English Tutorial Advanced Level as well as either Discussion Tutorial Advanced Level or English Tutorial Advance + and whose course grade must be A+ or A have more than 880 WeTEC score which is approximately 870 TOEIC score.

The document of CEFR consists of the nine chapters. We have adopted descriptors in Chapters 3, 4, and 5. ACPA skill item description was drawn up by arranging and reorganizing the illustrative descriptors described in these chapters of CEFR.

## 4.1 Example in the case of Tutorial English Course of Pre-intermediate

### ① Course Certification

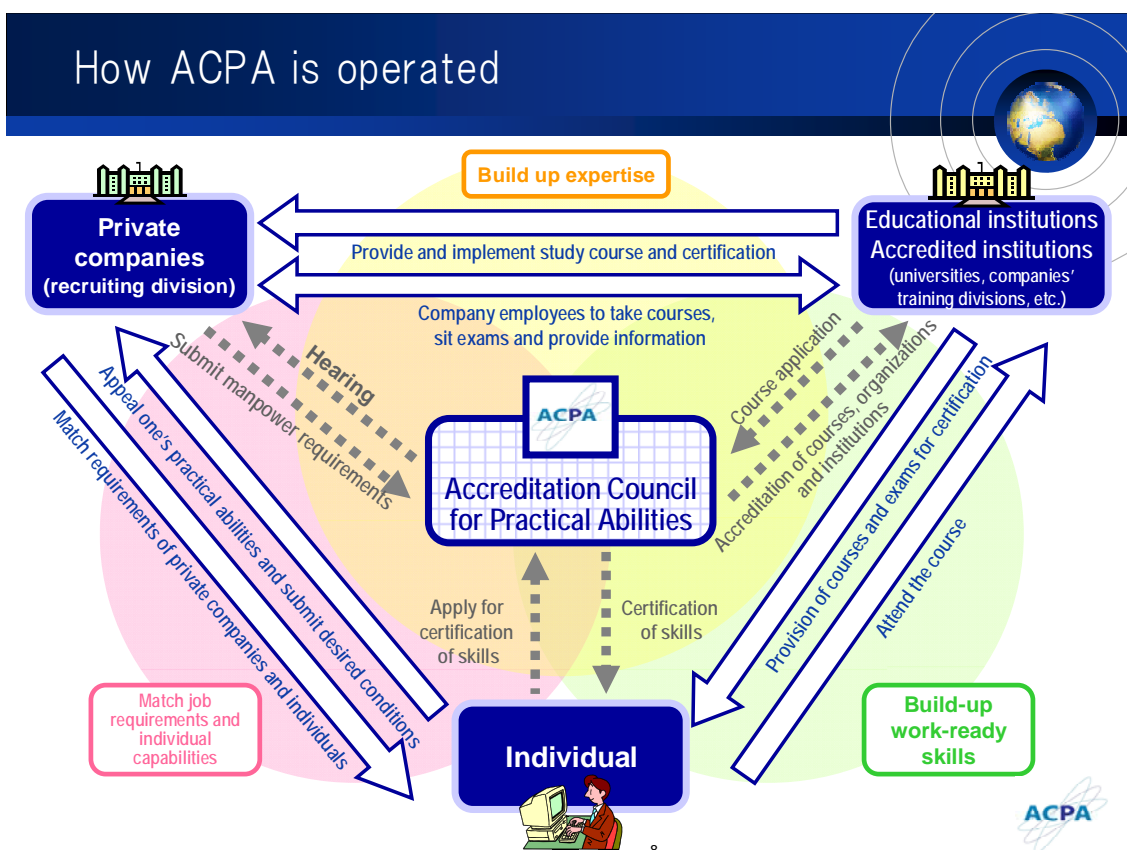
When it is judged that items marked with ● of a course of Intermediate meets the course certification screening criterion of ACPA, the skill of the course certified serves as items shown by ● marks.

### ② Skill Certification

When a student complete and pass the examination of the course of the Tutorial English Intermediate course, ACPA recognizes that the student acquired the skills of four items of language use capability indicated by mark ●, that is, Spoken Production, Production Strategy, Spoken Interaction, Interaction Strategy.

## Reference

European Council (2001). The Common European Framework of Reference for Languages” [http://www.coe.int/T/DG4/Portfolio/documents/Framework\\_EN.pdf](http://www.coe.int/T/DG4/Portfolio/documents/Framework_EN.pdf)



## Preliminary assessment of Cross-Cultural Distance Learning (CCDL) competence using Discourse Completion Tasks (DCT)

Michiko Nakano<sup>1</sup>, Junko Negishi<sup>1</sup>, Kazuharu Owada<sup>2</sup>, Remi Murao<sup>1</sup>, Masanori Oya<sup>1</sup> Tae Yamazaki,  
Miyasaka, N., and Norifumi Ueda<sup>1</sup>

<sup>1</sup>Waseda University, <sup>2</sup>Tokyo College of Music

nakanom@waseda.jp, jnegishi@dream.com, murao@aoni.waseda.jp,  
uedan@aoni.waseda.jp

### Abstract

In this paper, we discuss how to assess CCDL competence using Discourse Completion Tasks (DCT). This is preliminary to automatic assessment of oral responses to DCTs and we are interested in finding out whether DCT is suitable for our future research into automatic assessment of oral responses. We gave a list of 10 contexts of situations to our students asked what they would say in a given situation. Their responses were collected in textual format. These DCTs are distributed online via Learning Management System called Course N@vi. 66 students participated in this preliminary investigation; out of 66, 45 were CCDL participants and the remaining 21 students were non-CCDL students. The two types of participants were compared in terms of response patterns in the present study. Their responses are downloaded and analyzed according to the coding schema by Blum-Kuluka and Kasper's CCSARP project (1989) which has been used by various researchers in the past: Hill (1997), Nakano et al. (1999), Nakano(2003) and Negishi(2006). We also adopted speech act classifications by Aijmer(1996). We will discuss which coding is useful for the present situation.

### Keywords

CCDL, DCT, Intercultural communication

### Introduction

Cross-cultural Distance Learning Courses at Waseda University are joint courses with overseas partners in China, Korea and Taiwan. Theme-based Cross Cultural Distance Learning (CCDL) in 2009 Spring Semester deals with two themes: Social & Global Issues and Media. This is a part of the collaborative joint cyber seminar among Asian Universities initiated by Waseda University.

Waseda students are encouraged to take this CCDL course after they complete Intermediate Tutorial English (above WeTEC score 733, TOEIC score 703). In 2009 spring semester, average WeTEC scores are 730 among Waseda students and 880 among Yonsei students. The main purpose of this CCDL course is that Asian youths should be able to discuss current problems in the world in English and that they should offer their own tentative solutions to the problems which may be acceptable compromise among Asian students whose world views are diverse. For instance, the specific topics in Social and Global issues include Happiness Factors (Co-existence in Asia seen from individuals), Family roles in the changing society and Climate Change, Spring semester. The course also contains lessons on cross-cultural concepts and transferable skills as reported in Nakano et al. (2007a, 2007b, and 2007c). In this paper, it is supposed that cross-cultural concepts and skills can promote interpersonal skills which are related to speech acts realization patterns. According to Cohen (1996), "the first concern of SLA researchers has been to arrive at the set of realization patterns" which he called "speech act set": Cohen (1996:385). Discourse Completion Tasks should be able to define the preconditions and interactional goals of speech act explicitly, so that DCT should elicit what Cohen called "socio-cultural ability" which enables a speaker to achieve his intended perlocutionary effect. In this experiment we examine our DCTs can achieve our desired goal to maximize a learner's socio-cultural ability. Cohen's

socio-cultural ability includes a learner's choice of a relevant expression relative to not only such situational factors as the status of the speaker and hearer, age, power relationships, but also such personal factors as familiarity, and psychological distance. This means that our participant exhibits some sensitivity toward an addresser variability implicit in our DCT situations.

## 1 Background of Request Speech Acts

In discussing discourse situations, what Yule (1996:60) remarks on the relationship between situations and speech acts in general has been well-received among researchers:

As a technical term, **face** means the public self-image of a person. It refers to that emotional and social sense of self that everyone has and expects everyone else to recognize. **Politeness**, in an interaction, can then be defined as the means employed to show awareness of another person's face. In this sense, politeness can be accomplished in situations of social distance or closeness. (The two words were bold-faced in the original text.)

What is apparent in this extract is that our awareness of politeness is necessary for our daily interactions, and at the same time, our awareness of social situation is essential. Yule (1996:61) goes on to say:

Within their everyday social interactions, people generally behave as if their expectations concerning their public self-image, or their **face wants**, will be respected. If a speaker says something that represents a threat to another individual's expectations regarding self-image, it is described as a **face threatening act**. (The two words were bold-faced in the original text.)

Brown and Levinson (1987:76) present a formula in order to calculate the seriousness of a face threatening act (FTA) as follows:

$$W_x = D(S, H) + P(H, S) + R_x$$

$W_x$ : the numerical value that measures the weightiness of the FTA<sub>x</sub>

$D(S, H)$ : the value that measures the social distance between S and H

$P(H, S)$ : a measure of the power that H has over S

$R_x$  a value that measures the degree to which the FTA<sub>x</sub> is rated an imposition in that

culture

note: S stands for speaker, H stands for hearer

Requests are face threatening acts (Brown and Levinson 1987; Suh 1999), thus we can state that the choice of requestive strategies is influenced by the three factors in the formula above: social distance between a speaker and a hearer, power that a hearer has over a speaker, and degree of imposition that a speaker's request would put on a hearer. It is important to bear this claim in mind whenever we do research on requests. For this reason, Nakano et al. (1999) we have distributed the questionnaire about how a participant interprets social distance between S and H, power relationship between them, and a degree of imposition inherent in one's request, since a participant's interpretation varies individually. However, in this preliminary research we have omitted this aspect of investigation.

### 1.1 Request Speech Acts and CCDL

One of the speech act practice in CCDL courses is concerned with a facilitator role. Facilitating is one of the core CCDL skills that students have the opportunity to develop during the course. All students take turns facilitating activities and discussions. Students learn how to facilitate by a) observing others facilitate, b) experiencing being a participant in a discussion facilitated by another student, as well as c) facilitating a discussion themselves.

Depending on their learning styles, some may prefer to try first, others to watch first. Thus asking for volunteers is preferable to allocating roles during the first half of the course (Topics 1-3). In order to experience personal development and to contrast the different environments (face-to-face versus technology interface), the course has been structured to give all students the opportunity to practice facilitating at least twice in both 'Preparation' and 'Joint Lessons', four times in total, by the end of the semester.

Facilitation skills are introduced to the students using a metaphor of engine oil and cooking oil. Engine oil makes the different parts of an engine move smoothly together and prevents the engine seizing. This means that a facilitator helps the conversation in a group of 3 or more people go smoothly between people with different personalities and communication styles. Cooking oil stops the food from sticking to the frying pan. It helps to toss the food easily to stop one side burning while the other side is uncooked. A

facilitator stops the conversation getting stuck, clarifies misunderstandings, enables everyone to take part in the discussion, and has a pool of questions to keep the discussion moving in interesting directions. If the oil in the engine is overfilled, it will leak and the engine will not function at its best. If there is not enough oil the engine may seize. If there is too much cooking oil in the pan, the food itself will become oily and the flavor of the oil overpowers taste of the food. The facilitator's role is not necessarily to be as noticeable as that of a chairperson, although sometimes clear direction such as typing discussion questions into the Text Box, and clear structure such as asking for contributors by name can help. A facilitator helps their group to get the best out of the discussion and helps each person to give their best.

The following points are the responsibility of facilitators;

1. Being the engine/cooking oil in the system; helping things go smoothly; stopping things from getting stuck, keeping the conversation going.
2. Writing main questions in the Text Box.
3. Ensuring everyone is participating and can understand the discussion.
4. Encouraging a student to speak up or ask questions if they have been quiet for some time.
5. Offering opportunities to all students to contribute and preventing one person from dominating the discussion.
6. Deciding who should go first if two or more people want to speak at the same time.
7. Directing discussion by calling on individual people by name.
8. Typing in key words in the TextBox.

After briefly introducing the role of a facilitator, students will be given a facilitation exercise as a homework which asks them what they would do or say in certain situations in a discussion. For example, one student is very talkative and talking about 70 percent of the time while others are remaining quiet most of the time. In the preparation lesson, students will share their answers in group and discuss what they should say or do.

The instructor will elicit responses from different groups for the questions. The purpose of this activity is to a) give students phrases they can use if they encounter those situations, and b) get students used to the idea of asking themselves "why" they think a student was behaving in a certain way.

## 1.2 Speech Acts and Elicitation Materials

In this experiment, we have included five facilitator situations:

1 One student is talking in a very quiet voice and you cannot understand what she/he has just said to you. What are you going to say to this quiet person? [Facilitator's role: Request]

2 One student is very talkative and he/she has been talking about 70 percent of the time. Others are remaining quiet most of the time. What are you going to say to this talkative person? [Facilitator's role: request and remind ?]

3 Students start talking off topic, or about irrelevant things. What are you going to say to your partners? [Facilitator's role: Request]

4 When you are listening to someone, how do you show to your partner that you are paying attention to his/her talk? [Backchannel respect or non-verbal gesture → compliment]

5 When you are interested in what your partner is talking about, how do you communicate your interests to your partner? [Backchannel respect or non-verbal gesture → compliment]

Situations 4 and 5 are relevant to non-facilitator participants as well, but the facilitator needs to show his or her interest and sympathetic understanding in other participants most. In most speech act studies in the past, apology, thanking, request, offer, refusal and compliment have been studied. It is unique for us to include speech acts relevant to oral communication such as backchannel respect which may be subsumed under compliment speech act. These five DCTs have been practiced during the lessons, but we have elicited from students who have not been instructed explicitly.

As for the transferable skills, we included DCTs which have not been dealt with during the lessons:

6 You have arranged to meet your friends to study together for an exam. But you have forgotten the arrangement. What are you going to say to your friends afterward? [Apology, equal status]

7 Mika is arranging to have a study meeting for writing a report. You think that each of us should write a report in our own way. You want to refuse to join the meeting. What are you going to say? [Refusal, equal status]

8 You lent your friend 50 dollars three weeks ago. Though she promised to return it within a week, three weeks have already passed. What would you say to the friend? [Request, equal status]

9 You live in an apartment house. You are often irritated with the loud noise that your neighbor's children make. Today you decide you cannot endure any more. So you go to your neighbor and what would you say? [Request, status involves some subjective understanding the relationship with the

neighbor]

10 You are chair of the student union at your college. One of the members, who is the freshman, always comes late for the meeting. Today is the fifth time that he has come late. What would you say? [Request, higher status]

Table 1 Summary of DCTs

Situation	Social Distance	Dominance	Imposition	Speech Acts
1	-SD	-D	low	Request
2	-SD	-D	low	Request
3	-SD	-D	low	Request
4	-SD	-D	low	Request
5	-SD	-D	low	Request
6	-SD	-D	-	Apology
7	-SD	-D	subjective	Refusal
8	-SD	-D	subjective	Request
9	subjective	subjective	subjective	Request
10	+SD	+D	subjective	Request

As Table 1 indicates, we have some uncertainty in our data, since situational understanding varies individually and we have not included questionnaire asking about the self-perception of situational factors, particularly about the degree of imposition. One justification comes from our previous experiments in 1999 in which all the expression elicited were those learned at junior high school or high school textbooks on thanking, apology, requests and offers.

### 1.3 Coding Schema for Request

In this preliminary analysis, we adopt two methods: Bum-kuluka, House and Kasper (1989) and Aijimer (1996). The former is used for Requests. The coding schema adapted in the study had four categories:

Category 1: Alerters

Category 2: Head Acts categorised by directness level

Category 3: Internal modification

Category 4: External modification

In what follows, we follow Negishi (2006) and in contrast Nakano (2003) provides a shorter version.

An alerter in Category 1 refers to an opening device that alerts the hearer's attention prior to the actual request. Alerters comprised one of two types: terms of address and attention getters (e.g., Excuse me).

Category 2: Head Acts categorised by directness level

Head Acts, actual requests, were categorised into three types according to the senses of directness: direct requests, conventionally indirect requests and hints. **Direct Requests** are subdivided into four: Imperatives, Performatives (e.g., I ask you to help me.), Hedged Performatives (e.g., I need to ask you.) and Obligation (You should help me.). **Conventionally indirect requests** are most commonly used:

1	Want statement	<i>I'd like you to help me.</i>
2	Suggestory formula	<i>How about giving me a hand?</i>
3	Preparatory	
a	Ability	<i>Can/Could you lend me some money?</i>
b	Possibility	<i>Would/Could it be possible or is it possible...?</i>
c	Willingness	<i>Will/Would you help me?</i>
d	Permission	<i>May/Can I borrow some money?</i>

**Non-Conventionally Indirect request** stands for Hints such as I left my wallet at home rather than saying I want to borrow some money.

#### Category 3: Internal modification

Internal modification refers to downgraders, or upgraders; they mitigate or enhance the illocutionary force of a request. They are optionally used within the Head act. There are six syntactic downgrading devices and six types of lexical phrases. Upgraders function to increase the impact of a request.

1 Adverbial Intensifiers

2 Commiment Upgrader ( I am sure ~.)

3 Lexical Intensification

#### Category 4: External modification

External modification refers to supportive moves with which a speaker intends to mitigate or aggravate the request. Supportive moves are located before and/or after the Head Acts, not within the Head Act. There are eight sub-categories.

#### **1.4 Simpler Coding in Nakano (2003)**

Blum-Kulka, House, and Kasper (1989) suggested the following 9 strategies on the scale of indirectness.

1. mood derivable: utterances in which the grammatical mood of the verb signals illocutionary force ( 'Leave me alone'; 'Clean up that mess').

2. performatives: utterances in which the illocutionary force is explicitly named ( 'I am asking you to clean up the mess').

3. hedged performatives: utterances in which the naming of the illocutionary force is modified by hedging expressions ( 'I would like to ask you to give your presentation a week earlier than scheduled').

4. obligation statements: utterances which state the obligation of the hearer to carry out the act ( 'You'll have to move that car').

5 . want statements: utterances which state the speaker's desire that the hearer carries out the act ( 'I really wish you'd stop bothering me').

6. suggestory formulae: utterances which contain a suggestion to do x ( 'HOW about cleaning up?').

7 . query preparatory: utterances containing reference to preparatory conditions (e.g., ability,

willingness) as conventionalized in any specific language ('Could you clear up the kitchen, please?'; 'Would you mind moving your car?').

8. strong hints: utterances containing partial reference to object or element needed for the implementation of the act ('You have left the kitchen in a right mess').

9. mild hints: utterances that make no reference to the request proper (or any of its elements) but are interpretable as requests by context ('I am a nun' in response to a persistent hassler).

Blum-Kulka, House, Kasper: (1989:18).

Strategies 1-5 are regarded as "Direct", "Speaker-oriented" and "Impositives", Strategies 6 and 7 are "Conventionalized Indirect", "Hearer-oriented" and "Query Preparatory", Strategies 8 and 9 are "Nonconventionalized Indirect", "Mutually-oriented" and "Hints", according to Blum-Kulka, House, Kasper. Thus, it is reasonable to accept House (1989:102-105) who focused on the following three strategies: 1) Imperatives ('direct') e.g., *Leave me alone; Clean up this room*; 2) Query Preparatory ('conventionalized indirect'): utterances containing reference to preparatory conditions as conventionalized in any specific language e.g., *Could you clean up the room?*; 3) Hints ('nonconventionalized indirect') (e.g., *You have left the kitchen in a right mess.*)

Blum-kulka's classification is based on either the discourse orientation such as hearer-oriented, speaker-oriented, inclusive or impersonal (Blum-Kulka (1989:59)) or on the scale of directness – indirectness, ranging from Direct, Conventional Indirect to Non-conventional Indirect (Blum-Kulka (1989:47)). As we pointed out above, House (1989:102-103) related direct/indirect scale to syntactic realization patterns: Imperatives (direct), Query Preparatory (conventionalized indirect) and Hints (non-conventionalized indirect). We can suggest the following relationships:

Imperatives (direct) → speaker-oriented and impositives

Interrogatives and Negative Interrogatives → Query Preparatory, hearer-oriented, conventionalized

indirect

If the speaker uses *Let us* or *Shall we*, emphasizing the use of inclusive *we*, then the utterance is inclusive: e.g., it's time that we cleaned the room.

If the speaker uses impersonal pronouns such as *one* or the utterance is stated in general terms, then the utterance is impersonal; e.g., one needs to clean the room once in a while, it is our school policy to clean the room once a day, etc.

This simpler coding is easy to use and worked well for cross-cultural data comparison among British speakers, German speakers, Austrarilan English speakers, Argentinian Spanish speakers, Hebrew speakers, French speakers and Japanese Learners of English.

### 1.5 Coding Schema for Apology

Olshtain and Cohen (1990:46) defines the speech act of apology as follows:

An apology is a speech act which aims to provide support for the hearer (H) who was actually or potentially mal-affected by a violation (X) for which the speaker (S) is at least partially responsible. When apologizing, the S is willing to humiliate him/herself to some extent and to admit to fault and responsibility for X. Hence, the act of apologizing is face-saving for the H and face-threatening for the S, in Brown and Levinson's (1978) terms.

With regard to the apology speech act, we can assume that there are two participants involved: the speaker, who is responsible for causing the offense, and the hearer, who receive the apology. According to Olshtain and Cohen (1983:21), we can describe the apology speech act from the following five points of view:

1. The recipient's expectations based on his/her perception of the degree of severity of the offense
2. The offender's apology based on his/her perception of the degree of severity of the offense
3. The offender's apology based on the extent of reprimanding expected from the recipient
4. The interactive nature of both the initial apology and the recipient's response
5. The social status of the two participants
6. The way the tone of voice may function to convey meaning

As indicated above, we have not distributed self-perception questionnaires on the severity of the offense and the social status of the participants.

This is a serious weak point of the present study.

Although opinions vary as to the classification of apology strategies, we follow the one proposed by Olshtain and Cohen (1983:22-23).

They distinguish the case where the speaker recognizes the need to apologize, and the case where the speaker rejects the need to apologize. Concerning the former case, five strategies and some subcategories are presented as follows:

1. An expression of an apology
  - a. An expression of regret, e.g., *I'm sorry.*
  - b. An offer of apology, e.g., *I apologize.*
  - c. A request for forgiveness, e.g., *Excuse me. Please forgive me. Pardon me.*
2. An explanation or account of the situation  
e.g., *The bus was delayed.*
3. An acknowledgement of responsibility
  - a. Accepting the blame, e.g., *It's my fault.*
  - b. Expressing self-deficiency, e.g., *I was confused. I wasn't thinking. I didn't see you.*
  - c. Recognizing the other person as deserving apology, e.g., *You are right!*
  - d. Expressing lack of intent, e.g., *I didn't mean to.*
4. An offer of repair  
e.g., *I'll pay for the broken vase.*
5. A promise of forbearance  
e.g., *It won't happen again.*

On the one hand, the first two strategies are general, that is, they can be used in various situations. On the other hand, the last three strategies are situation specific.

The apology speech act realized by the strategies above might be modified internally or externally (Olshtain and Cohen 1990:47). The intensifiers such as *very, really, terrible, deeply*, and so on, usually occur internally to apology expressions. The external modification emerges in the form of a comment added to apology expressions. They can strengthen or diminish the degree of sincerity of a speaker.

As mentioned before, there might be the case where a speaker does not intend to apologize. In such cases, the person who caused some offense can express 1) no reaction, 2) a denial of the need to apologize, e.g., *there was no need for you to get insulted*, and 3) a denial of responsibility by not accepting the blame, e.g., *It wasn't my fault*, or by blaming the other participant for bringing the offense upon him/herself, e.g., *It's your own fault*.

(Olshtain and Cohen 1983:23)

For our coding, we compared Aijimer (1996) with Olshtain and Cohen (1990) mentioned above. Aijimer (1996) divided Apology expressions into two types: explicit and implicit. Each type is subdivided into emotional and non-emotional.

**Explicit emotional apologies** include: (A) explicitly apologizing, e.g. I apologize (for) and (D) expressing regret, e.g. I'm sorry, I'm afraid

**Explicit non-emotional apologies** include four expression patterns: B) offering or giving, presenting) one's apologies, e.g. I present my apologies, (C) acknowledging a debt of apology e.g. I owe you an apology, (E) demanding forgiveness (e.g. pardon me, excuse me), and (F) explicitly requesting the hearer's forgiveness, e.g. I beg your pardon.

**Implicit emotional apologies** include two expression patterns: (G) giving an explanation of account, e.g. (I'm sorry) it's so unusual and (J) expressing emotion, e.g. oh (I'm so sorry). **Implicit non-emotional apologies** refer to the following five expression patterns: (H) self-denigration or self-reproach, e.g. how stupid of me, how awful, I ought to know this; (I) minimizing responsibility, e.g. I didn't mean to ..., I thought this was ..., I was thinking it was...; (K) acknowledging responsibility for the offending act, e.g. that was my fault (Fraser 1981: 263); (L) promising forbearance from a similar offending act, e.g. I promise you that that will never happen again (Fraser 1981: 263); (M) offering redress, e.g. please let me pay for the damage I've done (Fraser 1981: 263).

## Partial Reference

Negishi, J. (2006). *Australians' Perceptions of Requests Employed by Japanese Speakers of English*. Unpublished M.A. thesis, Sydney University.

## Cyber course on World Englishes and ELF: Some tentative evidence

Michiko Nakano and Yuko Haraguchi

Director, Distance Learning Center, Waseda University

nakanom@waseda.jp

### Abstract

Since 2002, we have run the cyber course 'World Englishes and Miscommunications' in which nine varieties of English in Asia have been discussed online. We have since then obtained sufficient data both experimental and observational, for us to be able to discuss Asian Englishes in general. Firstly, in Asian Englishes there are grammatical features which are different from Native Speaker (NS) norms in our colloquial speech as well as in writing. Although our mother tongues are all different, our features different from the NS norms are all identical in Asia. In Section 1, we will compare our typical features presented in the on-demand lectures with those in ELF lexico-grammar generalized from Seidlhofer's Vienna-Oxford International Corpus of English (VOICE): Seidlhofer (2002). In Section 2, we will discuss Lingua Franca Core (LFC): Jenkins, 2000.

### Keywords

Asian Englishes, English as a Lingua Franca (ELF)

### Introduction

Firstly, in Asian Englishes there are grammatical features which are different from Native Speaker (NS) norms in our colloquial speech as well as in writing. Although our mother tongues are all different, our features different from the NS norms are all identical in Asia. In Section 1, we will compare our typical features presented in the on-demand lectures with those in ELF lexico-grammar generalized from Seidlhofer's Vienna-Oxford International Corpus of English (VOICE): Seidlhofer (2002). The present data appear to confirm ELF lexico-grammar. We will then argue that these features are the outcome of Grammar simplification. Kirkpatrick (2005 and 2007) points out that "any apparent systematic simplification of the inflectional system follows historical principles and is thus likely to be characteristic of a new variety of English": (Kirkpatrick, 2005, p.47). It is interesting to note that Englishes in the expanding circle follow the same pattern of simplification. The present section endorses that it also replicates the

developmental simplification patterns across three educational levels among Japanese learners of English.

In Section 2, we will discuss Lingua Franca Core (LFC): Jenkins, 2000. In order to avoid communication breakdown caused by lack of mutual intelligibility, Jenkins and Gimson make some suggestion for phonological core features for basic intelligibility (Jenkins, 2001 and Gimson, 2001). The greatest characteristic in the LFC presented by Jenkins lies in its practicality and it makes our goal of teaching feasible and practical for Asian users of English. On the other hand, in Gimson's proposal, learning inventory is traditional and we have to learn most of NS norms. With this background in mind, we examine the following features

1. The aspiration of /p, t, k/ in Experiment 1
2. The distinction of /i:/, /I/ in Experiment 2

The purpose of Experiment 1 is two-fold: 1) to investigate whether or not adult Japanese users of English (JUE) can pronounce aspirated and unaspirated voiceless stops in English distinctly, and 2) to examine whether or not the aspiration of the voiceless stops is learnable for adult JUEs without explicit instruction. Asian users of English who use English as a tool of communication are not often phonetically trained. In LFC, the duration is regarded as more crucial than the quality in [i:] and [I]. In our native language, we have durational difference between the two, but there is no qualitative difference. Thus, Experiment 2 is to investigate 1) whether adult JUEs can learn to distinguish the sound qualities of /i(:)/ and /I/ after the training and 2) whether Japanese teachers of English can recognize learners' improvement in quality. The result suggests that although learner can improve sound qualities to some extent, the experienced teachers failed to notice and recognize the improvement which was acoustically manifested in their spectrograms. This suggests that the training does not yield practical use, supporting LFC position that

duration is more helpful in intelligibility than quality at least for Japanese users of English.

# 1 Evidence for ELF Lexico-Grammar from Cyber materials

Seidlhofer (2004:220) proposes the following ELF Lexico-Grammar.

- 1 non-use of the third person present tense -s, (as in “She look very sad.”)
- 2 omissions of the definite and indefinite articles where they are obligatory in NS English, and insertion where they do not occur in NS English
- 3 heavy reliance on verbs of high semantic generality, such as get, make, have, do, etc
- 4 pluralization of nouns which are considered uncountable in NS English, e.g., informations, staffs, advices, furnitures, softwares
- 5 addition of unnecessary prepositions, such as “discuss about” or “study about.”
- 6 use of an all-purpose tag questions, e.g., isn’t it? or no?

Adapted from Seidlhofer (2004:220)

Cited by Jenkins (2005:5)

Japanese English in the on-demand lecture is explained by dividing it into three formal educational levels: junior high school level, high school level and university level. Interestingly, we have found that most grammatical features different from NS norms are similar all through the three educational stages. The on-demand lectures on other varieties of English in Asia indicate that those are also observed among English users of other Asian varieties. The following numbered list summarizes tendencies in Asian Englishes lectured by researchers, compared with ELF Lexico-Grammar (each number corresponds to Seidlhofer’s item number above):

- 1 Japan English, Korean English, China English, Thai English, Singapore English, Malay English, Hong Kong English Philippine English and India English
- 2 Japan English, Korean English, China English, Thai English, Singapore English, Malay English, Hong Kong English, Philippine English and India English
- 3 Not discussed except for Japan English and Englishes in Europe
- 4 Japan English, Korean English, China English, Thai English, Singapore English, Malay English, Hong Kong English, Philippine English and India English
- 5 Japan English, Korean English, China English, Thai English, Singapore English, Malay English, Hong Kong English, Philippine English and India English

6 Indian English and Estuary English [The Numbers correspond to ELF Lexico-Grammar cited above]

In order to support the present claim, I will refer to the corpus-based research done in my research group, to illustrate the point concerning Japan English. First, look at Baustista’s corpus analysis in Table 1 in which grammatical features seen in Philippine English are listed. The frequency list of features is extracted from Corpus of Written English consisting of 150 files each of which contains 2000 words up to 2300 words: see Baustista (2000).

Table 1 Number of Deviant Cases Found for Each Grammatical Category

Grammatical Categories	Number	Percent
Subject-Verb Agreement	136	27
Articles	110	22
Prepositions	77	15
Tenses	77	15
Mass and Count Nouns	46	9
Pronoun-Antecedent	23	5
Others	34	7
Total	503	100

Baustista (2000:36)

Observe the data for Japanese learners of English (JLEs) at the high school level in Fig. 1, Appendix 1. This spoken corpus is composed of Oral Speech among 58 high school students. Ano (2002) observed 644 non-NS features. In this corpus, Ano did not count the subject-verb number agreement errors, since it is difficult to detect such errors in oral speech. But frequent errors (i.e., features different from NS norms) occur in the use of tense, articles, and prepositions, which is the salient features in other Asian varieties. It is widely known that JLEs make mistakes in collective nouns, by regarding them as countable nouns. This is also common among other varieties in Asia.

The identical tendency was observed in Spoken English Corpus among university students. This time, 30 university students talked in English in such tasks as picture descriptions, role-play, and story-telling. They talked about 15 minutes. We obtained 1511 erroneous features: see Fig.2 Error types in oral English at the university level in Appendix 2.

Table 2 Comparison between JLEs and Phillipinnes

	JLEs (Speaking)	Phillipino (Writing)
Subject-Verb Agreement	93	136
Articles	370	10
Prepositions	202	77
Tense	201	77
Nouns (countable vs uncountable)	157	46
Pronouns	89	23
Others	399	34

These comparisons suggest that in Asian Englishes we exhibit similar grammatical features different

from NS norms. Particularly in Japan English the same error types persist at all the three educational stages. Although our mother tongues are all different in Asia, our features which are different from the NS norms are identical in both outer circle and expanding circle Englishes in Asia. The textbooks in Asia are however based on NS Englishes which do not contain any features noted in ELF Lexico-Grammar. For this reason, some researchers characterize the phenomena as restructuring: see Kirkpatrick, 2007. Error sources might be argued as the outcome of cognitive activities we all share. Nakano (2007b) argues that it is the outcome of universal mechanisms of grammar simplification, as shown in Bresnan (2000). Nakano (2007b) also demonstrates that English among JLEs reflects the influence of input our learners receive in their formal education and some language-independent mapping mechanisms of argument-structures onto grammatical functions. But in this section, we will focus on whether these features can be described as grammar simplification. Our features different from NS Englishes are those in VOICE, as we have seen in Section 1.

All the six items can be considered as grammar simplifications. If we can confirm that they are the outcome of internal grammar simplifications, it indicates that it is the inevitable result of cognitive activities, perhaps to reduce cognitive load and efforts on the part of non-native users of English, in order to meet the communicative pressure. For this reason, it is crucial to check whether the features can be described as the outcome of internal grammar simplifications. These simplifications do not occur systematically; therefore we cannot represent them as categorical rules. They are variable, depending on each lexical item. For this reason, Lexical Functional Syntax (LFS) provides optimal descriptive framework. In Feature 1, NS English can be described, following LFS format:

Look VI  
 -s: Infl<sub>v</sub> ↑ Tense = pres  
           ↑ subj = ↓  
                   ↓ pers = 3  
                   ↓ num = sg

This lexical rule states that if the verb 'look' is in the present tense and the subject noun is third person singular, the verb needs to be inflected as "looks." But this lexical rule is missing in colloquial Asian Englishes; hence it is a grammar simplification.

Feature 2 is variable use of definite and non-definite articles. In LFS, it can be represented as follows:

The Det  
 ↑ def = +  
 A Det  
 ↑ def = -

This lexical rule state that the determiner 'the' requires the notion of definiteness, while the determiner 'a' does not. This lexical rule is variably missing in colloquial Asian Englishes; hence it is a grammar simplification.

Feature 3 is also seen in many varieties of NNS Englishes: See Appendix 3, Tables 3 ~7. This overuse of the verbs of high semantic generality can be interpreted as simplification of verb subcategorizations. The same argument holds in the case of Features 4, 5, and 6. In Feature 4, NS English has the following lexical rule:

Furniture: ↑ Collective Noun ↑ pred =  
 'furniture'  
                   ↓ num = sg

The lexical rule states that if it is a collective noun, the number is necessarily singular. But in the colloquial Asian English and ELF, this lexical rule is variably missing and the lexical rule for the common nouns is applied; hence it is an instance of grammar simplifications. Feature 5 ( addition of preposition, about) is also a failure of sub-categorization with a touch of mother tongue semantic transfer. In Japanese, the following verbs all need to co-occur with the preposition, "about."

Talk about  
 Speak about  
 \*study about  
 \*discuss about

However, in NS English, study and discuss are transitive verbs. In our colloquial mode, we fail to notice that they are transitive verbs and our mother tongue semantics intuit us to say "study about ~" and "discuss about ~." But in terms of NS English, this is the case of the failure of subcategorization; hence, it is a form of grammar simplification.

Feature 6 deals with tag questions. In NS English, there are varieties of tag-questions, depending on the subject and verb of the main clause. If all the tag-questions are replaced by a single form, "isn't it?", it is the simplification of grammar.

In this section, we have provided some evidence to support ELF lexico-grammar whose 6 features are all observed among Asian varieties of English. I have characterized these features as internal grammar simplifications. Asian users of English receive the standard norms in their input at least at their formal education, but they appear to simplify the forms cognitively in order to meet the communicative pressure, which might account for the identical features among Asian users of English.

Kirkpatrick (2007) argues that “the simplifications of syntax ... are features of new varieties of English and mirror the way inflections have become simpler within traditional English itself”: Kirkpatrick, 2007, p. 163. Therefore, we might argue that simplification takes place, diachronically, synchronically as well as developmentally.

## 2 Experimental studies on ELF core phonetics

In this section, I will refer to the two experiments concerning Dr Jenkins’s proposal of ELF core phonetic inventory. The experiments reported here were carried out by Haraguchi and analyzed the data jointly by the author.

### 2.1 Background

In order to avoid communication breakdown caused by lack of mutual intelligibility, Jenkins and Gimson make some suggestion for phonological core features for basic intelligibility (Jenkins, 2001 and Gimson, 2001). Gimson argues the items shown below are required for basic intelligibility:

All consonants except /hw/

Six short monophthongs: /ɪ/, /ɛ/, /a/, /ɒ/, /ʊ/, /ə/

Seven long monophthongs: /i:/, /e:/, /a:/, /ɔ:/, /o:/, /u:/, /ə:/

Three diphthongs: /aɪ/, /aʊ/, /ɔɪ/  
(Gimson, 2001, p.310)

On the other hand, Jenkins proposes another concept of the Lingua Franca Core (LFC). LFC presented by her is as follows:

All consonantal phonemes except /θ/, /ð/, /ʃ/, /ʒ/

Allophones of unvoiced consonants (/p, t, k/) by the distinction of aspiration

Long-shot contrast (quantity) of vowels

Diphthongs /aʊ/, /aɪ/, /ɔɪ/

/ɜ:/ (not to confuse it with /a:/)

(Jenkins, 2001, pp.136-146)

The greatest characteristic in the LFC presented by Jenkins lies in its practicality and usefulness in classroom teaching, as we noted in Introduction. She also says the following items should be excluded from the LFC:

Consonants /θ/, /ð/, /ʃ/, /ʒ/

Vowel diphthongs /ɪe/, /eə/, /ʊə/

Vowel quality

(Jenkins, 2001, pp.137-138, p.145)

The core features shown by Gimson and Jenkins are common on the following points:

Six short monophthongs: /ɪ/, /ɛ/, /a/, /ɒ/, /ʊ/, /ə/

Seven long monophthongs: /i:/, /e:/, /a:/, /ɔ:/, /o:/, /u:/, /ə:/

Three diphthongs: /aɪ/, /aʊ/, /ɔɪ/

The core features by Gimson and Jenkins differ on some points. With respect to the consonants, Jenkins excludes some items from the LFC, unlike Gimson. The exclusion is reasonable regarding the number of the minimal pairs including those consonants. According to Higgins, there are only eight minimal pairs including /ə/ and /ð/ in “the 1974 edition of the Advanced Learners Dictionary, incorporating Mitton's 1990 additions to the word list” (<http://www.marlodge.supanet.com/wordlist/>). Therefore learners do not need to worry about the confusion of these sounds. Also /ɪ/ and /ɪ/ are, as the former is an allophonic element and the latter is not produced by all native speakers, less crucial in the achievement of mutual intelligibility. As for the vowels, though their opinions are congruous about the diphthongs, they differ in one crucial point that Jenkins prioritizes quantity over quality. Our experiment thus concentrates on

The aspiration of /p, t, k/ in Experiment 1

The distinction of /i:/, /ɪ/ in Experiment 2

### 2.2 Experiment 1: Aspiration of [p, t, k]

Experiment 1 is 1) to investigate whether or not adult JUEs’ can pronounce aspirated and unaspirated voiceless stops in English distinctly, and 2) to examine whether or not the aspiration of the voiceless stops is learnable for adult JUEs without explicit instruction. The second point relates to LFC, since Asian users who use English as a tool of communication are not often phonetically trained.

Ten participants are divided into two groups: 1) the novice group and 2) the advanced group, depending on 1) whether they have received phonetic instruction or not, and 2) whether they have stayed abroad or not. If a subject met neither of these conditions, the person was placed in the novice group. On the other hand, if a subject met either of the two conditions, the person was included in the advanced group. Voice Onset Time (VOT) of aspirated and unaspirated /p, t, k/ sounds were extracted from the subjects’ running speech during their picture descriptions. The pictures were designed to elicit words which contain /p, t, k/ sounds. Tables 8, 9 and 11 summarize the result. Table 10 represents VOT among the four NSs in (Lisker and Abramson, 1967, p.10).

Table 8 VOT values by the novice group.

	Mean	SD	N
[p <sup>h</sup> ]	26.29	11.45	11
[p]	30.88	19.97	11
[t <sup>h</sup> ]	32.9	18.44	47

[t]	21.48	7.84	10
[k <sup>h</sup> ]	66.42	26.49	42
[k]	38.74	9.55	8

*N* stands for the number of the tokens.

Table 9 VOT values by the advanced group.

	Mean	SD	N
[p <sup>h</sup> ]	34.00	20.64	56
[p]	25.01	8.26	12
[t <sup>h</sup> ]	63.22	23.42	90
[t]	25.1	5.37	22
[k <sup>h</sup> ]	66.42	26.49	52
[k]	38.74	9.55	16

*N* stands for the number of the tokens.

Table 10 VOT in English word-initial /p, t, k/ (running speech) among 4 native speakers.

	/p/	/t/	/k/
Mean	33	38	48
Range	10–70	10–90	10–91
N.	110	125	163

*N.* in the first column stands for the Number of Tokens.

(Lisker and Abramson, 1967, p.10)

It may be noticed that VOT in the unaspirated sounds among JUEs are close to the aspirated sounds by the native speakers. Especially the subjects of the advanced group, in production of aspirated voiceless stops, they showed much greater VOT values than those by native speakers in running speech. This may be the case of hypercorrection. The difference of VOT values by aspiration was smaller in the utterances of the novice group and as Table 11 indicates, they were confused with aspirated [p<sup>h</sup>].

Table 11 Aspirated – Unaspirated VOT values (msec.).

	Novice	Advanced
[p <sup>h</sup> ]-[p]	-3.59	8.99
[t <sup>h</sup> ]-[t]	11.42	38.12
[k <sup>h</sup> ]-[k]	11.55	27.68

The answer to the research question 1) is partially ‘Yes,’ because the advanced group showed high proficiency in distinguishing aspirated and unaspirated /t/ and /k/. However, they showed much smaller VOT difference in /p/, and they confused aspirated and unaspirated sounds in production of /p/ in some words. Therefore we can say that adult JUEs can acquire the distinction of aspirated and unaspirated voiceless stops, but /p/ is difficult for them to pronounce with proper amount of aspiration.

The answer to the research question 2) is ‘No,’ because the subjects of the novice group, who had not received phonetic instruction, reveal poor proficiency in discriminating aspirated and unaspirated voiceless stops. In all the phonemes they showed similar VOT distributions irrespective

of whether the sound is aspirated or unaspirated, and they even showed confusion of [p<sup>h</sup>] and [p]. On the other hand, the subjects in the advanced group showed greater difference of VOT by aspiration in /t/ and /k/ than the novice group. Advanced group had received phonetic instruction more than one year, and the effect of the instruction may be suggested in their performance. In conclusion, we may argue that JUEs cannot achieve distinction between aspirated and unaspirated voiceless stops without explicit instruction. The present experiment supports Jenkins’s view that all allophones of the voiceless stops with the distinction of aspiration should be taught as ELF core items.

## 2.3 Experiment 2: [i:] vs [I]

In LFC, the duration is regarded as more crucial than the quality in [i:] and [I]. This small experiment is to investigate 1) whether adult JUEs can learn to distinguish the sound qualities of /i(:)/ and /I/ after the training and 2) whether Japanese teachers of English can recognize learners’ improvement in quality.

The qualitative contrast of /i(:)/ and /I/ is observed visually through spectrographic analysis. By observing the first and the second formant (F1 and F2), we can confirm qualitative difference of vowels: see Table 12. The native speakers in Table 12 demonstrate that the qualitative difference is revealed in the difference between Formant 2 and Formant 1. The difference ( F2 – F1) in [i:] is always larger than the difference (F2 – F1) in [I]:

$$(F2 - F1) \text{ in } [i:] > (F2 - F1) \text{ in } [I:]$$

Table 12 Formant Frequency for /i(:)/ and /I/ by native English speakers (Hz.).

	/i(:)/		/I/
	F1	F2	F1
			F2
RP (male)	280	2249	367
RP (female)	303	2654	374
GA 280		2250	400
			1920

The data of RP speakers are from Gimson (2001, p.99), and the data of GA speakers are from Ladefoged (1975, p.99). The data of RP speakers were taken from 5 male and 5 female speakers. The number and the details of the GA speakers are not shown in the source.

### 2.3.1 Procedure

The present experiment was run by Haraguchi (2005) in the following way: 1. the first Perception test, 2. the first production test, 3. instruction, 4. the second perception test, and 5. the second production test. Three to seven days of intervals were inserted between each step, but First Production Test and Instruction were held on the same day. The perception tests are identification tasks. Each of the perception tests has twenty items and the production test, eight tokens.

The traditional training group had studied

video-taped instruction by the author, and after having watched the video, they pronounced the target phonemes and their pronunciations were corrected. The correction was done up to ten times for each subject. In the instruction, we emphasized in the traditional-training group that /i(:)/ and /I/ are different in quality.

The computer assisted instruction was given as follows. The software consists of the following six phases: 1) visual instruction with description of sounds, anatomical diagram and the image of movement of lips to show the movement of sound organs, 2) practice with model pronunciation of words with contrasting sounds, 3) practice with model speech of sentences, 4) listening discrimination tasks on contrasting sounds in sentences, 5) practice of the target phonemes in sentences, and 6) visual feedback with waveforms. The subjects can both listen to the model speech by native speakers and record their own voices for comparison. They were told to have training whenever convenient for them, but not to practice more than ten times in order to control the experimental condition. The subjects in the computer-assisted training group do not receive explicit instruction about the difference of the sound quality between /i(:)/ and /I/.

The perception test is to test participants to what extent they can discriminate the vowels aurally given. Both perception and production tests are composed of the four sections:

Section 1: /easy consonants/ + /i(:)/, /I/ + /easy consonants/

Section 2: /easy consonants/ + /i(:)/, /I/ + /difficult consonants/

Section 3: /difficult consonants/ + /i(:)/, /I/ + /easy consonants/

Section 4: /difficult consonants/ + /i(:)/, /I/ + /difficult consonants/

The ease or difficulty of consonants for JLEs is based on Fujii (1986) and Ono(1986).

### 2.3.2 Result and discussion

Table 13 indicates that the scores slightly dropped in the second test in both groups. The only exception is Section 2; the subjects in both groups show higher scores in the second perception test than in the first only in this section. This suggests that the explicit instruction given by traditional method and by computer assisted instruction were not effective to the participants.

Table 13 the scores in the perception tests.

First Perception Test					
Section	1	2	3	4	Total
Computer	4.4	3.3	4.5	4.3	16.5
	(0.52)	(1.06)	(0.7)	(0.7)	(1.5)
Traditional	3.9	3.8	4.4	4.4	15.6

	(0.88)	(1.32)	(1.07)	(0.85)	(3.5)
Second Perception Test					
Section	1	2	3	4	Total
Computer	3.7	3.7	4	4.2	15.6
	(0.95)	(0.82)	(0.94)	(0.63)	(2.76)
Traditional	3.4	4.2	4	3.6	15.5
	(1.17)	(0.82)	(1.1)	(0.99)	(2.59)

The numbers without parenthesis show mean values of the scores. The numbers in the parenthesis show SDs.

In order to see whether JUEs were successful in discriminating /i(:)/ and /I/ in quality, (F2 values – F1 values of /i/) – (F2 values – F1 values of /I/) was calculated. Their performance was regarded as successful, if they pronounce /i(:)/ with larger difference between F1 and F2 frequency than /I/.

Table 14 implies that the traditional instruction is more effective for improvement in distinguishing vowel quality than the computer-assisted training.

Table 14 Quality Improvement by the computer assisted and traditional training groups (Hz.)

	1st Production Test	Mean	SD
Computer		38.33	673.5
Traditional		60.24	423.44
	2 <sup>nd</sup> Production Test		
Computer		349.26	831.49
Traditional		438.33	739.84

Table 15 refers to the duration measurement.

Table 15 Duration of /i(:)/ and /I/.

	/i(:)/	/I/	/i(:)/	/I/
1st Production Test				
Computer-assisted				
	28.53	12.93	8.63	4.86
Traditional				
	28.96	10.23	18.5	3.62
2nd Production Test				
Computer-assisted				
	33.08	11.84	34.26	4.57
Traditional				
	30.72	11.12	20.59	4.37

The subjects in both groups show almost identical duration in producing the phonemes in the first and second production tests. This may indicate that they still rely on duration in discriminating /i(:)/ and /I/ even after the training. In order to verify this assumption, we examined the data with two-way ANOVA. With respect to the computer-assisted training group, there was a main effect of Phoneme (/i(:)/ and /I/),  $F(1, 159) = 83.92$ ,  $p = 0.00 < 0.01$ . There was no statistically significant main effect of Test (the first and second production tests),  $F(1, 159) = 0.74$ ,  $p = 0.39 > 0.05$ , nor interaction between Phoneme and Test,  $F(1, 159) = 1.97$ ,  $p = 0.16 > 0.05$ . As for the traditional training

group, there was a main effect of Phoneme,  $F(1, 159) = 147.22$ ,  $p = 0.00 < 0.01$ . However there was no main effect of Test,  $F(1, 159) = 0.71$ ,  $p = 0.4 > 0.05$ , nor interaction of Phoneme and Test,  $F(1, 159) = 0.08$ ,  $p = 0.08 > 0.05$ . ANOVA thus suggests the similar tendencies between the two groups of subjects; that is, there was a main effect of Phoneme but not main effect of Test and the interaction between Phoneme and Test in both groups. Therefore, one can claim that the subjects showed significant differences in duration between the two phonemes in the two production tests, but they did not change their performance in producing each phoneme with respect to quality. Therefore, we can confirm that the subjects in both groups relied on duration in distinguishing /i(:)/ and /I/ regardless of the training and test settings.

As for the human judgments, the ten raters with MA holders were asked to judge whether the subjects were successful in discriminating /i(:)/ and /I/ in the production tests. In order to examine internal-consistency reliability of the judgments and test reliability, Cronbach  $\alpha$  and Kuder-Richardson formula 21 (K-R 21) were estimated: Cronbach  $\alpha$  (0.92) and KR-21 0.995 in the first production test, and 0.99 in the second test. Thus, internal-consistency of the judgment data is very high. KR-21 verifies high test reliability of the evaluation sheet. However, except for two raters, they failed to recognize Quality improvement. They told us that they put priority to sound quality in evaluation, since they had studied Gimson for 13 years, whereas the two exceptional raters took duration into account. The raters' focus on quality than quantity may be the reason why most subjects were judged to be unsuccessful in discriminating /i(:)/ and /I/ in both first and second production tests.

In comparison of acoustic analysis and human judgments by the raters, it was revealed that human evaluation does not reflect acoustic analysis accurately. The statistic analysis verified high inter-rater consistency in their judgments by the ten raters, but the evaluation was contradictory to the result of the acoustic analysis. We should note that even expert English teachers could not perceive the improvement of sound quality in subjects' pronunciation of /i(:)/ and /I/. There seem no noticeable merits in providing the quality training in this case. This supports the proposal of LFC which asserts the priority of durational distinction rather than the quality. Although subjects showed some improvement in production of these phonemes in acoustic analysis, the qualitative improvement was not recognized sufficiently by the ten experienced teachers. It indicates invalidity of human judgments in evaluating vowel quality; although the raters

were all expert English teachers, they could not perceive spectrographic improvement by the subjects, and this confirms the difficulty of acquisition of L2 vowel quality.

### 2.3.3 Summary and implications in English Language Teaching

This chapter examined ELF lexico-grammar and a part of ELF Lingua Franca Core. The former was examined in terms of cyber materials and discussion as well as some corpus-based research by my graduate seminar group at Graduate School of Education, Waseda University. It was found that ELF lexico grammar describes our features in our spoken mode among Asian varieties of English. This suggests that at least in our spoken mode Non-NS features among Asian users of English could be tolerated.

As for ELF LFC, our experimental findings confirms LFC in the area of aspiration in stops and duration in [i:] and [I]. It was interesting to notice that even experienced teachers who are phonetically trained failed to notice the acoustic improvement exhibited by Japanese users of English.

If a learner is content with achieving his own goal, say, as a businessman [Toeic 630] he or she could rely on durational difference of vowels as Jenkins (2001) suggests. There is also a time factor: a business person would not have time to spare for such task, so he or she should not be bothered by acquisition of the quality distinction. However, if a learner is to develop his or her career as a language teacher, the person should know qualitative distinction of the vowels to some extent, although native-like quality production may not be attainable. As Widdowson (2003, 2004 and 2005) proposes, each of us should appropriate one's goal of English Language Learning to suit our purpose. Most of Japanese users of English aim to be competent in English as a lingua franca or as an International Language. It is our attitude and belief toward one's English whether we despise our use of English as Inter-language with errors or whether we accept our English as a legitimate variety consisting of international-situation specific utterances.

## 3 References and appendices

- Ano, K. (2002). 「高校生英語学習者の発話における流暢さと正確さの関係」 (A study of the relationship between fluency and accuracy in speech production among high school students), *STEP BULLETIN* Vol.14 pp.39-49.
- Bautista, MA Lourdes S. (2000). *Defining Standard Philippine English: Its Status and Grammatical Features*. De La Salle University Press: Manila.
- Bresnan, J. (2000). "Pidgin Genesis and Optimality Theory". In *Processes of Language Contact*:

- Case Studies from Australia and the Pacific*, edited by Jeff Siegel. Montreal: Les Editions Fides, 145--173.
- Corder, S. Pit. (1980). *Error analysis*. Oxford: Oxford University Press.
- Fujii, K. (1986). *The Basis for Pronunciation of Present-day English: Phonological Comparison of Japanese and English*. Kenkyusha Press: Tokyo.
- Gimson, A.C. (2001). *Gimson's Pronunciation of English*. Oxford University Press: New York.
- Haraguchi, Y. (2006). Acquisition of English Phonemes by Japanese English Learners, Unpublished M.A. thesis, Graduate School of Education, Waseda University.
- Higgins's Minimal Pairs of RP <http://www.marldodge.supanet.com/wordlist/>
- Jenkins, J. (2000). *The Phonology of English as an International Language*. Oxford University Press: Oxford.
- Jenkins, J. (2003). *World Englishes*. Routledge: London.
- Jenkins, J. (2005). English as a Lingua Franca: Past Empirical, Present Controversial, Future Uncertain, Ed. J.A. Foley, *New Directions in the Teaching of Oral Communication*, SEAMEO Regional Language Centre, Singapore. Pp1-19.
- Kirkpatrick, A. (2005). Oral Communication and Intelligibility among ASEAN Speakers of English.
- Ed. J.A. Foley, *New Directions in the Teaching of Oral Communication*, SEAMEO Regional Language Centre, Singapore. Pp33-54.
- Kirkpatrick, A. (2007). *World Englishes: Implication for international communication and English language teaching*. Cambridge: Cambridge University Press.
- Ladefoged, P. (1975). *A Course in Phonetics*. Harcourt Brace Jovanovich: New York.
- Ladefoged, P. (1993). *A Course in Phonetics*. Harcourt Brace & Company: Florida.
- Ladefoged, P. (2001). *A Course in Phonetics*. Fourth Edition. Heinle & Heinle. Boston.
- Lisker, L., Abramson, A.S. (1967) 'Some effects on context on voice onset time in English stops.' *Language and Speech*. 10. pp.1-28.
- Nakano, M. 2000. "An Experimental Study of a Relationship between Lexical Functional Grammar and Learner-Language Data." *Selected Papers from AILA'99 TOKYO*. Waseda University Press, 344-361.
- Nakano, M., Miyasaka, N., Yamazaki, T (2000). A Study of EFL Discourse Using Corpora: An Analysis of Discourse Completion Tasks. *Journal of Pan-Pacific Association of Applied Linguistics* 4(2), 273-297
- Nakano, M., Yamazaki, T., Miyasaka, N. and Saito, T. (1999). A Study of EFL Discourse Using Corpora(7): An Analysis of E-mail Discourse and Variation of Expressions. *Proceedings of the 4th Conference of Pan-Pacific Association of Applied Linguistics*, 55-59.
- Nakano, Michiko (2003) 'Japanese Learners of English are not necessarily polite in English? – some evidence from Discourse Completion Tasks and Naturally Occurring Data.' *The 31<sup>st</sup> Jacet Summer Seminar Proceedings*, Japan Association of College English Teachers, No. 3. pp39-60.
- Nakano, M. (2007a). Japan English. In Asia Englishes and Miscommunication, Eds. Park, K and Nakano, M. Korea University Press, pp 319-428.
- Nakano, M. (Ed.) (2007b). *On-demand Cyber Course Book: World Englishes and Miscommunications*. Waseda University International.
- Ono, S. (1986). *Introduction to English Phonetics*. Liber Press: Tokyo.
- Ringbom, H. (1998). Vocabulary frequencies in advanced learner English: a cross-linguistic approach. In S. Granger (Ed.), *Learner English on Computer*. Longman
- Seidlehofer, B. 2001. Closing a conceptual gap: the case for a description of English as a lingua franca. *International Journal of Applied Linguistics*. 11/2. pp133-158.
- Seidlehofer, B. (2004). Research perspectives on teaching English as a lingua franca. *Annual Review of Applied Linguistics*. Vol. 24, pp. 209-239.
- Widdowson, H. G. 2003. *Defining Issues in English Language Teaching*. Oxford: Oxford University Press.
- Widdowson, H. G. 2004. *Text, Context, Pretext*. Oxford: Oxford University Press.
- Widdowson, H. G. 2005. 'On the subject of English.' Invited Lecture given at Waseda University.

# Assessment of Teaching Practice of an EFL teacher – A Working Model of a Practical Scheme

Taehee Choi

Guryong Middle School

calm91@chol.com

## Abstract

Recently in Korea the media and parents associations have raised criticism that public school education has become a disaster. They claim that public schools do not meet students' needs or care for them, and that parents are forced to pay extra for education and for private tutoring or education abroad. They demand the reformation of education, blaming its failure mostly on teachers, claiming that the majority of public school teachers are inadequate or half-hearted.

In response to this criticism, the government has revised the existing teacher evaluation scheme, called Teacher Promotion Scheme (TPS)<sup>1</sup>, which is a basis for promotion, relocation, awards, and distribution of performance pay. In TPS, the proportion of the score for teaching practice has increased from 12% to 20%. The government has also instituted a new teacher evaluation scheme called Teacher Development Appraisal (TDA). The criteria of TDA are lesson planning, teaching, evaluation and feedback, and satisfaction of students and parents. The details for this appraisal are left flexible to meet the needs of individual schools. TDA was originally planned to be used by the educational offices in different regions for deciding pay and in dealings with individual teachers. However, because of incompleteness of the system and protests from teachers' associations, it is currently being used only for encouraging personal development, thus indirectly improving the quality of education.

Research which was done after the government's intervention concluded that TPS still did not ensure the quality of education. Overlapping between the

two schemes, TPS and TDA, resulted in a waste of resources for the government, the teacher evaluators and the teachers. It was thought that TDA would not improve the quality of education since it was not linked to any rewards or penalties (SNU & KICE, 2008). These results indicated a need for incorporating both evaluation schemes into a rigorous scheme that promotes changes in teacher behavior.

The purpose of this research is to develop tools that constitute a part of that scheme. The objective of the tools is to evaluate teaching practice of a secondary school, English as a Foreign Language (EFL) subject teacher<sup>2</sup>. The scope of the tools, however, is not limited to the area of teaching practice in TPS but, in a sense, covers all the other areas since it will look at teaching practice in terms of the whole curriculum, in which case a teacher's personality, attitude, guidance skills, scholarship & duties are all related to teaching practice. However, research into other areas of TPS is limited to the extent that these other areas affect students' learning of the EFL subject.

## A WORKING MODEL TO ASSESS THE TEACHING PRACTICE OF EFL TEACHERS

The scheme includes both evaluation tools and a standard procedure for evaluation using the tools, which evaluates the teaching practice of a secondary level EFL teacher. The scheme is intended for in-school use and is a practical scheme in that it measures practical knowledge, rather than declarative knowledge, and would require a reasonable amount of time for evaluation.

The tools consist of three components: assessment kits, assessment manuals for the assessor and the assessee, and items for assessing the assessment itself. The tools cover the three phases of the

---

<sup>1</sup> The evaluation is relative and is composed of 5 areas: Personality, Attitude, Teaching Practice, Guidance, and Scholarship & Duties. There are four grades, the percentage for each grade being set: Su, the highest, can be given to 20% of the teachers; Wu, 40%; Mi, 30%; Yang, the lowest, 10%. The result of the appraisal is communicated to the individual only when it is requested.

---

<sup>2</sup> In this research, EFL means learning English as a subject within the school environment. The research is concerned with secondary level which includes both middle and high schools; in Korea teachers move freely between the two levels of schools.

assessment: pre-assessment, assessment, and post-assessment. The composition of the tools was decided after research into teacher effectiveness and teacher assessment research which include Campbell, 2004; Choi, 2006; Hill et al, 2004, 2005; Jeong, 2005; Jin et al, 2001, 2002; Kim et al, 2004; Leshem & Bar-Hama, 2007; Pasko, 1992; . Sinnema, 2005; Yu & Kim, 2005.

The justification of rather complex composition of the tools follows herewith. In Korea, the frequency of teacher assessment/supervision is increasing, especially for EFL subject teachers. However, many supervisors did not receive systematic training for assessing or supervising teaching/teachers. The case is truer for subject supervisors in particular, who are consulted on an irregular basis. On top of that, recruitment from July of the year of 2009 of a vast number of English conversation teachers with various backgrounds entails an excessive workload on supervisors. This will probably lead to commissioning veteran teachers (who did not have professional training for supervision or assessment of teaching/teachers) to supervise or assess those teachers either on a permanent or temporary basis. Thus, standardization of the procedure of EFL teacher assessment, as well as building in measures to improve the reliability of the assessment, is mandatory.

## 1 Methods

As the first step in developing such tools, the researcher examined current practices through an investigation of related documents. The investigation found that TPS is an evaluation based on impressions for the most part. The evaluators<sup>3</sup> ask four questions about the evaluatee: Does the person prepare well for the class? Is the person enthusiastic about their teaching and developing their teaching skills? Does the person reorganize the curriculum to meet the needs of the students and use teaching resources efficiently? Does the person evaluate students well and make use of the results? The principal and the vice principal evaluate evaluatees referring to a one-page, self-reporting form<sup>4</sup>, and the evaluation committee

<sup>3</sup> The evaluators are the principal, the vice principal, and the evaluation committee composed of three teachers who know the teacher best; their evaluation accounts for 40%, 30%, and 30% of the evaluation score, respectively.

<sup>4</sup> The form requests reports on self-set goals regarding three items: duties, performance and innovations. At the bottom there is a table for self-evaluation of four categories; degree of achievement of the goals, creativity, punctuality and

evaluate without asking for any materials to be used in the evaluation. Also, Baek's study (as cited in SNU & KICE, 2008) found that the evaluation is done perfunctorily for most teachers and only the evaluations for those few who are considered for promotion are done properly.

In the case of TDA, in some schools it was based on real data such as video recording of one class and handouts from the class; however, the evaluation was limited to each individual class and did not look at teaching practice in terms of the whole curriculum. These results show the need for valid and reliable teacher evaluation tools that look at teaching practice throughout the curriculum. This will satisfy the recent demands from educational parties, teachers and parents alike<sup>5</sup>.

Also, an analysis of recent studies on teacher evaluation showed that assessment systems are not transferable from one context to another (Lee et al, 2003; Teddlie et al, 2006). Therefore, the prerequisite for the defining work is research into the needs and wants of different educational stakeholders in Korea. The research data include documents produced by the media, the educational offices, the schools, and teachers' associations, etc.

Lastly, the assessment tools are to be piloted to a sample of the target population using multi-stage sampling technique. To determine the validity and reliability of the tools, factor analyses are to be conducted.

## References

- Campbell, R. J. Kyriakides, L., Muijs, D., & Robinson, W. (2004). Assessing teacher effectiveness: developing a differentiated model. New York: RoutledgeFalmer.
- Choi, C. (2006). *A study on a plan to improve the performance appraisal system as compared with the essence of teacher evaluation.*

---

efforts. Evaluatees evaluate themselves using three grades: satisfied, okay, unsatisfied. If needs be, an evaluatee can attach extra pages to the form.

<sup>5</sup> The government has increased the percentage of performance pay to 30% of the whole salary and banned distributing it evenly among teachers, against which teacher associations are protesting. Teacher associations request a transparent and reliable evaluation system as a prerequisite for a pay system tied to performance. On the other hand, several parents associations insist that the entire salary be paid according to performance based on evaluation, and that those who do not qualify be either reeducated or discharged.

- Unpublished master's thesis, Seoul National University of Education, Seoul, Korea.
- Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teacher's mathematics knowledge for teaching. *Elementary School Journal*, 105, 11-30.
- Hill, H. C., Rowan, B., & Ball, D. L. (2005). **Effects of teachers' mathematical knowledge** for teaching on student achievement. *American Educational Research Journal*, 42(2), 371-406.
- Jeong, K. (2005). *외국의 교원 평정 제도* [A case study of teacher evaluation systems of foreign countries]. Seoul: Korea Education Development Institute.
- Jin, K., Kim, M., Yi, W., Brown, A., & McNamara, T. (2001). *초등 영어 교수 능력 신장을 위한 진단 도구 개발 연구* [Measuring an elementary school teacher's ability to teach English]. (Available from Korea Institute of Curriculum & Evaluation, 142 Gahoe-ro, Jongno-gu, Seoul 110-230 Korea).
- Jin, K., Kim, J., Brown, A., & McNamara, T. (2002). *중등 영어 교수 능력 진단 프로그램 개발 연구* [Measuring a middle school teacher's ability to teach English]. (Available from Korea Institute of Curriculum & Evaluation, 142 Gahoe-ro, Jongno-gu, Seoul 110-230 Korea).
- Kim, I., Yu, K., Yi, T., Baek, S., Jeong, K., & Park, S. (2004, December 31). *교사 평가 시스템 연구* [A study on the teacher appraisal system]. Retrieved April 3, 2009 from <http://public.kedi.re.kr/Report/LuPublic/Gz> PubViv.php?PI\_Num0=4392&Pc\_Code=010301&PageNum=1&S\_Part=A00&S\_Menu=&S\_Key=&Pc\_Part=A00&S\_Kind=RR.
- Lee, J.C.K.; Lam, W.P.; Li, Y.Y. (2003). Teacher evaluation and effectiveness in Hong Kong: issues and challenges. *Journal of Personnel Evaluation in Education*, 17(1), 41-65.
- Leshem S., & Bar-Hama, R. (2007). Evaluating teaching practice. *ELT Journal*, 62(3), 257-265.
- Pasko, E. A. (1992). *Handbook for the evaluation of teachers and foreign language programs*. Boston University. (ERIC Document Reproduction Service No. ED376698).
- Sinnema, C. E. L. (2005). *Teacher appraisal: missed opportunities for learning*. Unpublished doctoral dissertation, University of Auckland, New Zealand.
- SNU (Seoul National University) & KEDI (Korea Education Development Institute), (2008). *공교육 정상화를 위한 기초 연구* [Fundamental study for educational reformation]. Seoul: Korea Education Development Institute.
- Teddlie, C., Creemersb, B., Kyriakides, L., Muijs, D., & Yua, F. (2006). The International System for Teacher Observation and Feedback: Evolution of an international study of teacher effectiveness constructs 1. *Educational Research and Evaluation*, 12(6), 561 – 582.
- Yu, K., & Kim, I. (2005, October 10). *교수 평가 시스템 분석 및 향후 개선 방향* [Analysis of teacher appraisal system and ways to improve the system]. *KEDI Position Paper*, 8, Article KP2005-08. Retrieved April 3, 2009, from

[http://public.kedi.re.kr/Report/LuPublic/GzPubViw.php?Pl\\_Num0=5351&Pc\\_Code=010301&PageNum=1&S\\_Part=A00&S\\_Menu=&S\\_Key=&Pc\\_Part=A00&S\\_Kind=KP](http://public.kedi.re.kr/Report/LuPublic/GzPubViw.php?Pl_Num0=5351&Pc_Code=010301&PageNum=1&S_Part=A00&S_Menu=&S_Key=&Pc_Part=A00&S_Kind=KP).

# A Research on Children's Motivation

Eunpyo Lee

<sup>1</sup>School of Medicine, Eulji University

elee@eulji.ac.kr

## Abstract

This study examines how stamp-awards affect children for English class participation, assignment and performance on tests. Two 18-student groups of the similar level children were the subjects and one group was stamped as awards when they did their assignment well and participation was good. The results showed the stamp-awarded group outperformed on the post-test, and close rapport was built between the teacher and the students. As stamps boost children's motivation, teachers need to create a positive motivation-inducing atmosphere so that the young learners participate and do their work voluntarily to learn.

## Keywords

Elementary school education, stamp-awards, motivation

## 1 Introduction

Offering English classes to elementary students has been a controversial, yet demanding issue, and has become part of the curriculum in Korea. The Ministry of Education, Science, & Technology(MOEST) stipulates the objectives of elementary school English as acquiring basic communicative skills and inducing interest along with confidence in students. To meet such demand and fulfill the objectives of English education, the MOEST implemented one-hour-a-week English program for 3<sup>rd</sup> and 4<sup>th</sup> graders and two-hour for 5<sup>th</sup> and 6<sup>th</sup> graders. With the emphasis on the new English curriculum, the on-going English camp for elementary students wins parents' support and still plays an important part in elementary school English.

The compelling fact that America plays a role of super power in world economy makes more nations focus on communicative aspect of the English language. Even in the nations where English is learned as a foreign language, a special priority is given to English education because of its global status in the international society (Graddol, 1997, 2006; MaKay, 2002).

Like any other type of learning, second language

learning requires time, effort, and persistence on the part of learner as Gardner (2001) states whereas his earlier socio-educational model of motivation (1985) suggests that effort, desire, and positive affect, are necessary factors distinguishing successful L2 learners from less successful ones.

Concerning motivation, Chomsky (1988) stresses the importance of activating learners' motivation and claims that "the truth of the matter is that about 99 percent of teaching is making the students feel interested in the material" (p. 181). Also Schumann (1997) asserts in his book, *The Neurobiology of Affect in Language*, that second language acquisition is primarily emotionally driven and emotion underlies most, if not all, cognition.

The primary purpose of the study is to examine L2 learning motivation in an environment in which stamp-awards are given to the students who actively participate in class and/or outperform their homework. It will examine how the students in stamp-awarded group do on the English pre- and post-test and if they are better motivated to do their assignment. It further examines on the degree of interaction between the teacher and the learners of English.

## 2 Literature Review

Arnold and Brown (2005) state that providing an emotionally safe atmosphere is becoming an important issue that neuro-science delivers substantial research into the mechanism for motivation and memory.

An alternative motivational framework has been formulated based on a self-determination theory (Deci & Ryan, 1985, 1995). According to the theory, they types of motivation can be categorized in terms of learners' goal setting for performing an activity: intrinsic and extrinsic motivation and amotivation. Intrinsic motivation, according to Deci & Tyan (1985, 1995), refers to motivation to engage in an activity for itself, and the pleasure, fun, interest, and satisfaction derived from the participation. Intrinsic motivated activity is considered to be fully self-determined without

any external enforcement or rewards (Noels, Pelletier, Clement, & Vallerand, 2000, recited from Kim (2004)).

On the other hand, extrinsic motivated behaviors are controlled by external constraints and less self-determined than intrinsically motivated ones (Deci & Ryan, 1985, 1995; Deci, Vallerand, Pelletier & Ryan, 1991).

Williams and Burden (1997) proposed three stages of motivational process along a continuum in L2 learning. They focused on time and separated stages into 1) reasons for doing something, 2) deciding to do something, and 3) sustaining the effort or persisting.

Motivating learners is a continuing practice evolving from creating the condition and generating an initial motivation to maintaining and protecting it and encouraging positive self-evaluation (Dornyei, 2001a). In formal instructional settings, Cameron (2003) underlines the task of sustaining or sometimes restoring learner motivation over longer periods of time is challenging and effort-taking. Understanding what motivates children in English education will be an important concern and task for the language teachers as well as professionals involved in developing textbooks.

### **3 Method**

#### **3.1 Subjects**

Two 18-student classes of the similar level children based on the placement test were selected among the ones participating in the 2008 summer English camp at H University, sponsored by Daejeon District Office and Education Office. Camp participants were chosen by their homeroom teachers from various elementary schools located in the eastern and western districts of the city.

Students in Class Eagle(E), experimental group, were from 11 schools, 6 males and 12 females, mostly 5<sup>th</sup> and 6<sup>th</sup> graders. This group was stamp-awarded upon active participation and/or satisfactory performance on homework assignment. Students in Class Shark(S), control group, were from 15 schools, 7 males and 11 females, 5<sup>th</sup> and 6<sup>th</sup> grades except for one 4<sup>th</sup> grader, and this group was not stamp-awarded.

Both groups were encouraged to work hard and learn by participating actively. They had four hours of study from Monday to Friday for 3 weeks, 2 hours of native speaking teachers and 2 hours of Korean English (bilingual) teachers each day. The primary language in the classroom was English.

The study was based on the pre- and post-test and the questionnaire to learn how the students progressed, how the stamp-awarded group felt about such awards and if the stamp-awards helped build rapport between the teacher and the students.

#### **3.2 Study Design**

The three steps were involved in the study. On the first day of the program, student orientation was given to the two classes as to how the class was to be conducted, what they were expected to do in class, and how much homework would be given. The pre-test was administered to measure the subjects' background knowledge of animals, weather and season related adjectives, and ability to discern different sounds of the language. The test items included matching the same sound word animals, counting the words that start with the given sound, writing the first letter of weather related adjectives and distinguishing <s> and <sh> sounds of 16 words. The total number of the test questions was 42.

From the second day on, students were taught on the various sounds using the minimal pairs along with the textbook prepared by the District Office. The students were asked to write three words of each sound that was practiced in class. And the ones who did their assignment well were stamp-awarded. Their performance and the progress were measured by comparing pre- and post-test.

#### **3.3 Data Analysis Procedure**

The results of the pre- and post-test of the two classes were analyzed. Also the results of the questionnaires were analyzed to learn of the students' opinion.

### **4 Results and Discussion**

The noticeable difference in the two groups was the class atmosphere. The experimental group was very active to show what they knew and participated with great enthusiasm. Their performance on homework was better than that of the control group. There was no single student who had not done homework. About half of the class did more than asked for, and they were extremely happy to receive stamp-awards. It was evident that most of the students, except for the three shy ones who always sat at the back of the classroom, were eager to show their vocabulary lists of knowledge.

The results of the stamp-awarded group were 39.82 on the post-test showing 1.88

points improvement whereas 35.06 with -2.28 point difference among the non-stamped group. These differences could be explained that the students who did their assignment and got praised with the stamp-awards performed better by remembering what they learned in class. Since the post-test questions were based on the material the students learned during the English camp, the results were likely to decrease if they did not put in effort to their study.

The t-test was done on the improvement of the two groups, and the two-sided p-value was 0.001294. Therefore, the improvement of the stamp-awarded group showed statistical significance as its  $P(T \leq t)$  shows a lot less than 0.05.

Though the overall improvement was not statistically significant, it showed stamped group was better performing and doing their assignment better. The questionnaire answers showed they were favorable toward stamp-awards.

## 5 Conclusion and Limitation

As was already pointed out in Kym's (2008) study, L2 teaching at a primary level should pay more attention to the potential development of the child. Dornyei (1990, 2001a) maintains that children are at an early stage of the language development and their initial experiences in L2 learning and attitudes toward it crucially influence the ultimate achievement of their language learning. Young learners seemed very enthused in getting any form of awards and such an award-giving triggers them to do their work enthusiastically and participate in class more actively.

Graddol (2006) points out that the institutional education is to provide the generic skills needed to acquire new knowledge and specialist skills in the future. English is considered an important and useful skill for modern people.

The findings of the current study indicate that stamp-awards or any kind of awards for that matter provide students with such motivation in English class that students actively participate in the classroom and outperform in their assignment. At the same time, close rapport was built between the teacher and the students.

Since the study is exploratory with a rather small sample size students, the results are not to be generalized. A bigger sample size in a different context might yield a different finding. And the pre- and post-test were not comprehensive to determine students' lexical knowledge. There might have been

unexpected external effects during the period between the two tests.

Having the objectives of English education for the young learners as acquiring basic communicative skills and inducing interest and confidence in students, teachers need to implement ways to motivate young learners in classroom so they can participate more actively and learn efficiently. Along with awarding stamp, praising each student on his/her achievement, progress or performance also gives positive impact on the students.

## References

- Arnold, J., & Brown, D. (2005). A map of the terrain. In J. Arnold. (Eds.) *Affect in language learning*. Cambridge University Press.
- Cameron, L. (2003). Challenges for ELT from the expansion in teaching children. *ELT Journal*, 57(2), 105-112.
- Chomsky, N. (1988). *Language and problems of knowledge*. Cambridge, MA: MIT Press.
- Covington, M. (1999). Caring about learning: The nature and nurturing of subject-matter appreciation. *Educational Psychologist*, 34, 127-136.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.
- Dornyei, Z. (2001). *Teaching and researching motivation*. New York: Longman.
- Dornyei, Z. (1990). Conceptualizing motivation in foreign-language learning. *Language Learning*, 40, 45-78.
- Gardner, R. (1985). *Social psychology and second language learning: The role of attitude and motivation*. London: Edward Arnold.
- Gardner, R. C. & Lambert, W. E. (1959). Motivational variables in second language acquisition. *Canadian Journal of Psychology*, 13, 266-272.
- Graddol, D. (2006). *English next*. London: The British Council.
- Kym, I (2008). Motivational structures among learners with different levels of L2 proficiency: The case of Korean young learners of English. *English Teaching*, 63(1), 101-129.
- McKay, S. (2002). *Teaching English as an international language*. Oxford University Press.
- Murray, B. (2007). Classroom Motivation of Korean EFL Students from the Perspective of Self-determination Theory. *English Teaching*, 62(4), 391-409.
- Oxford, R. & Nyikos, M. (1989). Variables

affecting choice of language learning strategies by university students. *The Modern Language Journal*, 73, 292-300.

- Pintrich, P., & Schunk, D. (2002). *Motivation in education: Theory, research and applications* (2<sup>nd</sup> Edition). Upper Saddle River, N.J.: Merrill Prentice Hall.
- Ryan, R. M., & Deci, E. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25, 54-67.
- Schumann, J. H., & Wood, L. (2004). The neurobiology of motivation. In J. Schumann (Eds.), *The neurobiology of learning: Perspectives from second language acquisition*, Nahwah, NJ: Lawrence Erlbaum Associates.
- Schumann, J. H. (1997). *The Neurobiology of Affect in Language*. Malden, MA: Blackwell.
- Williams, J., & Burden, R. (1997). *Psychology for language teachers*. Cambridge: Cambridge University Press.
- .

# Award-giving as a Means of Motivation and English Grades

Eunpyo Lee

<sup>1</sup>School of Medicine, Eulji University

elee@eulji.ac.kr

## Abstract

This study examines how award-giving affects students' performance in English class by comparing their grades before and after the awards. 49 students were the subjects and those who wrote grammatical, coherent, and thematically-organized essays, and who did their best to present were awarded. The results show that the awarded group improved significantly than the non-awarded. It can be concluded that inducing students' motivation through awards results in better performance and efficient language learning.

## Keywords

Award-giving, intrinsic & extrinsic motivation, English grades

The proceedings are the records of the conference. PAAL hopes to give these by-products a high-quality appearance. We would like authors to make the papers look exactly like this document. The easiest way to achieve this is to replace the content with you own material.

## 1 Introduction

Fluency in English was perceived as a symbol of the so-called exclusive elite group in Korean society in the past several decades. As America plays a leading role in world academia, economy, science and technology, English is considered one of the most important subjects regardless of learners' major. Furthermore, as Koreans expand their capacity outside the nation, carrying out prominent positions like the U.N. secretary-general, it is obvious that English is becoming a world language spoken by more people than ever before.

The need of English for medical students is especially on the rise since English textbooks are used for their study and research papers are required to be written in the language. Competence in the English language is no longer a symbol of social status but a requirement for the modern competitive world called the global village. Moreover in Korea,

English knowledge and skills are tested for college entrance, prestigious high school admittance, job recruitment, graduate school admission, and other fields.

In any kind of education, learners' motivation is what leads them to learn and acquire the skills of learning. In this context, understanding what motivation is, what it does to learners, where it comes from, and how to instill motivation to learners is a very crucial element for teachers.

The issue of motivation is an essential part of our daily lives, and few would overlook the important role it plays in human affairs in general, and in education and language learning in particular (Dornyei, 2001).

In regards to motivation, medical students are already considered highly motivated academic group and their class performance generally excels in the fields of biology, chemistry, physics and other science subjects. However, since their workload, especially of second year and on in medical school is heavy, and study demands extensive time and concentration in the basic medical subjects such as anatomy, embryology, organic chemistry and osteology, non-medical subjects like English are often neglected. Therefore, conducting an English class to the seemingly non-interested group of students may be an energy-depleting task for any language teacher.

Considering unique circumstance of somewhat non-interested group of students whose major is medicine, administering a language class that motivates students should be a pertinent task for the instructors who teach English for ESP. This study aims to examine if award-giving is actually effective for students as extrinsic motivation, how it affects students' performance in English, and how their grades are different from those of the non-awarded ones.

## 2 Literature Review

There have been numerous research studies on motivation. The foundation of motivation theory was laid by Gardner & Lambert (1959) and they

identified two motivational orientations: integrative motivation and instrumental motivation. The learners' desire to identify with members from the L2 community is characterized as the integrative motivation, while the learners' interest in learning L2 for the pragmatic benefit of L2 proficiency is highlighted by the instrumental motivation (Dornyei, 1990).

Oxford & Nyikos (1989) have shown that motivation directly influences the use of L2 strategies. It also influences the higher level of achievement in language learning (Clement, Major, Gardner, & Smythe, 1977; Gardner, 1985). Motivation is identified primarily with the individual person's orientation toward the goal of second language learning within Gardner's framework, known as the Socio-Educational Model (Gardner, 1985).

The word "motivation" is derived from the Latin verb *move*, which suggests the idea of movement or a drive to complete a particular task (Pintrich & Schunk, 2002). Brophy (1999) refers to motivation as a theoretical concept used to describe the initiation, direction, intensity, and persistence of some activity, especially if that activity is goal-directed.

Ryan and Deci (2000) describe motivation as an impulse or inspiration to do something and to complete the task. The concept of motivation was outlined by Deci and Ryan's (1985) Intrinsic Motivation and Self-Determination in Human Behavior. Their intrinsic-extrinsic motivation explains that human motivation can be seen to exist on a six-point continuum, from amotivation that is non-internalized, non-regulation through 4 categories of extrinsic motivation of external, introjected, identified, and integrated regulation, to intrinsic motivation that is fully internalized, intrinsic regulation.

Intrinsic motivation is defined as "the doing of an activity, for its inherent satisfactions rather than for some separable reason" (Ryan & Deci, 2000). They further elaborate that "When intrinsically motivated, a person is moved to act for the fun or challenge entailed rather than because of external prods, pressures or reward." On the other hand, extrinsic motivation is explained with four natural process where humans actively seek to transform extrinsic or external regulation into a more internalized type of self-regulation by Ryan (1995).

### 3 Method

#### 3.1 Subjects

The second year medical students in English Reading and Writing were the subjects. The class comprised 20 female(41%) and 29 male(59%), and

they met twice a week for 2 hours each for 16 weeks. English reading was emphasized particularly with acquisition of college level vocabulary including commonly used medical terminology. Along with the textbook articles, additional reading material related to new medical procedures, extraordinary reconstructive surgeries, and other life and death related articles were supplemented to read for class discussion and writing.

Since reading was considered a passive activity and each class met immediately after lunch, it was inevitable for students to fall asleep, an understandable physiological phenomenon. So unless the material was interesting enough for the students to keep awake and/or students were demanded to participate in class presentation, they were likely to remain asleep. Consequently, the students asked to write an essay or article after each reading material either related to the textbook or supplementary handouts. Such writing requirement was one of the reasons why these students were not very favorable in terms of the course evaluation and the instructor stating the class was too demanding with much writing work.

#### 3.2 Procedures

For students' class presentation, over 50 topics were selected and each student picked one topic and prepared class presentation. Also other written assignment was given such as writing a paragraph-long description and/or answers to the textbook questions, and short essay. Their assignment was collected, reviewed, and given back with the teacher feedback.

The award winners were selected based on originality, creativity, and grammaticality of the written context. Other technical aspect such as great Power Point, interesting video clips with the ability to draw audience's attention was taken into consideration.

English final grades were compared between the awarded and non-awarded groups.

### 4 Results and Discussion

The total twelve students were awarded, 6 females and 6 males. But two of them were repeaters, one returned from two and a half year's leave of absence, and the other a year. They were excluded from the analysis because their last grades were not retraced.

It showed no gender difference in award winners but the proportion of the female awardees was slightly higher by 10%(female 30%, male 20%). They showed positive response toward the awards

and the medical students were motivated by the awards.

Ten awarded students' post-grades were 92.7 showing 4.9 points improvement whereas non-awarded students' post-grades 81.45 with only 1.09 points improvement. Among the total 33 students, ten students (33%) showed points drop, one particular student showed 14 points drop from the previous class, from B+ to C.

The comparison of the two groups shows the t-value of 1.733275 and two-sided p-value of 0.090563. As the p-value is greater than 0.05, it may not be considered of statistical significance due to the small sample size of the awarded group. However, the improvement of 4.9 points can be interpreted as significant difference while the non-awarded group showed only 1.09 points improvement.

The awarded students' feedback about the awards revealed positive and motivational. They stated that it boosted their pride, made them write better and pay closer attention in class. Korean students rarely had a chance to receive an English award. Therefore, award-giving provided them with great pleasure, and the awardees were envied by their non-awarded peers.

## 5 Conclusion

Student motivation is an important factor for any kind of learning and it plays especially an essential role in language learning. Without motivation, learners easily shut off their interest and neglect engrossing in learning. It is one of the English teachers' tasks to stimulate students to actively engage in classroom language learning.

Although medical students are a highly motivated group, their performance in the language classroom depends on the degree of how motivated they are in the class. For the students who learn the language as ESP, providing an environment in which they feel the desire to learn and participate is a very essential task for language teachers especially the ones who teach English for non-English majors.

Conducting an ideal reading and writing class that invites all the students to the maximum level of participation and concentration may not be feasible. However, understanding what can be done for students' extrinsic motivation and what impact it can have on their class performance and further language acquisition is a primary step for better teaching English so that the learners

are fully equipped with a necessary tool for their future career.

Even for the highly motivated medical students, award-giving acts as an important means to induce motivation for studying English harder and affects their performance positively and in fact, shows better grades at the end of the semester.

However, the present study has some limitations. The sample size was too small to make any generalization from the study. And the number of the awardees was only ten (as two repeaters were excluded from the study) and the comparison of the two groups did not provide statistically significant results. The study was done on a homogeneous group of students who were already highly motivated in learning and being engaged in the academic activities. Further study on a larger sample of both experimental and control groups may render more inclusive conclusion.

## References

- Brophy, J. (1999). *Research on motivation in education: Theory, research and applications (2<sup>nd</sup> Edition)*. *Language Research*. In T. C. Urdan (Ed.), *The context* (pp. 1-44). Stanford, CT: JAI Press.
- Clement, R., Major, L., Gardner, R. C., & Smythe, P. C. (1977). Attitudes and motivation in second language acquisition: An investigation of Ontario francophones. *Working Papers on Bilingualism*, 12, 1-20.
- Dornyei, Z. (2001). *Teaching and researching motivation*. New York: Longman.
- Dornyei, Z. (1990). Conceptualizing motivation in foreign-language learning. *Language Learning*, 40, 45-78.
- Gardner, R. (1985). *Social psychology and second language learning: The role of attitude and motivation*. London: Edward Arnold.
- Gardner, R. C. & Lambert, W. E. (1959). Motivational variables in second language acquisition. *Canadian Journal of Psychology*, 13, 266-272.
- Murray, B. (2007). Classroom Motivation of Korean EFL Students from the Perspective of Self-determination Theory. *English Teaching*, 62(4), 391-409.
- Oxford, R. & Nyikos, M. (1989). Variables affecting choice of language learning strategies by university students. *The Modern Language Journal*, 73, 292-300.
- Pintrich, P., & Schunk, D. (2002). *Motivation in education: Theory, research and*

- applications* (2<sup>nd</sup> Edition). Upper Saddle River, N.J.: Merrill Prentice Hall.
- Ryan, R. M., & Deci, E. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25, 54-67.
- Ryan, R. M. (1995). Psychological needs and the facilitation of integrative processes. *Journal of Personality*, 63, 397-427.
- Van Lier, L. (1996). *Interaction in the language classroom*. New York: Longman

# Durational Analysis of the Acoustical Characteristics of English Speech by Japanese Learners based on the Contrast between the Stressed and the Unstressed

Shizuka Nakamura

Graduate School of Global Information and Telecommunication Studies, Waseda University

shizuka@akane.waseda.jp

## Abstract

In this paper, we analyzed the durational differences between learners and native speakers in various speech units from the perspective of that the contrast between the stressed and the unstressed is one of the most important features to characterize stress-timing of English by comparison with mora-timing of Japanese. The results showed that the lengthening and shortening of learner speech were not enough to convey the difference between the stressed and the unstressed. Finally, it was confirmed that these durational differences strongly affected the subjective evaluation scores given by English language teachers.

## Keywords

rhythm control, mora-timing, stress-timing, syllable, Japanese, English

## Introduction

We have studied the objective evaluation of English rhythm control by learners whose native language is Japanese (called “Japanese learners” or “learners” in this paper), and constructed an objective evaluation model which could simulate the process of evaluating rhythm control of learners by English language teachers (called “evaluators” in this paper) [1-4].

Durational information, which, among the acoustical features, can be calculated automatically in the relatively simple way, has been used for objective evaluation. Durational differences between native speakers and learners are correlated with subjective evaluation scores. Speech units, in which durational differences show strong correlations, have been selected from the phonetic viewpoint as parameters for the objective evaluation model. Such speech units included the phoneme, the vowel (the stressed/unstressed vowel, the vowel in the content/function word), the consonant (the voiced/unvoiced consonant, the consonant in the content/function word), the

syllable (the strong/weak syllable, the closed/open syllable, the syllable in the content/function word), the word (the content/function word), the sentence, and the pause.

Speech samples were selected from the speech database of second language learners [5]. We used 480 samples uttered by a large number of Japanese learners (127 subjects), and these were compared to 215 samples uttered by English native speakers (21 subjects). Five text groups of four English sentences, 20 in total, were used to construct the speech materials. Each of the text groups consisted of four text lengths, namely, VS (Very Short), S (Short), L (Long), and VL (Very Long); these were edited by simply adding new phrases/clauses/words. Examples of the text are shown in Table 1.

In the process of these previous studies, we noticed that the ingrained mora-timing of Japanese learners might unfavourably affect the stress-timing of English speech. In this paper, we analyzed the durational differences between Japanese learners and native speakers in various speech units from the perspective of that the contrast between the stressed and the unstressed is one of the most important features to characterize stress-timing by comparison with mora-timing.

The purpose of this study is to analyze the

Table 1: Sample texts of VS (Very Short), S (Short), L (Long) and VL (Very Long) varieties in text group A, and VL in group B, C and D. (‘: primary stress, ,: secondary stress, .: syllable boundary, and /: phrase boundary)

[Group, Text length] Text
[A, VS (Very Short)] 'Thank you.
[A, S (Short)] ,Thank you ,ver.y 'much.
[A, L (Long)] ,Thank you ,ver.y 'much / for 'eve.ry.thing.
[A, VL (Very Long)] ,Thank you ,ver.y 'much / for 'eve.ry.thing / that you 'did for us.
[B, VL] I'm a.'mused / by the 'man / and his ,ver.y ,fun.ny 'jokes.
[C, VL] Why won't you 'wait / un.,til 'Fri.day / when he's 'back ?
[D, VL] I was ,ter.ri.bly an.'noyed / with the 'man / for ,beat.ing the 'dog.

properties in detail of the speech units used for the objective evaluation model from the phonetic viewpoint. An increase of correlation coefficients between durational differences and subjective evaluation scores attributed this study can also increase an accuracy of a prediction. Furthermore, the results of this study are expected to be useful for the study to specify the effects of mora-timing in the future.

## 1 Durational characteristics of learner speech

In this chapter, in order to analyze the actual conditions of learner speech, the durational characteristics in speech units—sentences, pauses and syllables—are examined by comparing them with the corresponding durations of native speech.

### 1.1 Sentence duration

First, sentence durations which consist of all speech units are analyzed. For English texts being used as evaluation material, only the VL texts of each text group in Table 1 were used, since it was statistically confirmed that a similar or higher accuracy of analysis could be obtained by this reduced set of material texts.

Although variations by text are observed, the average sentence duration of learners is about 1.5 times that of native speakers. In addition, the individual variety of learners extends to three times as wide as that of native speakers. In the following sections, in such a long sentence duration uttered by learners, the cause of different rhythm control by learners from that by native speakers is analyzed by dividing the sentence duration into the various speech units.

### 1.2 Pause duration

In this section, the sentence durations are separated into pause and non-pause durations in order to analyze the ratio of each to the overall sentence duration. While pauses are exceptionally inserted into native speech, the sum of the pause durations in the learner speech was found to increase in proportion with the increase in the text length.

On one hand, the percentage of pause durations in the VS (Very Short) texts is low; on the other hand, that in VL (Very Long) texts is around 40%. On the contrary, the percentage of non-pause durations is not variable unlike that in the case of the pause durations.

### 1.3 Syllable duration

The syllable duration of English native speakers, as required by the place and the manner of articulation

of the preceding consonant, can be lengthened and shortened depending on factors such as whether the syllable is stressed or not, whether it is part of a content or function word, its position within a word or its context in a sentence. However, the innate mora-timing of Japanese learners causes them to have difficulty in lengthening or shortening the syllable durations.

First, the syllable durations of learners and that of native speakers are compared to analyze the influences on the durational difference. The syllable durations of learners in general are longer than those of native speakers.

Next, the phrase durations of all speakers are normalized by that of the typical native speaker to compare the ratios of the syllable duration to that of the phrase duration. An example of normalized syllable duration is shown in Figure 1—for the phrase “for everything” in a VL (Very Long) text in text group A, “Thank you very much for everything that you did for us.” The learner characteristics of lengthening the first syllable “for” can be observed more clearly after normalization in the areas encircled by dotted line.

This observation may be affected by the contrast between the stressed and unstressed syllables, and by the content and function words. Such learner characteristics of rhythm control, which differ from the native styles, are analyzed in the next chapter.

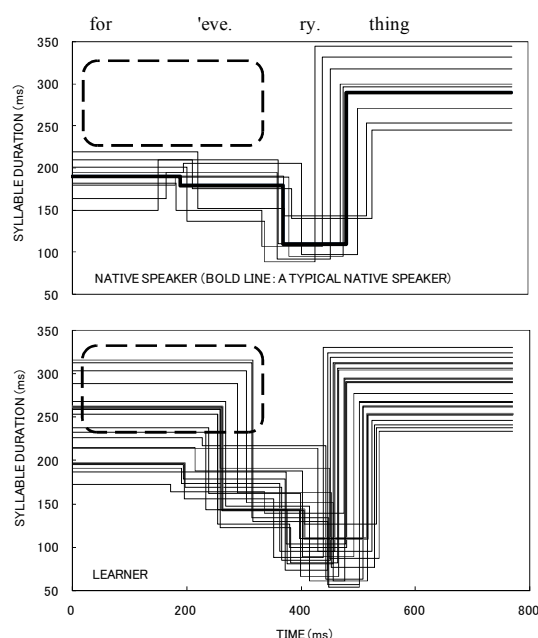


Figure 1: Syllable durations are plotted alongside time for the phrase “for everything,” uttered by native speakers (top figure) and learners (bottom figure). The phrase durations are normalized by that of the typical native speaker (bold line).

## 2 The contrast between the stressed and the unstressed

Compared to the mora-timing of Japanese, one of the most important features to characterize stress-timing of English is the contrast between the stressed and the unstressed. In the following section, the characteristics of learner speech are analyzed by comparing them to the native style from the perspective of the contrast between the stressed and the unstressed.

### 2.1 Stressed/unstressed syllable

Considering the fact that a mora tends to be uttered for just as long as other morae in Japanese, it can be thought that shortened unstressed syllables help to make equal intervals between stresses in English. It is expected therefore that the difficulty in adjusting these durations can be shown in the stressed/unstressed syllable durations and the relationship.

The contrast between the average durations of the stressed and unstressed syllables is shown in Figure 2. The numbers of the stressed and unstressed syllables involved in the four VL texts are 19 and 34, respectively. The durations of the stressed syllables uttered by native speakers are twice as long as those of the unstressed ones, and they gather in the small solid ellipse drawn with the average as its center and an axis along the solid line; on the other hand, there is no significant tendency among those of learners, and they expand in the large dotted circle on the dotted line with the

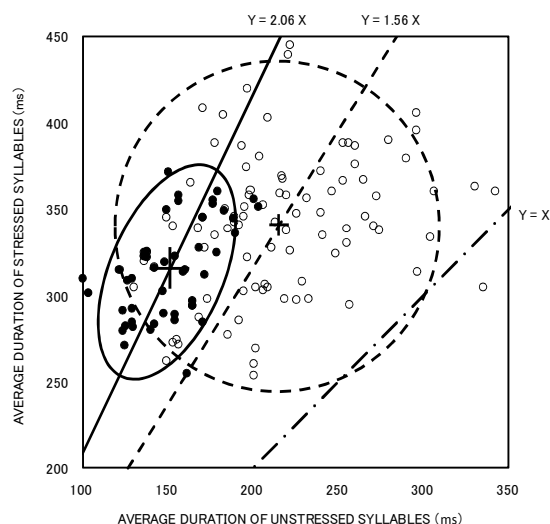


Figure 2: Relationship between the intra-speaker average durations of stressed and unstressed syllables, uttered by native speakers (closed circle) and learners (open circle) in each of the four VL (Very Long) texts. + marks indicate the averages of native speakers (large) and learners (small). The lines (solid: native speakers, and dotted: learners) are drawn from the origin through their averages.

ratio of the stressed to unstressed syllable duration of about one and a half.

### 2.2 Strong/weak vowel

The durational differences between the stressed and the unstressed syllables in the previous section are affected, primarily, by not the consonants but the vowels, which lengthen and shorten more easily. In this section, the vowels in stressed and unstressed syllables, that is, the strong and weak vowels, are analyzed to study the results spelt out in the last section in greater detail.

The inter-speaker average sentence durations and the vowel durations were analyzed. The difference between the strong and the weak vowels uttered by native speakers was approximately 100 ms. On the contrary, the difference among them by learners was only about 50 ms. This result, that is, the degree of the decrease of learner speech was smaller than that of native one, indicates that lengthening and shortening of learner speech were not enough to convey the difference between strong and weak vowels.

### 2.3 Syllables in content/function words

These contrasts between the stressed and unstressed syllables and vowel durations are also affected by the differences between content and function words, since the vowel in content rather than in function words tends to be stressed. In this section, the vowel durations, which greatly influence syllable durations in content and function words, are analyzed. The number of vowels involved in the content and function words in the four VL (Very Long) texts are 29 and 27, respectively.

The ratios of the intra-speaker average vowel duration in content words to that in function words in native speech gathered around two, while those in learner speech expanded from two to one. In order to analyze how these ratios were consistent among the learners, both values were divided by the average vowel duration of the native speakers to normalize the differences across different texts.

As shown in Figure 3, the data on native speakers gathers in a small solid ellipse with the axis along the function to content ratio of about one; in comparison, the data on learners expands over a large dotted ellipse with the axis along the ratio of about one and a half. The ratios deviate from the native data up to almost two. These results show that the shortening of the vowel duration by learners are not enough to convey the contrast between content and function words, in addition, the deviation from the native data are not consistent among the learners.

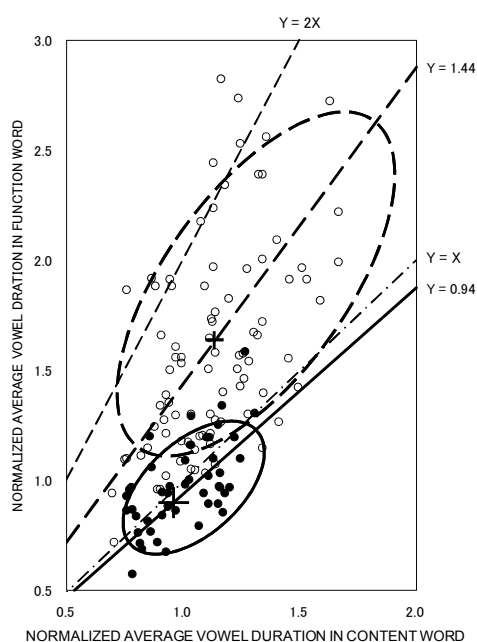


Figure 3: Relationship between the intra-speaker average durations of vowels in content words and function words uttered by native speakers (closed circle) and learners (open circle). The values are normalized by dividing by their respective native average durations of each text. + marks and the lines are drawn in the same way in Figure 2.

### 3 Correlations between durational differences with subjective evaluation scores

In this chapter, we shall study the relationship between the unnatural speech durations of learners and their subjective evaluation by evaluators based on the correlation coefficients between durational differences and subjective evaluation scores.

The durational differences were the RMS scores calculated by using the corresponding typical native speech taken in each speech unit by texts. If the durational differences affect subjective evaluation, the former should have negative correlation with the subjective evaluation scores, whereby smaller scores would indicate a greater degree of unnaturalness.

As a result of the correlation analysis, it was confirmed that durational differences strongly affected subjective evaluation scores. Although variation by texts was observed, for the speech units referred to in the last chapter, the results showed that it was the unstressed (correlation coefficients: -0.49) rather than the stressed (-0.18) syllable, the weak (-0.58) rather than strong (-0.18) vowel, and vowels in function (-0.56) rather than content (-0.31) words had the stronger correlation with the subjective evaluation scores.

## 4 Conclusions

In this paper, we analyzed durational differences between learners and native speakers in various speech units from the perspective of the contrast between stressed and unstressed syllables, devoting attention to the fact that stress-timing of English speech is unfavourably affected by the ingrained was not enough to convey the difference between the stressed and the unstressed. Furthermore, it was confirmed that durational differences between learners and native speakers strongly affected the subjective evaluation scores.

Since the characteristics of English rhythm control in speech units such as stressed/unstressed syllable, strong/weak vowel and vowel in content/function word by Japanese learners should be effective in the evaluation of learner proficiency at rhythm control, the results of this study is expected to be taken in the objective evaluation model.

## Acknowledgements

This study was supported in part by the Grant-in-Aid for Research fellow No. 21 · 06162 and for Scientific Research (B) No. 20300069, Japan Society for the Promotion of Science. We express our thanks to Prof. Y. Sagisaka (Waseda University), H. Kato (NICT), Prof. M. Tsuzaki (Kyoto City University of Arts) and Dr. S. Matsuda (NICT) for their cooperation in this study, and Prof. S. Hiki (Waseda University) for his advice.

## References

- [1] M. Muto et al. (2003). Corpus-based modeling of naturalness estimation in timing control for non-native speech. *Proceedings of EUROSPEECH*, 498-501.
- [2] S. Nakamura et al. (2007). Tempo-normalized measurement and test set dependency in objective evaluation of English learners' timing characteristics. *Proceedings of International Congress of Phonetic Science (ICPhS)*, 1733-1736.
- [3] S. Nakamura et al. (2009). Objective evaluation of English learners' timing control based on a measure reflecting perceptual characteristics. *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 4837-4840.
- [4] S. Nakamura et al. (2009). Effects of mora-timing in English rhythm control by Japanese learners. *Proceedings of INTERSPEECH*, (in press).
- [5] N. Minematsu et al. (2004). Development of English speech database read by Japanese to support CALL research. *Proceedings of International Congress on Acoustics (ICA)*, 557-560.

# The relation between familiarity rating and productive knowledge of academic words

Toru Nakamura

Tokyo Metropolitan Aoyama High School

shobota@gmail.com

## Abstract

In this study, we focus on the relation between familiarity rating and productive knowledge of academic words. The research questions are as follows.

- 1) Is there any relation between the frequency of occurrence of the academic words and the number and the variety of associations written by the participants?
- 2) Is there any relation between familiarity ratings and the number and the variety of associations written by the participants?
- 3) Are the word associations with the low familiarity rating words similar to stimulus words in spelling and sound, whereas are the associations with the high familiarity rating words similar in meaning?

Results suggest that there is a close relation between the familiarity ratings and the number and the variety of response words written by the participants, namely participants produced more responses to high familiarity rating words than to low familiarity rating words. In addition, associations with the low familiarity rating words were similar to stimulus words in spelling only. However, associations with high familiarity rating words were similar in meaning, such as synonyms.

## 1 Background to the current research

### 1.1 What is involved in knowing a word?

In the last 30 years, there has been a growing interest in vocabulary learning and teaching. At the same time, researchers have become aware that there is much more to knowing a word than just learning its meaning and form. Richards (1976) made the first attempt to list the different aspects of knowledge that are necessary to fully know a word. Nation (2001) developed this list further and covered all the aspects of various types of knowledge that one must possess. His model is very

attractive because he distinguished productive knowledge of words from receptive one in his components. According to his assumption, the word to use requires extended knowledge beyond what you need just to understand it. That is, the number of words native speakers can recognize and understand is rather larger than the number they use in their own speech and writing. L2 learners also have different mastery of the various kinds of word knowledge (Schmitt N. and Meara P., 1997).

### 1.2 Modern tests based on corpora

Nation's framework is purely descriptive and cannot explain either the processes of acquisition for the different kinds of word knowledge or the mechanisms by which they interrelate. Lexical knowledge is not a directly accessible quality. Therefore, some researchers designed the tests which use Nation's definition of knowing a word such as *X-Lex* (Meara and Milton, 2006) and the Levels Test (Nation, 2001). *X-Lex* samples the 5,000 most frequent words of English, and then draws 20 words from each of the five 1,000-word frequency bands within this list and uses this to make an estimate of the number of words known out of these 5,000 words. Like *X-Lex*, the Levels Test samples a range of frequency bands in order to measure learners' overall vocabulary knowledge. Eighteen words are tested at each of the 1,000 word frequency bands. It is similar to *X-Lex* in that it tests vocabulary breadth and takes advantage of frequency information in its design. One of the advantages of these tests is that they can be tested empirically, that is, they tend to be relatively quick to administer and mark, and produce a numeral score. As a consequence of this, they were used to make judgments about the language level of non-native speakers.

However, there are three factors that are potential threats to the validity of these tests. First, they calculate vocabulary size and are based on the assumption that the more frequently a word is in a language then the more easily, and earlier, it is

likely to be learned. Indeed, a typical L1 learner's knowledge is high in the frequent words and lower in the less frequent words. Nevertheless, we need to deal with this assumption with extreme caution because frequency information drawn from native speakers may not be relevant to L2 learners who are not exposed to the target language except textbooks (Nakamura, 2007). Second, these tests focus on receptive rather than productive vocabulary because it is difficult to measure the productive knowledge effectively. *X-Lex* is receptive since it is a Yes/No or checklist test which requires learners merely to say if they recognize a target word as a word. For the Levels Test, Laufer and Nation (1999) themselves describe the blank-filling version as a measure of active or productive vocabulary knowledge. However, they have limited evidence to support this interpretation. Thus, it may simply be an alternative way of asserting receptive knowledge rather than a measure of productive ability (Reed, 2000). Third, these tests did not explore the interrelationship among the various types of word knowledge and the results of these tests have been isolated. Furthermore, few researches have explicitly explored the interrelationships among the various types of word knowledge. This is unfortunate, because a better understanding of these interrelationships enables us to explain the model of L2 vocabulary acquisition.

### 1.3 The design of this study

This study examines how two types of word knowledge, familiarity and word associations, interrelate with each other. As this is one of the first pilot studies exploring the interrelationship among the various types of word knowledge, it was impossible to design a study that could capture all of the word knowledge categories Nation (2001) pointed out. Therefore, it was decided to limit the current study to two types of word knowledge.

The framework of the familiarity rating follows Familiarity Ratings among Japanese learners of English (J-EFL) by Yokokawa (2006). J-EFL has the following features. 1) 3,000 words were compiled from JACET 8000. 2) The scores are rated on seven-point scale. 3) It does not intend to ask knowledge of words.

Word associations are the links that connect or relate words in some manner in a person's mind. A common way of drawing out them is to have a tester give a stimulus word and have the participant write the first word that comes to mind. Meara (1983) gave an outline of the work done on word association and concluded that L2 learners tend to give more varied responses than native speakers, even though their vocabulary size is smaller. He

also noted that learners tend to produce non-related but similar-sounding words, instead of the semantically related responses that adult native speakers typically produce. Overall, L2 associations are often very different from those of native speakers. We used the design of the *Lex30* (Meara and Fitzpatrick, 2000) as the test of productive knowledge of vocabulary. The *Lex30* is a word association task, in which participants are presented with a list of stimulus words, and required to produce responses to the stimuli.

### 1.4 Research questions

In the present study, the following research questions were identified:

- 1) Is there any relation between the frequency of occurrence of the academic words and the number and the variety of response words written by the participants?
- 2) Is there any relation between familiarity ratings and the number and the variety of response words written by the participants?
- 3) Are the word associations with the low familiarity rating words similar to stimulus words in spelling and sound, whereas are the associations with the high familiarity rating words similar in meaning?

## 2 Experiment

### 2.1 Participants and Instruments

There were 34 participants. They were students of a high school in Tokyo, Japan; all of these were from Grade 12. All participants had Japanese as their L1. All had not lived in English-speaking countries during childhood. All stimulus words were from Academic Word List (Coxhead, 2000). The word list is divided into ten sublists based on the frequency of occurrence of the words in the Academic Corpus. We used sublist 1 (60 more frequent words) and sublist 10 (30 less frequent words) in this study.

### 2.2 Procedure

First, the participants were asked to rate on a 7-point scale how familiar they regarded the academic words. The participants were asked to rate words as to the number of times they had experienced them, with 1 indicating that the participant had never seen, heard, or used the word in his or her life, 4 indicating that he or she had seen it before, but is not sure of the meaning and 7 indicating that he or she had seen, heard, or used it nearly every day of his or her life. Second, the

participants were asked to write a series of response words at least three if possible for each stimulus word, using word association. Participants had 30 seconds to write responses to each stimulus word. The participants were shown examples of how to complete the test before beginning.

After the experiment, response words were classified into three categories, that is, Form, Meaning, and Use according to Nation's framework (2001). After that these categories were classified into more detail eight categories (Table 1).

Table 1: Eight Categories of Word Associations

Form
A similar in spelling (persist-insist)
B part of the stimulus word (likewise-like)
C derivative and inflection (reluctant-reluctance)
Meaning
D synonym given by Concise Oxford Thesaurus (odd-strange)
E not synonym but similar in meaning (depress-sad)
F same semantic networks (encounter-accident)
Use
G collocation 1 (number-large)
H collocation 2 (reluctant-to)

*Note:* (stimulus word-Response word)

### 3 Results and Discussion

First, the averages of the familiarity ratings for the stimulus words (F) were compared. Surprisingly, there was no significant difference between sublist 1 and sublist 10.

Sublist 1	(more frequent words)	F=4.89
Sublist 10	(less frequent words)	F=4.37

#### 3.1 Results with respect to research question 1

Then, the averages of the total number of associations (X) with each stimulus were compared. In this comparison, every association is counted.

Sublist 1	(more frequent words)	X=43.22
Sublist 10	(less frequent words)	X=24.70

The participants produced more associations with frequent words (Sublist 1) than with infrequent words (Sublist 10).

In addition, the averages of the number of

different associations (Y) with each stimulus were compared. In this comparison, if we see the same association, we do not count it again.

Sublist 1	(more frequent words)	Y=28.62
Sublist 10	(less frequent words)	Y=17.20

The participants also produced more varied associations with frequent words (Sublist 1) than with infrequent words (Sublist 10). However, the ratio between X and Y ( $Z = Y/X$ ) was not different significantly.

Sublist 1	(more frequent words)	Z=0.69
Sublist 10	(less frequent words)	Z=0.73

#### 3.2 Results with respect to research question 2

The results of the analyses are shown in Table 2 and 3. The Pearson correlation between the three variables (familiarity (F) and the number of associations (X) (Y) ) appears to be fairly high for sublist 1 and sublist 10. It indicates that as familiarity with stimulus word increases, participants easily produce more and varied associations.

Table 2: Correlation between familiarity and associations (Sublist 1)

	F	X	Y
F	1	.834	.729
X	.834	1	.802
Y	.729	.802	1

*Note:* All coefficients are significant at  $p < .01$ .

F=familiarity rating of stimulus word

X=averages of the total number of associations

Y=averages of the number of different associations

Table 3: Correlation between familiarity and associations (Sublist 10)

	F	X	Y
F	1	.892	.818
X	.892	1	.876
Y	.818	.876	1

*Note:* All coefficients are significant at  $p < .01$ .

F=familiarity rating of stimulus word

X=averages of the total number of associations

Y=averages of the number of different associations

#### 3.3 Results with respect to research question 3

The part of the results of the experiment can be seen in **Appendix**. In the paper-based test, it is possible for participants to produce infrequent words randomly. Therefore, associations produced more than two participants were counted. For each stimulus word, the participants produced more and varied associations with familiar words (convince:

F=6.59, X=33, Y=25) than with unfamiliar words (adjacent: F=1.81, X=16, Y=7).

While associations with the low familiarity rating words (levy: F=2.04) were similar to stimulus words in spelling only (levy-envy) and the number of associations were very small (levy: X=9, Y=6), associations with high familiarity rating words (encounter: F=6.3) were similar in meaning such as synonyms (encounter-meet), and the number of associations were very large (encounter: X=50, Y=31).

Summarizing the answers for the research questions, we have found that the frequency effect on vocabulary learning is very strong. In addition, participants produced more responses to the high familiarity rating words than to the low familiarity rating words. While they tend to produce non-related but similar-sounding associations with the low familiarity rating words, they produced the semantically related associations with the high familiarity rating words.

#### 4 Directions for Future Research

There appears to be a strong relation between the familiarity ratings and the number and the variety of response words written by the participants. As for the quality of associations, we arrived at the same conclusion as Meara's (1983) that "clang associations" are likely to result from a lack of familiarity with the stimulus words. However, if we investigate the features of two tasks used in this study more closely, we can find another possible explanation for these results.

The first problem is that there are a number of issues concerning the validity of the test used in this study and its results. Familiarity is not the only factor which can influence the number of associations produced by participants. In a familiarity test, the participants' task is closer to self-assessment than to a real language task. The response bias is in fact caused by the element of self-assessment that is part of the test format. In addition, the method we used was based on written forms of knowledge and only predicts writing-based language ability. The tasks should have focused on actual language use. Therefore phonologically based test and the interviews needed to assess face validity or content validity.

A second problem is that the results were influenced by the learner attitude. It is clear that the computer-based test has more control over the test situation. Through computer programming, time limits per item can be built in, the possibility of omitted responses can be ruled out, test-takers can be denied an overview of the test, etc. All these factors may affect the individual's test performance.

Even in the paper-based test, it is useful to use a particular correction formula in order to deal with the response bias issue.

#### 5 References

- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly* 34, 213-238.
- Laufer, B. and P. Nation. (1999). A vocabulary-size test of controlled productive ability. *Language Testing* 16 /1, 33-51.
- Meara, P. (1983). Word associations in a foreign language, *Nottingham Linguistic Circular* 11, 29-38.
- Meara, P. and Fitzpatrick. T. (2000). Lex 30: an improved method of assessing productive vocabulary in an L2. *System* 28, 19-30.
- Meara, P. and Milton, J. (2006). *X\_Lex: The Swansa Levels Test*. Newbury, UK: Express Publishing.
- Nation, I.S.P. (2001). *Learning Vocabulary in Another Language*. Cambridge University Press.
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University.
- Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quarterly* 10, 77-89.
- Schmitt, N. and Meara, P. (1997). Researching Vocabulary Through a Word Knowledge Framework: word associations and verbal suffixes. *Studies in Second Language Acquisition* 19, 17-36.
- Nakamura, T. (2007). Conceptual Representation in Word Recognition. *STEP BULLETIN* vol.19, 23-36.
- Yokokawa, H. (2006). *Familiarity Ratings among Japanese learners of English*. Kuroshio Publishing Company.

Appendix A (Sublist 1) N=the number of participants

		F	X	Y	Z	1	N		2	N		3	N		4	N	
1	analyze	4.24	48	34	0.71	data	3	H	study	3	D	investigate	5	D	examine	2	D
2	approach	5.97	66	41	0.62	goal	2	F	come	3	E	near	8	E	tennis	2	F
3	area	6.53	69	41	0.59	country	4	E	teritory	2	D	place	14	D	field	3	D
4	assess	2.79	21	21	1.00												
5	assume	5.35	39	27	0.69	think	13	D									
6	authority	3.88	41	30	0.73	author	5	B	power	5	D	book	3	F	teacher	2	F
7	available	4.38	45	33	0.73	can	5	E	get	4	E	use	4	E	easy	2	E
8	benefit	5.24	50	31	0.62	company	4	F	profit	6	D	money	10	F	good	2	D
9	concept	5.59	53	41	0.77	policy	2	E	notion	3	D	thought	4	E	image	2	D
10	consist	5.71	39	31	0.79	compose	5	E	insist	2	A	make	2	E	people	2	F
11	constitute	4.35	24	17	0.71	constitution	4	C	institute	2	A	America	2	F	consist	2	A
12	context	3.41	20	12	0.60	book	5	F	back	2	F	word	2	F	sentence	3	F
13	contract	4.18	28	24	0.86	paper	2	?	contrast	3	A	color	2	?			
14	create	6.06	63	35	0.56	art	6	F	make	18	D	music	3	F	creater	2	C
15	data	6.15	65	45	0.69	investigate	2	F	information	3	D	research	4	F	graph	3	F
16	define	5.26	42	31	0.74	dictionary	2	F	eye	2	?	contact	2	?	country	2	?
17	derive	3.24	25	20	0.80	name	4	F	from	3	H	grandmother		F			
18	distribute	3.18	21	19	0.90	distribution	3	C									
19	economy	6.32	72	44	0.61	company	2	F	money	13	F	save	4	F	class	2	F
20	environment	6.41	67	39	0.58	ecology	2	F	destruction	2	F	pollution	8	F	forest	4	F
21	establish	4.41	45	23	0.51	found	7	D	school	6	F	college	2	F	make	4	E
22	estimate	3.85	30	25	0.83	money	5	F	budget	2	F						
23	evident	3.79	29	24	0.83	evidence	3	C	plain	2	D	clear	2	D	accident	2	F
24	export	4.79	60	29	0.48	import	20	F	world	2	F	sea	2	F	track	2	F
25	factor	4.97	34	31	0.91	fact	2	B	cause	3	F	result	3	F			
26	finance	4.32	53	24	0.45	money	20	D	economy	2	E	company	3	F	financial	2	C
27	formula	3.41	21	18	0.86	formal	2	A	car	3	F						
28	function	4.71	36	34	0.94	car	2	F	move	2	E						
29	identify	5.68	49	37	0.76	person	2	F	identity	5	C	mind	2	F	individual	3	E
30	income	5.68	65	28	0.43	high	2	F	father	2	F	money	12	E	salary	4	D
31	indicate	4.65	26	22	0.85	problem	2	F	show	2	D	indication	3	C			
32	individual	5.59	46	36	0.78	personal	7	D	person	2	D	oneself	2	E	see	2	?
33	interpret	4.50	31	23	0.74	understand	4	D	translate	3	E	book	2	F	English	3	F
34	involve	4.85	36	30	0.83	accident	3	F	concern	3	D	incident	2	F	problem	2	F
35	issue	5.85	57	25	0.44	question	6	D	difficult	2	G	problem	19	D	solve	5	G
36	labor	4.24	43	29	0.67	hard	2	G	work	10	D	worker	3	F	company	3	F
37	legal	4.50	56	36	0.64	crime	2	F	police	2	F	criminal	2	F	illegal	6	C
38	legislate	2.71	22	20	0.91	law	3	F									
39	major	6.26	67	34	0.51	baseball	9	F	minor	20	F	specialize	2	D	many	3	E
40	method	5.88	48	28	0.58	way	18	E	use	2	G	science	2	F	means	2	E
41	occur	4.85	38	25	0.66	happen	11	D	issue	2	F	accident	3	F			

42	percent	6.24	58	32	0.55	average	3	E	proportion	6	E	number	9	F	graph	2	
43	period	6.03	61	29	0.48	sentence	3	F	term	10	D	era	2	D	point	3	
44	policy	5.65	47	35	0.74	thought	3	E	politic	2	C	important	3	G	pride	4	
45	principle	4.65	47	31	0.66	school	8	?	important	4	G	one	2	?	principal	2	
46	proceed	4.06	25	16	0.64	advance	6	D	progress	4	D	way	2	F			
47	process	6.44	51	33	0.65	produce	2	A	make	2	?	result	7	F	way	8	
48	require	5.03	45	30	0.67	demand	4	A	request	2	E	ask	2	E	need	6	
49	research	6.24	53	31	0.58	study	7	D	examine	2	D	find	2	E	data	4	
50	respond	5.68	36	25	0.69	call	2	F	question	4	H	answer	5	D	letter	2	
51	role	5.21	50	27	0.54	musical	2	F	part	5	D	movie	3	F	cast	3	
52	section	5.32	48	31	0.65	part	10	D	group	4	E	company	2	E	devide	2	
53	sector	3.00	15	14	0.93	section	2	D									
54	significant	3.74	25	15	0.60	important	11	D									
55	similar	5.76	46	24	0.52	brother	2	F	like	5	D	familiar	4	A	resemble	4	
56	source	4.85	40	29	0.73	resource	6	A	forest	2	F	oil	2	F	water	2	
57	specific	3.26	27	21	0.78	life	2	?	space	4	?	special	3	E			
58	structure	3.97	35	32	0.91	building	2	F	house	2	F	architecture	2	F			
59	theory	6.00	53	42	0.79	science	5	F	philosophy	2	D	thinking	2	E	practice	3	
60	vary	4.56	41	23	0.56	color	2	F	many	2	F	change	10	D	variety	2	

Average 4.89 43.22 28.62 0.69

Appendix B (Sublist 10) \*=non-word

		F	X	Y	Z	1	N		2	N		3	N		4	N	
1	adjacent	1.81	16	7	0.44	* adjust	6	A	adjust	5	A						
2	albeit	1.74	9	8	0.89	work	2	?									
3	assemble	5.26	31	17	0.55	resemble	5	A	gather	3	D	like	3	?	to	3	?
4	collapse	5.81	35	27	0.77	break	5	E	destroy	4	E	down	2	E			
5	colleague	5.33	37	28	0.76	friend	7	E	company	4	F						
6	compile	2.73	6	6	1.00												
7	conceive	5.96	31	28	0.90	think	3	D	suddenly	2	?						
8	convince	6.59	33	25	0.76	persuade	6	D	think	2	?	understand	2	?	of	2	H
9	depress	5.63	28	19	0.68	impress	5	A	discourage	2	D	sad	2	E	dislike	2	F
10	encounter	6.3	50	31	0.62	meet	10	D	come	4	E	happen	3	E	accident	3	F
11	enormous	6.26	42	19	0.45	big	8	E	large	7	E	many	5	E	huge	4	D
12	forthcoming	2.63	8	8	1.00												
13	incline	5.15	25	17	0.68	decline	6	A	to	3	H	tend	2	D			
14	integrity	3.56	15	12	0.80	integrate	2	C	in	2	?	integral	2	C			
15	intrinsic	2.56	15	15	1.00												
16	invoke	2.48	17	14	0.82	involve	4	A									
17	levy	2.04	9	6	0.67	envy	4	A									
18	likewise	3.85	22	14	0.64	othrewise	6	A	like	3	B	wise	2	B			
19	nonetheless	5.56	36	19	0.53	nevertheless	8	D	but	6	D	though	4	D	however	3	D
20	notwithstanding	3.04	21	15	0.71	withstand	4	B	withstanding	3	B	stand	2	B			
21	odd	6.04	41	27	0.66	strange	8	D	add	4	A	stupid	2	E	ugly	2	E
22	ongoing	3.41	19	16	0.84	continue	3	E	go	2	E						
23	panel	4.44	19	18	0.95	board	2	E									
24	persist	6.19	35	18	0.51	insist	10	A	on	3	?	say	2	E	cling	2	E
25	pose	5.59	35	22	0.63	stop	9	?	picture	3	F	to	2	?	model	2	F
26	reluctance	4.59	21	11	0.52	reluctant	5	C	unwilling	4	E	* relax	2	A	to	2	H
27	so-called	4.59	21	17	0.81	what	3	?	called	2	B	name	2	?			
28	straightforward	3.04	20	19	0.95	road	2	?									
29	undergo	5.33	31	23	0.74	experience	7	D	undertake	3	D						
30	whereby	3.67	13	10	0.77	where	2	B	* therefor	2	A	nearby	2	A			
Average		4.37	24.70	17.20	0.73												

# A Study of the Effectiveness of the CALL Program, 'Adjective Sommelier', as a Learning Tool to Improve Learners' Analytical Approach to the Polysemous Senses of TL Adjectives

Noriko Aotani<sup>1</sup>, Taichi Kameyama<sup>2</sup>, Naoki Sugino<sup>3</sup> and Yuko Amaya<sup>4</sup>

<sup>1</sup>Tokai Gakuen University, <sup>2</sup>Gifu National College of Technology,

<sup>3</sup>Ritsumeikan University, <sup>4</sup>Tokai Gakuen University

aotani@tokaigakuen-u.ac.jp, kame@gifu-nct.ac.jp

gwisno@is.ritsumei.ac.jp, amaya@tokaigakuen-u.ac.jp

## Abstract

This study aims to analyze the effectiveness of the CALL program 'Adjective Sommelier', which we have developed for Japanese learners of English to experience an analytical, inquiring approach to expanding the polysemous senses of adjectives (Aotani & Kameyama, 2008). The program was developed to help encourage learners who are suspicious of transferability to notice underlying cross-linguistic similarities between L1 and TL, to develop the necessary confidence to decide when it is appropriate to transfer the word senses and when it is not, and to develop learner autonomy over their metaphorical thinking process.

The effectiveness of the program was examined using data from 182 Japanese college and university students by analyzing changes in their learning strategies after using this program. The scale of learning strategies consists of three subscales: analogical strategy, consideration strategy, and working strategy (Mori, 2004).

Results indicated the scores of the post survey on consideration strategy and working strategy were significantly higher than the pre survey scores regardless of learners' English proficiency level. The effectiveness of 'Adjective Sommelier' was discussed in relation to the findings.

## Keywords

language transferability, polysemous adjectives, learning strategies

## Introduction

Japanese and English are regarded as distant

languages. Ringbom (2007) defines the relation of distant languages as a zero relation, which means that items and patterns in the TL at an early stage of learning appear to have little or no perceptible relation to the L1 or any other language the learner knows. Aotani (2003, 2004, 2007) investigated the results of the language transferability tests regarding pair adjectives as 'omoi' (*heavy*) or 'karui' (*light*), 'fukai' (*deep*) or 'asai' (*shallow*), 'hiroi' (*wide*) or 'semai' (*narrow*), among others, administered to Japanese high school students and university students. The results showed that the metaphorical expansion of word senses in their TL mental lexicon did not occur properly, even if they continued studying English through university level. They judged the transferability as acceptable only within the basic sense of the word, and most of the metaphorical expansion was related to language-oriented expressions.

The zero relation, however, does not mean there is nothing relevant between two languages, since there are always some linguistic universals common to all languages. In the world of meanings that involves the way humans perceive the world, most words especially adjectives and verbs are polysemous in all languages. Their meanings have been metaphorically expanding throughout their long history. Some expansions are culturally oriented, and others are more general, having underlying concepts that are universal. As Ringbom (2007) suggested, universal similarities need to be pointed out to the learner in an explicit way since the level of abstraction in those universals is too high to notice. Without any appropriate guidance, learners have to make trial and error

continuously, which is not the most productive approach to language learning.

### 1. The aim of the program

Figure 1 illustrates the world of meaning in L1 and TL, and their relations. Expressions which are peculiar to L1-speaking people are categorized in *Area A*. Expressions which are peculiar to TL-speaking people are categorized in *Area C*. *Area B* is a cross-linguistic area, and expressions which are common in L1 and TL are categorized in this area. As described above, *Area B* for Japanese average learners of English is unnecessarily small.

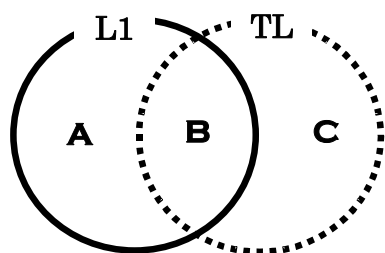


Figure 1: A model of the L1-TL relations

The program, ‘Adjective Sommelier (hereafter AS)’, consists of ten units each of which consists of three stages; the stage of perceiving, the stage of deeper processing, and the stage of new association. The first two stages aim to expand a learner’s *Area B*, which simultaneously shrinks his/her *Area A*, and the third stage aims to expand a learner’s *Area C*, which means perceiving the difference (Aotani & Kameyama, 2008).

At the first stage, the stage of perceiving, learners are given, one by one at random order, 15 Japanese phrases that contain an adjective such as ‘*omoi*’ and are instructed to make a judgment as to whether the English equivalent adjective (*i.e.* in the case of ‘*omoi*’, *heavy* is an equivalent adjective) could be transferable as a translation. Some of the given phrases are literal (core) usage of the adjective such as ‘*omoi nimotsu*’ (*heavy luggage*), and others are metaphorical usage such as ‘*omoi sekinin*’ (*heavy responsibility*). Some phrases are transferable, and others are nontransferable. When learners finish the analogical judgment for all the given phrases, the results of their analogical judgments are shown. The ultimate aim at this stage is to give learners the opportunity to acknowledge that there is a metaphorical expansion of the meaning in L1, and there may also be a metaphorical expansion

in TL, whether similar to or different from their L1’s.

At the second stage, the stage of deeper processing, learners are given ten English sentences one by one with the target adjective blanked out. A Japanese translation of the sentences is also given. They are instructed to choose the correct adjective. When a question shows up on the screen, a direct translation adjective is already given in the answer box as a default answer. If a learner thinks the direct translation is the right answer, he or she can just click the ‘OK’ button to proceed to the next questions. If he or she thinks the direct translation is not appropriate, he or she can choose the correct adjective among the three choices. After finishing all questions, the learner’s result and also a professional sommelier’s example answer are shown next to each other to let the learner compare his or her results with the professional’s. The ultimate aim at this stage is to encourage learners to experience a hypothesis-making process, which makes them deeply involved in the world of metaphorical expressions. This stage is also aiming to let learners develop confidence in the production of the TL. If they can start noticing underlying common schema of the meaning, the distance they perceive to exist between Japanese and English will be shortened.

The aim of this study is to analyze the effectiveness of these two stages by investigating if any changes in their learning strategies take place after experiencing these two stages of AS.

## 2. Method

### 2.1 Subjects

Subjects were 182 Japanese college and university students. Ages ranged from 17 to 22 years old. They were not English majors, and they were enrolled in compulsory English classes.

The subjects were required to take the Measures of English Grammar Test (Shimizu, *et al.*, 2006) to assess their English proficiency level before the experiment.

### 2.2 Procedure

The on-line surveys were administered to regular English classes during class hours. The subjects were assessed at two points: at the first week before experiencing AS (S1), and at the fifth week after finishing all units of AS (S2). A questionnaire was given in Japanese.

### 2.3 Materials and variables

The scale of learning strategies that was constructed by Kubo (1999) was used. It is a 22-item scale and consists of three subscales: analogical strategy, consideration strategy, and working strategy (Mori, 2004). The analogical strategy consists of seven items, such as *When I do not understand the sentences I read or hear in English, I guess the meanings from the context*, *To understand unfamiliar words, I make guess*, *When I talk in English or read English, I try to grasp the general meaning instead of sticking to details*, and *I find the meaning of an English word by dividing it into parts that I understand*. The consideration strategy consists of four items: *I consider the reason why I cannot understand*, *I try to notice my language errors and find out the reasons for them*, *I make clear what I understand and what I do not*, and *I do not leave the questions half-done*. The working strategy consists of six items, such as *I classify words based on their meanings*, *I create associations between items I learned*, *I use the English words I know in different ways*, and *I classify words based on their forms*.

Responses were scored on a six-point scale from 1 (never use) to 6 (always use). This scale was administered at two different points, S1 and S2.

### 3. Result

The design was two levels of the score of MEG (the high/low score group of MEG), and two level of time (S1/S2). The data of MEG was divided into “high score group (N=76)” and “low score group (N=106)” based on the means

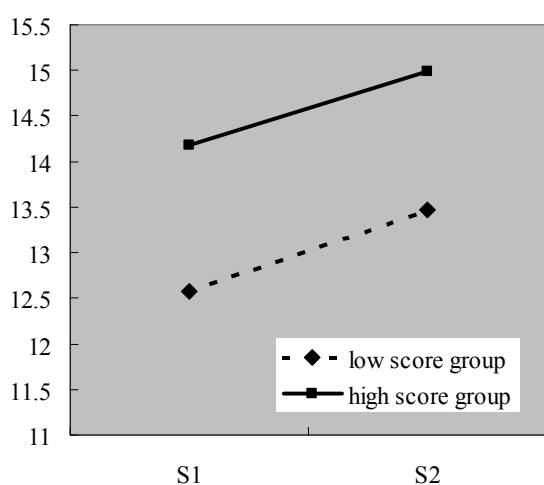


Figure2: Each group of Mean for the consideration strategy

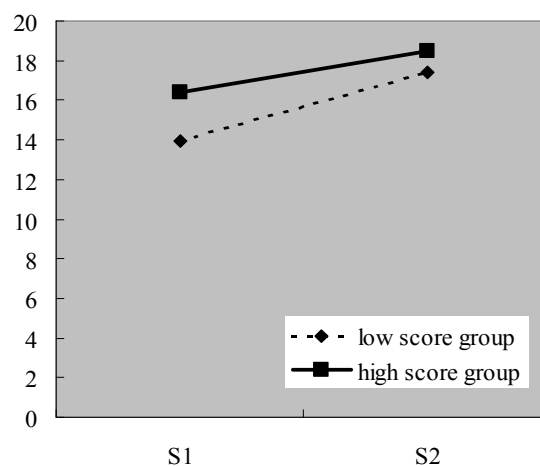


Figure3: Each group of Mean for the working strategy

of MEG (M=11.0, SD=3.6). The data for time was the repeated condition.

A two (the score of MEG) x two (time) ANOVA was carried out on the learning strategy scales. The analogical strategy ( $\alpha=.74$ ) indicated no significant main effects of time and the score of MEG.

The consideration strategy ( $\alpha=.77$ ) revealed significant main effects of time ( $F(1,180)=12.67$ ,  $p<.001$ ) and the score of MEG ( $F(1,180)=10.05$ ,  $p<.01$ , Figure 2). The interaction between time and the score of MEG was non-significant.

The working strategy ( $\alpha=.80$ ) revealed significant main effects of time ( $F(1,180)=54.01$ ,  $p<.001$ ) and the score of MEG ( $F(1,180)=6.99$ ,  $p<.01$ , Figure 3). The interaction between time and the score of MEG was non-significant.

### 4. Discussion

The aim of the first two stages of AS was to give learners the opportunity to consciously engage in the psychological process of noticing metaphorical uses of the languages. From this study we found some evidence that the learners began noticing the universal similarities in metaphorical expansions of the meaning between L1 and TL, and tried to find out the underlying schema as a clue to the solution.

Learning strategies are steps taken by students to enhance their own learning. The consideration strategy is considered to be affected when learners recognize careful consideration is effective in the face of solving problems. Changes in this strategy after using AS imply the program gave the learners the chance to psychologically involve in the world

of meanings more deeply and carefully than before. The working strategy is considered to be stimulated when learners recognize information they get is worth classifying for future use. Changes in this strategy imply during the lessons of AS the learners try to work out the strategy to classify the meanings of words, to detect similarities between L1 and TL meanings.

This study also gave evidence that AS was effective to all the subjects regardless of their English proficiency level from a viewpoint of changes in their learning strategies. When the characteristic of the task whose achievement level largely depends on the learners' English abilities, post-task impressions of the lower-level learners are usually different from ones of the higher-level learners. In case of AS as a task, however, its aim was not to test the learners' English abilities, but to stimulate the learners' cognitive process, regardless of their English proficiency level.

Although the surveys were conducted in a short term, the learners' learning strategies were still affected significantly. When we consider, however, the fact that the changes in learning strategies usually need a certain period of learning experience to be firmly established, the continuous lessons are considered to be more effective. Danesi (1992) stated the importance of including the study of metaphorical expansion of meanings on the agenda of second language acquisition in the classroom environment. However, if we look at the learning situation in Japan, it might be difficult to spare time for this kind of activity. As a result, the typical classroom Japanese learners of English have virtually no access to the metaphorical expansion of the TL even after six years of study at junior and senior high school.

We believe using the program such as AS in the everyday classroom environment can provide a learner-led learning environment for learning metaphorical expansion in the TL, and help construct their learning strategies regarding vocabulary acquisition.

The effectiveness of AS should be assessed further by different scales, such as learners' motivation, self-esteem, self-evaluation, or enjoyment of learning.

### Acknowledgments

This work was supported by KAKENHI (Grant-in-Aid for Scientific Research (C); No. 20520535).

### References

- Aotani, N. (2003). A psycholinguistic analysis of the expansion of semantic network of ESL learners: A questionnaire survey of the Japanese adjective *omoi*. *The Bulletin of Chubu English Language Education Society*, 33, 9-16.
- Aotani, N. (2004). An analysis on the analogical semantic network of the Japanese adjective *omoi*. *The Bulletin of Tokai Gakuen University*, 9-2, 67-80.
- Aotani, N. (2007). A comparative study of transferability between 'measure' adjective pairs, *omoi* and *karui*. *The Bulletin of Tokai Gakuen University*, 12-2, 1-12.
- Aotani, N. and Kameyama, T. (2008). Development of a CALL program to improve learner's analytical approach to the polysemous senses of L2 adjectives. *Proceedings of WorldCALL 2008*. ([http://www.j-let.org/~wcf/modules/tiny\\_d12/](http://www.j-let.org/~wcf/modules/tiny_d12/))
- Danesi, M. (1992). Metaphorical competence in second language acquisition and second language teaching: The neglected dimension. In J.E. Alatis (ed.) *Language Communication and Social Meaning* (pp. 489-500). Washington, DC: Georgetown University Round Table on Languages and Linguistics.
- Kubo, N. (1999). Orientation-appraisal model for Japanese university students' learning of English: Relation among learning motive, cognitive appraisal, learning behavior, and performance. *Japanese Journal of Educational Psychology*, 47, 511-520.
- Mori, Y. (2004). Relationships between self-efficacy and learning strategies of English in college students. *Japan Journal of Educational Technology*, 28, 45-48.
- Ringbom, H. (2007). *Cross-linguistic similarity in foreign language learning*. Multilingual Matters Ltd
- Shimizu, Y., Yamakawa, K., Sugino, N., Ohba, H., and Nakano, M. (2006). Developing a sample-free grammatical proficiency test for SLA research. *Proceedings of the 10th Conference of Pan-Pacific Association of Applied Linguistics*, 227-236

# Learning Strategies: A Theoretical Assumption

Wakako Kobayashi

Faculty of Commerce, Chuo University

wakakoba@mbh.nifty.com

## Abstract

Of the four language skills, listening is often considered to be the most crucial for the language learning since listening is a highly interactive skill and many SLA researchers have demonstrated its significant role on language acquisition (Vandergrift, 1996). The appropriate and effective use of listening strategies can help learners to facilitate the language process. Oxford (1989) states that there is the interaction between learner's variables and the choice of language learning strategies (LLSs). Therefore, the use of appropriate LLSs, especially metacognitive strategies hold a key to improve and advance learners' listening proficiency as their language levels rise.

## Keywords

Listening strategies, language learning strategies (LLSs), metacognitive strategies, teacher education

## Introduction

Krashen (1982) asserted that comprehensible input is a necessary and indispensable condition for language learning when considering the relationship between input and adjustments and message comprehension. In his input hypothesis, Krashen (1985) states that if the learners can comprehend the language that contains linguistic items (lexis, syntax, morphology) at a slightly above the learner's current level ( $i + 1$ ), then, s/he can achieve the further development in

learning, especially in listening and reading ability (Rost 2001). Although Krashen does not refer to strategic adjustments, which learners makes in understanding the new information, I believe his assertion – in spite of the endless controversy for the reliability and validity, which implies that the notion of  $i + 1$  is vague and inaccurate – has implies many second language acquisition (SLA) researchers into learning strategies researchers, at least to some extent.

## 1 Hypothesis

Here, this study investigate the effects of language learning strategies, especially listening strategies. I hypothesize that 1) implementation of teaching learning strategies is effective for learners to improve their listening proficiency, 2) students can use more metacognitive strategies than cognitive and socio-affective strategies as their levels of comprehension increase. After the literature reviews, this study concludes that fostering the acquisition of the appropriate listening strategies, particularly metacognitive strategies, hold the key of enhancing success in listening comprehension (O'Malley & Chamot, 1990, Vandergrift, 1996, Rost, 2001).

## 2 Background of the study in listening strategies

As for the definition of the listening strategies, Rost (2001) notes that they are conscious plans to deal with incoming speech, especially when the listener tries to compensate for the incomplete

input or partial understanding. For the representative researchers in this, the following three are to be picked up: 1) Rost & Ross 1991; 2) Kasper, 1984 and 3) Vandergrift, 1996.

1) Rost & Ross (1991) discuss that more proficient listeners tend to use more hypothetical testing (asking about information about the story), rather than lexical push down (asking about the word's meaning) and global reprises (asking for repetition). They also report that if listeners follow the training sessions, they could ask more hypothesis questions. 2) Kasper (1984) asserts that in think aloud protocols study that L2 listeners tend to form an initial interpretation of a topic (a frame) and then stick to it, trying to fit incoming words and propositions into that frame. 3) Vandergrift (1990) classifies the strategies which consist of metacognitive, cognitive and socio-affective strategies. He finds that the higher the learners' levels are, the more they use the metacognitive strategies. Then, based on the findings, Vandergrift proposes a pedagogic plan for encouraging the use of metacognitive strategies at all proficiency levels in order to increase the learner's comprehension.

Of the four language skills, listening is often considered to be the most crucial for language learning since listening is a highly interactive skill and many SLA studies have demonstrated its significant role in language acquisition (Vandergrift, 1996). Rost (2001) notes that listening involves both bottom-up processing and top-down processing, which takes place at various levels of cognitive organization: phonological, grammatical, lexical and propositional. Listening strategies, therefore, are the most beneficial for teachers to nurture (Vandergrift, 1996). I think the appropriate and effective use of the listening strategies can help learners to facilitate the language process.

### **3 Training of learners in using appropriate strategies**

As far as I have discussed, the use of appropriate LLSs, especially metacognitive strategies, holds a key to improve and advance learner's listening proficiency as their language levels rise (O'Malley & Chamot, 1990, Ellis, 1994, Vandergrift, 1996), because I think with this effective use of the skills and strategies, learners could have more comprehensible input for SLA (Krashen, 1985), learners could lessen their anxiety in listening (Durey, Burt and Krashen, 1982) and could even increase their motivation by having stronger confidence (Anderson & Lynch, 1988). Therefore, here I consider how teachers could teach those kinds of strategies effectively to learners by reviewing related articles. First, Rubin (1994) and many other researchers (Mendelson, 1998) assert that the importance of listening strategies training in the classroom teaching. Furthermore, Mendelson (1998) notes that the materials have increasingly come to include the strategy training, particularly the significance of the schemata, prior to listening, which leads to pre-, while- and post-listening phases (Underwood, 1989). Then Rost (1994) presents a framework for incorporating five types of listening strategies into classroom instruction, which consists of 1) predicting, 2) monitoring, 3) inferencing, 4) clarifying, and 5) responding. As for the way of the strategy training itself, perhaps, we can pick up three ways: 1) explicit training, direct training, 2) embedded strategy training, 3) combined strategy training.

[Note: Though Bialstock (1985) expresses his doubt, restricting their effects, that LLSs training is effective only under specific conditions and teachers should employ teaching strategies that are incongruous with useful experiences. I understand that it does not mean he completely

rejects the effect of teaching LLSs. Training is effective only under specific condition and teachers should employ teaching strategies that are incongruous with useful experience. However, it is somewhat surprising that there have been few empirical quantitative studies that have attempted to evaluate the success of the strategies training, as Ellis (1994) and Larsen &Freeman (1991) note. ]

#### 4 Summary and implications for teachers

As far as I have discussed and reviewed LLSs related articles, it is clear that teaching learning strategies will receive an increasing attention in teacher education, since the SLA researchers have not yet developed this academic field (Ellis, 1994). Empowering learners by having them develop LLSs might help them to not only cope with classroom tasks, but also continue to learn, which is to promote self- autonomous learning. Moreover, learners variables, such as age, gender proficiency, cultural and educational background, etc. and situational factors such as task difficulty, informal and formal setting and goals, etc. inevitably influence and interact with the choice of LLSs. Therefore, teachers should pay attention to them in teaching strategies. Lastly, the right and appropriate choice of LLSs may determine the rate and level of the achievement as well which is considered to be significant for both teachers and learners.

#### 5 References

- Anderson, A. & Lynch, T. (1988). Graded development of listening. *Listening* Oxford: Oxford University Press.
- Bialystock, E. (1985). The compatibility of teaching and learning strategies. *Applied Linguistics* 6: 255-262
- Duley, J.B. & Krashen, S. (1982). *Language two*. Oxford: Oxford University Press.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Kasper, G. (1984). Pragmatic comprehension in learner – native speaker discourse. *Language Learning*, 34..1-20.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. New York: Pergamon.
- Krashen, S. (1985). *The input hypothesis: issues and implication*. London: Longman.
- Larsen-Freeman, D. &Long, M. (1991). *An introduction to second language acquisition research*. London: Longman.
- Mendelson, D. (1998). Teaching listening. In W. Grabe (Ed). *Foundation of second language teaching* : Cambridge: Cambridge University Press.
- O'Malley, M. & Chamot, A. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.
- Oxford, R. (1989). Use of language learning strategies: A synthesis of studies with implication for teacher training. *System* 17:235-247.
- Rost, M. & Ross, S. (1991). Learner use of strategies in interaction: typology and teachability. *Language Learning* 41. 235-273.
- Rost, M. (2001). Listening. In R. Carter & D.Nunan (Eds). *The Cambridge guide to teaching English speakers of other language*. Cambridge: Cambridge University Press.
- Rubin, J. (1994). A review of second language listening comprehension research. *Modern Language Journal*, 78. 199-221.

Underwood, M. (1989). Teaching listening.

London: Longman.

Vandergrift, L. (1996). The listening comprehension strategies of core French high school students. *The Canadian Modern Language Review*, 52.200-223.

# The Development and Implementation of Task-based Writing Performance Assessment for Japanese Learners of English: (1) A Pilot Experiment

Yoshihito Sugita<sup>1</sup> and Michiko Nakano<sup>2</sup>

<sup>1</sup> Faculty of Nursing, Yamanashi Prefectural University, <sup>2</sup> School of Education, Waseda University

sugita@yamanashi-ken.ac.jp, nakanom@waseda.jp

## Abstract

This article describes a pilot study conducted to explore the development of task-based writing assessment for Japanese learners of English. The main purpose of the research is 1) to establish a framework for the test development and the constructs of writing performance test, 2) to conduct a pilot test of a developed writing performance assessment, and 3) to examine the degree of reliability of the assessment tasks and rating scales. We found that there is considerable validity in the integration of construct-based task development and task implementation based on the operation of the processing factors and the influence of the processing conditions, and thus the construct-based processing approach to testing resulted in a comprehensive framework for our test development. The constructs of our task-based writing test developed for this study were determined to be accuracy and communicability, and the test development proceeded according to the three stages. The pre-testing was intended as an examination into the assessment tasks and rating scales, and the analyses were done using FACETS. The results showed that 1) the students' writing ability was effectively measured, 2) the difficulty of the two assessment tasks were equivalent, and 3) rating scales for accuracy and communicability need to be optimized.

## Keywords

writing performance, task-based assessment, FACETS

## Introduction

As Bachman and Palmer mentioned (1996), the primary purpose of a language test is to make inferences about language ability. The ability that we want to test is defined as a construct, and describing the construct is one of the most fundamental concerns in test development. When assessing writing, it is therefore necessary to

address the issue of how much importance we place on the ability of our students to write. For example, the construct of TOEFL Test of Written English (TWE) is an applicant's ability to compose academic English. The TWE is holistically scored, using a criterion-referenced scale to provide information about an examinee's ability to generate and organize ideas on paper, to support those ideas with evidence or examples and to use the conventions of standard written English (ETS, 1996).

The constructs of the task-based writing performance test developed for this pre-testing are assumed to be accuracy and communicability. Accuracy is comprised of organizational skills and linguistic accuracy. Organizational skills can be defined as the ability to organize logical structure which enables the content to be accurately acquired, and linguistic accuracy concerns errors of vocabulary, spelling, punctuation or grammar. Communicability is comprised of communicative quality and effect. Communicative quality refers to the ability to communicate without causing the reader any difficulty, and communicative effect concerns the quantity of ideas necessary to develop the response as well as the relevance of the content to the proposed task (Sugita, 2008).

The present writing test contains two tasks as elicitation devices. Those tasks are designed to reify the underlying constructs based on the dual-mode system proposed by Skehan (2001). Since the task that elicits accuracy places more emphasis on writing accurately in adequate time, the rule-based system can be accessed. The task that elicits communicability emphasizes the importance of conveying a message in very limited time, so that a memory-based system will be appropriate. Specialized rating scales for the constructs were also developed, each accompanied by descriptors of the accuracy and communicability, respectively. By conforming one construct closely to the definition of its rating scale, it is fair to say that raters would

use the scale appropriately and consistently, ensuring the reliability and validity of assessing writing.

## 2. The Study

### 2.1 Purposes and research questions

The purpose of this study is to conduct a pilot test of the TBWT. The pre-testing is intended as an examination into the degree of reliability of assessment tasks and rating scales. The following research questions were addressed:

- 1) Is student ability effectively measured?
- 2) Are teacher-raters equally severe?
- 3) How much do tasks that are designed to be equivalent actually differ in difficulty?
- 4) How well scales conform to expectations about their use? Do raters use all parts of them, and use them consistently?

### 2.2 Test participants and materials

The data for this study were 30 scripts (15 scripts on each of two tasks) collected from 15 undergraduate students of a Japanese University (6 males and 9 females) who took a task-based writing performance test on the first day of the basic English course in the fall semester, 2006. The subjects were mostly freshmen (93.3%) whose majors varied. All of the subjects were native speakers of Japanese with the intermediate level of English language proficiency.

The writing performance test developed for the pre-testing contained two elicitation tasks: 20 minutes for Task 1 and 10 minutes for Task 2. The features of the two tasks are described below:

- Task 1 — Writing a 100-120 word letter introducing oneself to a host family, focusing on accuracy.
- Task 2 — Making notes about a discussion topic, “Why do you study English?” focusing on communicability.

### 2.3 Scoring procedure

Each of the thirty scripts was scored by five raters, who were all experienced Japanese high school teachers of English. They were all native speakers of Japanese, and they shared similar backgrounds in terms of qualifications of ten or more years of teaching experience. Since rating sessions were held separately in each location, a task-based writing test guide (TBWT Guide) was developed to provide a full explanation for assessment tasks, procedures and rating scales. Both scripts and the TBWT Guide were given to the teacher raters by mail at the end of December, 2007. Each of the five raters rated the entire set of thirty scripts and sent them back by the end of January, 2008. They were instructed to rate the 15 scripts of Task 1 first, and

then to rate the 15 scripts of Task 2. Finally, they were asked to holistically rate each of the participants’ writing proficiency at six levels, 1-6.

### 2.4 Data analysis

Tables 1, 2 and 3 show the descriptive statistics for the scores of the two test tasks and the impressionistic scoring. Since the average of the inter-rater coefficients for each scoring is relatively high (0.75, 0.79, 0.80), the five raters appear to be of acceptable reliability.

Table 1: Descriptive statistics of scoring Task 1

Rater	1	2	3	4	5
<i>Mean</i>	3.53	4.07	3.73	3.13	3.13
<i>SD</i>	1.35	1.48	1.09	1.40	1.64
<i>Min.</i>	1.0	1.0	2.0	1.0	1.0
<i>Max.</i>	6.0	6.0	6.0	6.0	6.0

Table 2: Descriptive statistics of scoring Task 2

Rater	1	2	3	4	5
<i>Mean</i>	3.86	4.13	3.46	3.86	3.40
<i>SD</i>	1.64	1.40	1.06	2.40	2.97
<i>Min.</i>	1.0	2.0	2.0	2.0	1.0
<i>Max.</i>	6.0	6.0	6.0	6.0	6.0

Table 3: Descriptive statistics of impression

Rater	1	2	3	4	5
<i>Mean</i>	3.86	4.13	3.60	3.66	3.06
<i>SD</i>	1.35	2.55	0.98	1.34	1.57
<i>Min.</i>	1.0	1.0	2.0	2.0	1.0
<i>Max.</i>	6.0	6.0	6.0	6.0	6.0

Table 4: Descriptive statistics of different scoring

	Task 1	Task 2	Impression
<i>Mean</i>	3.52	3.74	3.66
<i>SD</i>	1.40	1.48	1.39
<i>Min.</i>	1.0	1.0	1.0
<i>Max.</i>	6.0	6.0	6.0

Table 4 reports results for each test task and the impressionistic scoring, including its mean and standard deviation. It can be seen that the mean scores for all variables are very close, ranging from 3.52 to 3.74. The alpha coefficients for the test tasks and the holistic scoring were calculated. Davies (1990) recommended the cut-off point (.90) as an acceptable level of internal consistency on a high-stakes test. Each Cronbach’s  $\alpha$  would meet the point: .9349, .9466 and .9443 for Task 1, Task 2 and impressionistic scoring, respectively.

The correlation coefficients between the scores provide a preliminary estimate of the parallel-form reliability of each test task. As seen in Table 5, the correlation coefficients between each task and the impressionistic score fall in a range of .846 to .913, which are all significant at the 0.01 level. The correlation between the two test tasks (.701) is,

however, lower than the established estimate of reliability (.80). There is a possibility that it can be influenced by errors of measurement resulting from variation in rater severity and test tasks, as well as by the nature of the rating scale used and by the range of ability of the subjects who are being assessed. Therefore, it was necessary to use statistical models which take into account all of the factors that might affect a student's final score.

Table 5: Pearson correlation coefficients

	Task 1	Task 2
Task 2	.701**	
Impression	.913**	.846**

Note: \*\*all correlations significant at 0.01 level.

The analyses for the present study were done using FACETS, version 3.63 (Linacre, 2008). To examine the measurement characteristics of the pre-testing, the data was specified as having three facets, namely, the ability of subjects, the difficulty of tasks and the severity of raters. The partial-credit model was chosen because the scoring criteria for the rating scales were qualitatively different.

### 3. Results

#### 3.1 FACETS summary

Figure 1 shows a graphical summary of all facets and their elements. They are positioned on a common logit scale which appears in the first column labeled "measure." The second column shows the severity variation among raters. The most severe rater (ID: 2) is at the top, and the least severe rater (ID: 5) is at the bottom. The third column shows the ability variation among the 15 subjects. The subjects are ranked with high ability at the top (IDs 11 and 12) and low ability at the bottom (ID: 8). The fourth column shows the difficulty variation among tasks. The most severely scored task (Communicability) is at the top and the least severely scored task (Accuracy) is at the bottom. The last three columns graphically describe the three rating scales. Each of the two tasks and total impression has its own scale. The most likely scale score for each ability level is shown.

#### 3.2 FACETS analysis

##### 1) Is student ability effectively measured?

As shown in Figure 1, subject ability estimates range from a high of 3 logits to a low of -5 logits, indicating a spread of 8 logits in terms of students' ability. Subject separation value was 4.61, meaning that populations like these students in this study can be spread into about five levels. The reliability index was .95, which demonstrates it is possible to achieve reliable ability scores.

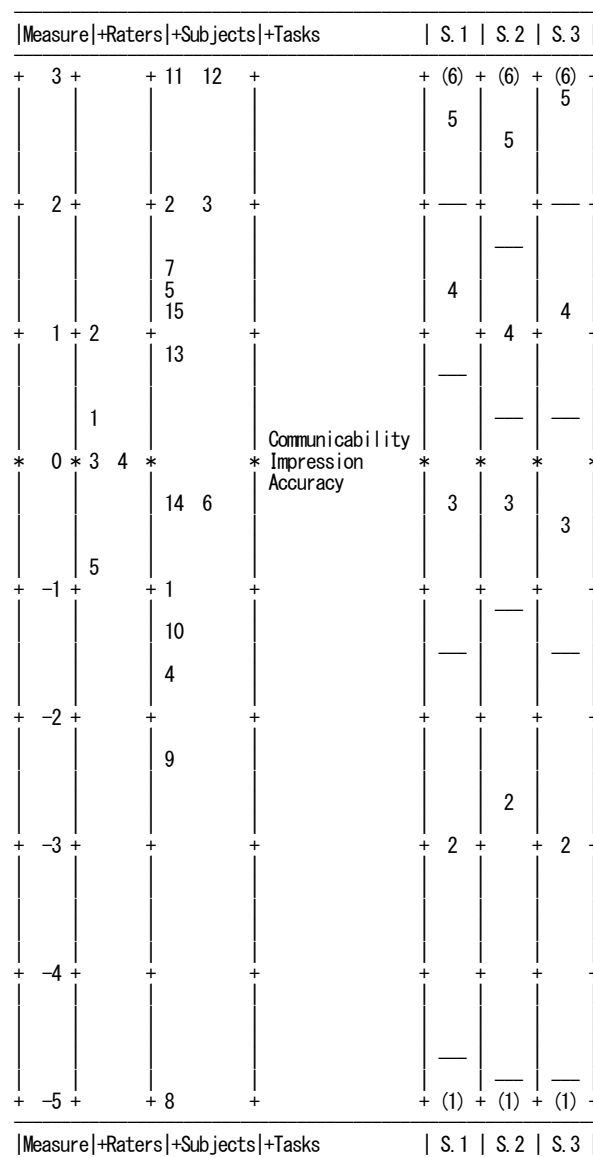


Figure 1 FACETS summary

##### 2) Are teacher-raters equally severe?

Table 6 provides information on the characteristics of raters. From the left, each column shows rater IDs, fair average scores, rater severity, error and fit mean square value. The second column indicates that the severity span between the most severe rater and the most lenient rater was 1.79 and the difference, based on fair average scores in the first column, is 1.13 of one grade in the scale. The reliability of separation index (which indicates the likelihood to which raters consistently differ from one another in overall severity) was high (.89). The chi-square of 37.4 with 4 df was significant at  $p < .00$  and, therefore, the null hypothesis that all raters were equally severe must be rejected. There was a significant difference in severity among raters. On the other hand, the Infit Mean Square column indicates that no raters were identified as misfitting: all raters fall within the range of two standard deviations around the mean ( $M - 2SD < Infit < M + 2SD$ ).

In other words, all raters are considered to behave consistently in scoring.

Table 6: FACETS analysis of rater characteristics

Rater	Fair-M average	Severity (logits)	Error	Infit (mean square)
1	3.43	.35	.21	.94
2	3.90	1.03	.21	.73
3	3.23	.05	.21	.83
4	3.18	-.03	.21	.77
5	2.77	-.76	.22	1.44
Mean	3.30	.13	.21	.94
SD	.41	.65	.00	.29

Note: Reliability of separation index=.89; fixed (all same) chi-square: 37.4, df:4; significance: p=.00

3) *How much do tasks that are designed to be equivalent actually differ in difficulty?*

The analysis of the two test tasks and holistic scoring in Table 7 shows that no significant variation in difficulty exists among them. Raters are considered to be self-consistent in scoring and the tasks do not appear to separate the subjects to a significant degree, meaning that the difficulty of the two tasks and the total impression of the tasks can be considered equivalent.

Table 7: Descriptive statistics on the different scoring

	Difficulty (logits)	Error	Infit (mean square)
Task 1	-.18	.16	1.01
Task 2	.14	.16	1.10
Impression	.04	.17	.68
Mean	.00	.16	.93
SD	.16	.00	.22

Note: Reliability of separation index=.00; fixed (all same) chi-square: 2.0, df:2; significance: p=.36

4) *How well do scales conform to expectations about their use? Do raters use all parts of them, and use them consistently?*

Linacre (1997) has proposed guidelines for a rating scale: (1) average category measures should advance monotonically with each category, (2) outfit mean-squares should be less than 2.0, and (3) the step difficulty of each scale should advance by at least 1.4 logits and by no more than 5.0 logits. Table 8 shows the rating scale statistics for accuracy. Since higher category scores are intended to reflect higher measures, the average category measures are expected to rise. The average measure for category 6 is, however, slightly less than for category 5. All outfit mean-squares are less than 2.0 meaning that each of the six categories has expected randomness in choosing categories. Step difficulties between 5 and 6 rose by .48, which does not meet (3).

Table 8: Rating scale statistics for Accuracy

Category Score	Average Measure	Infit (mean square)	Step Difficulty
1	-5.02	.4	
2	-1.76	.9	-4.48
3	-.07	1.0	-1.36
4	.92	1.2	.83
5	2.31	.6	2.26
6	2.29*	1.4	2.74

Table 9: Rating scale statistics for Communicability

Category Score	Average Measure	Infit (mean square)	Step Difficulty
1	-3.11	.1.9	
2	-2.08	.8	-4.74
3	-.03	1.3	-.66
4	1.12	.8	.52
5	2.13	.9	1.22
6	2.05*	1.5	3.66

Table 9 shows the rating scale statistics for communicability. The average measure for category 6 is also slightly less than for category 5. All outfit mean-squares are less than 2.0, which meet (2). Step difficulties between 3 and 4 rose by 1.18, which is less than 1.4. The spread between 4 and 5 is .70, which does not meet (3).

Table 10: Rating scale statistics for Impression

Category Score	Average Measure	Infit (mean square)	Step Difficulty
1	-3.11	.1.9	
2	-2.08	.8	-4.74
3	-.03	1.3	-.66
4	1.12	.8	.52
5	2.13	.9	1.22
6	2.05*	1.5	3.66

Table 10 shows the rating scale statistics for Impression. Average measures advanced monotonically with each category. All outfit mean-squares are less than 2.0. All step difficulty increases fall within 1.4 and 5.0, which does meet (3). In sum, the rating scale for impression conformed to the best of expectations about its use.

## 4. Discussion

### 4.1 Implications for scoring procedure

The five raters in this study were all experienced Japanese high school teachers of English and shared similar backgrounds in terms of qualifications of ten or more years of teaching experience. The results of the pre-testing, however, indicate an influence by errors of measurement resulting from variation in rater severity. Related to the previous studies on rater variation, the importance and effect of rater training have begun to be recognized (Shohamy, Gordon and Kraemer, 1992; Weigle, 1994). The implication is that the function of

training is not to force raters into agreement with each other, but rather to train raters to be self-consistent. This internal consistency will make the subjects' measurement of ability more accurate, and differences in rater severity will be compensated for mathematically (Lunz, Wright and Linacre, 1990; Weigle, 1998). Therefore, as in this pre-testing, if the raters displayed acceptable levels of consistency, a thorough understanding of a definition of the ability being measured by the test may be a central aspect of the training process. This view of the function of training addresses the concern that the TBWT Guide needs to be revised in order to give raters a shared understanding of each construct. The scoring guide which gives clearer understanding of the constructs and rating scales may effectively reduce the differences or biases caused by variation among raters.

#### 4.2 Implications for the development of rating scales

Many large-scale tests of writing such as the TOEFL and FCE use a six point scale. Bachman and Palmer (1996) recommend using more score points than there are decisions to be made in terms of reliability. Pollitt (1990) also points out that the number of points on a writing rating scale can be verified as reliable among five scale points or more. Subsequently, the rating scales for this pre-testing are determined to have six levels.

The correlational analysis showed that the five teacher raters appeared to be of acceptable reliability with each rating scale. The results of the FACETS analysis, however, suggested that rating scales for task 1 and 2 should be optimized. It would be necessary to examine the step difficulties of modified scales utilizing FACETS in order to construct sufficient discernable distance between steps in the scale (Linacre, 2002). As for the scale for task 1, it would be reasonable to combine category 5 and 6 so that a modified 5-point scale would be established. For the developmental scale for task 2, there are probably some options to combine adjacent categories such as 3 and 4, 4 and 5, or 5 and 6. In this way, the modified scales for the TBWT may be determined to have five levels. Since Japanese teachers are familiar with giving students a score from 1 to 5, most teachers would accept that a five point scale is more suitable to the grading system at Japanese schools. Further research is recommended to see if these modified 5-point scales can be a more reliable tool in determining the estimate of subjects' writing ability.

#### 5. Conclusion

This research was conducted in order to develop a

task-based writing performance assessment for Japanese learners of English. We found that there is considerable validity in the integration of construct-based task development and task implementation based on the operation of the processing factors and the influences of the processing conditions. As a result, the TBWT was developed on the basis of this construct-based processing approach to testing and was pilot-tested. The pre-testing provided us with the following findings: 1) the students' ability was effectively measured using the developed assessment tasks and five teacher raters; 2) all raters displayed acceptable levels of consistency with themselves, but there were significant differences among raters in terms of severity as was indicated in the previous studies; 3) although the correlation between the two test tasks was lower than the established estimate of reliability, the FACETS analysis showed that the difficulty of the two tasks and the impressionistic scoring were considered equivalent; 4) the equivalence of task difficulty may indicate that the test tasks measure the unified writing ability to be assumed as a construct; and 5) the FACETS analysis also showed that both rating scales need to be optimized. In further research it would be better to establish the 5-point scales based on the results, and to see if the modified scales can be a more reliable tool in determining the estimate of subjects' writing ability.

#### References

- Alderson, J.C., Clapham, C. and Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Bachman, L.F. and Palmer, A.S. (1996). *Language testing in practice: designing and developing useful language tests*. Oxford: Oxford University Press.
- Bachman, L.F. (2002). Some reflections on task-based language performance assessment. *Language Testing* 19, 453-76.
- Brown, J.D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall Regents.
- Connor, U. and Carrell, P. (1993). The interpretation of tasks by writers and readers in holistically rated direct assessment of writing. In J.G. Carson & I. Leki (Eds.), *Reading in the composition classroom* (pp. 141-160). Boston, MA: Heinle and Heinle.
- Cumming, A. (1989). Writing expertise and second language proficiency. *Language Learning* 39, 81-141.
- Cumming, A. (1990). Expertise in evaluating

second language composition. *Language Testing* 7, 31-51.

Davies, A. (1990). *A Principles of language testing*. Basil Blackwell.

Hyland, K. (2003). *Second language writing*. Cambridge: Cambridge University Press.

Linacre, J. (2002). Guidelines for rating scales. Mesa Research Note 2 (Online). Available at <http://www.rasch.org/rn2.htm> (accessed 18 March 2008).

Linacre, J. (2008). *Facets*, version no. 3.63. Computer program. Chicago, IL: MESA Press.

Lunz, M., Wright, B. and Linacre, J. (1990). Measuring the impact of judge severity on examination scores. *Applied Measurement in Education* 3, 331-45.

Pollitt, A. (1990). Response to Charles Alderson's paper: 'Bands and scores.' Alderson, J. C. (1991). In J. C. Alderson and B. North (Eds.), *Language Testing in the 1990s*. (pp. 87-91). London: Macmillan.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Shohamy, E., Gordon, C. and Kraemer, R. (1992). The effect of raters' background and training on the reliability of direct writing tests. *Modern Language Journal* 76, 27-33.

Sugita, Y. (2008). *Task-based performance assessment of Japanese second language writing*, Paper presented at 42<sup>nd</sup> International Annual IATEFL Conference, Exeter, the UK, April.

Weigle, S. (1994). Effects of training on raters of ESL compositions. *Language Testing* 11, 197-223.

Weigle, S. (1998). Using FACETS to model rater training effects. *Language Testing* 15, 263-87.

Weigle, S. (2002). *Assessing writing*. Cambridge: Cambridge University Press.

paragraphing (ESL: Linguistic accuracy )

**B+(5)** An essay at this level is between A and B

[Communicability]

**A(6)** An essay at this level

- displays consistent facility in use of language (TWE: Communicative quality )
- contains well-chosen vocabulary to express the ideas and to carry out the intentions(M: Communicative quality )
- effectively addresses the writing task (TWE: Communicative effect )
- has a very positive effect on the target reader with adequately organized relevant ideas (FCE: Communicative effect )

**B+(5)** An essay at this level is between A and B

-----  
TWE: TOEFL writing scoring guide (ETS, 2000)

FCE: FCE scoring rubrics (UCLES, 1997)

M: Michigan writing assessment scoring guide (Hamp-Lyons, 1990)

ESL: ESL composition profile (Jacobs *et al.*'s, 1981)

## Appendix A: Rating scales for pre-testing

[Accuracy]

**A(6)** An essay at this level

- is well organized and well developed (TWE: Organizational skills)
- shows strong rhetorical control and is well managed (M: Organizational skills)
- demonstrates clear organization with a variety of linking devices (FCE: Organizational skills )
- demonstrates appropriate word choice though it may have occasional errors (TWE: Linguistic accuracy)
- has few errors of agreement, tense, number, word order/function, articles pronouns, prepositions spelling, punctuation, capitalization,

# **The Development and Implementation of Task-based Writing Performance Assessment for Japanese Learners of English: (2) How to improve rating scales**

Yoshihito Sugita<sup>1</sup> and Michiko Nakano<sup>2</sup>

<sup>1</sup> Faculty of Nursing, Yamanashi Prefectural University, <sup>2</sup> School of Education, Waseda University

sugita@yamanashi-ken.ac.jp, nakanom@waseda.jp

## **Abstract**

This research describes a part of the development process of a task-based writing performance assessment for Japanese learners of English. In the pilot experiment, two assessment tasks to elicit the underlying constructs and rating scales were developed. The pre-testing was conducted on the 15 university students and five experienced teacher raters, who provided scores for two sets of 15 texts. The correlational analysis showed that the teacher raters could produce relatively high scores with each rating scale. The results of FACETS analysis, however, suggested that rating scales for the tasks should be optimized. In this research, we explored how rating scales can be developed and improved based on the results of the pre-testing. First, we optimized the rating scales on the basis of the criteria that Linacre (2002) proposed. Then, the scoring guide was revised in each section. As a result, the new rating scales are comprised of clearer descriptions of each construct and of 5-point Likert scales. The descriptors of each category are also provided with written samples as an explanatory part of the scale in order to give raters a shared understanding of the constructs and to make them more self-consistent in scoring.

## **Keywords**

writing performance, task-based assessment, rating scales

## **Introduction**

The main purpose of this research is to modify and improve the rating scales that were pilot-tested in The Development and Implementation of Task-based Writing Performance Assessment: (1) A Pilot Experiment, so that it can conform to expectations about their use. In the pre-testing, an elicitation task (writing a letter) was chosen, and specific topics of self-introduction were given in the task. A situation is supposed in which the

student is going to stay with a host family in Britain, and is suggested to write a letter, so that students can focus on writing accuracy. On the contrary, communicability tasks need form-oriented support and meaning-focused stakes in order to write with a focus on meaning. A discussion task was designed because it encourages students to write their opinions or ideas about the topic, and it lays emphasis on meaning-focused response.

Five raters, who were all experienced Japanese high school teachers of English, scored each of the thirty scripts collected from 15 undergraduate students of a Japanese University. They were instructed to rate the 15 scripts of Task 1 first and then to rate the 15 scripts of Task 2. Finally, they were asked to rate each of the participants' writing proficiency based on the total impression at six levels, 1-6. The correlative analysis of the testing showed that the five teacher raters appeared to be of acceptable reliability with each rating scale. The results of the FACETS analysis, however, suggested that rating scales for task 1 and 2 should be optimized. This article discusses how rating scales can be developed and improved based on the results of the pre-testing.

## **1. The study**

### **1.1 FACETS analysis of the rating scales for pre-testing**

Linacre (2002) has proposed the guidelines for a rating scale: (1) average category measures should rise monotonically with category, (2) outfit mean-squares should be less than 2.0, and (3) step difficulty advances should be larger than 1.4 logits and smaller than 5.0 logits. Table 1 shows the rating scale statistics for accuracy. Since higher categories are intended to reflect higher measures, the average category measures are expected to increase. Average measure for category 6 is, however, slightly less than for category 5. All outfit mean-squares are less than 2.0 meaning that each of

the six categories has expected randomness in choosing categories. Step difficulties between 5 and 6 rise by .48, which does not meet (3).

Table 1: Rating scale statistics for Accuracy

Category Score	Average Measure	Infit (mean square)	Step Difficulty
1	-5.02	.4	
2	-1.76	.9	-4.48
3	-.07	1.0	-1.36
4	.92	1.2	.83
5	2.31	.6	2.26
6	2.29*	1.4	2.74

Table 2 shows the rating scale statistics for communicability. Average measure for category 6 is also slightly less than for category 5. All outfit mean-squares are less than 2.0, which meet (2). Step difficulties between 3 and 4 rise by 1.18, which is smaller than 1.4. The spread between 4 and 5 was .70, which does not meet (3).

Table 2: Rating scale statistics for Communicability

Category Score	Average Measure	Infit (mean square)	Step Difficulty
1	-3.11	.1.9	
2	-2.08	.8	-4.74
3	-.03	1.3	-.66
4	1.12	.8	.52
5	2.13	.9	1.22
6	2.05*	1.5	3.66

The results of the pre-testing indicated that it was relatively difficult for raters to share understanding of B+category with 5 points whose level is between A and B. Therefore, it may be best to combine adjacent categories such as 5 and 6 for the scale of task 1, and 4 and 5 for the scale of task 2 because the spread between these categories is narrow. As for the scale of task 2, there is another option to combine 3 and 4, and thus we examine the both options utilizing FACETS.

## 1.2 Developmental procedures

The first stage is to combine the two scores from category 5 and 6 of the scale for task 1 (accuracy). The second stage is to combine the two scores from category 3 and 4, or 4 and 5 of the scale for task 2 (communicability). The third stage is to conduct a FACETS analysis on each data set. The final stage is to check the results based on Linacre's rating scale guidelines. In this way, the modified 5-point scales are verified utilizing FACETS. The step difficulties of modified scales are examined in order to construct a sufficient discernable distance between steps in the scale.

After optimizing the rating scales, the TBWT scoring guide should be revised in each section. The

first section is the background of the TBWT. The second section is the explanation of assessment tasks. The third section is the implementation method of the testing. The fourth section is the evaluation criteria and written samples of each score. In terms of modifying rating scales, the fourth section needs to be carefully revised. There are four possible steps for revising the section. The first step is to describe the constructs (i.e. accuracy and communicability) more clearly by indicating each sub-skill: organizational ability and linguistic accuracy for *accuracy* and communicative quality and effect for *communicability*. The second step is to choose the most appropriate written sample for each category on the basis of the specified constructs and the modified score points at five levels. The third step is to construct a new 5-point Likert scale and to provide the descriptors of each category with the selected five written samples as an explanatory part of the rating scales.

## 1.3 Developing rating scales

### 1.3.1 Combining categories 5 and 6 (accuracy), and 3 and 4 (communicability)

Table 3 shows the rating scale statistics for accuracy. "Average measure" values advance with category, so that higher measures correspond to higher category scores. All outfit mean-squares are less than 2.0 meaning each of the five categories has expected randomness in choosing categories. Most step difficulty increases fall within 1.4 and 5.0, but step difficulties between 4 and 5 rose by 1.03, which does not meet (3) of Linacre's guidelines for rating scales.

The scale structure probability curves are shown in Figure 1. Starting from the left, category 1 is most likely to be observed for low-measure scripts. Then as script measures increase, the probability of observing category 2 increases. With increasing measure, category 3 becomes most probable, then 4, and finally 5. According to Linacre (1999), if the modeled category probability curves depict a succession of "hills", the step difficulties successively increase with category scores, meaning that each category in turn is most likely to be chosen. Tyndall and Kenyon (1995) mentioned that the obvious peaks and divisions between the categories indicate that the scales conform to the expectations regarding their use. In Figure 1, one peak (category 4) is not like the expected succession of hills and the division between 4 and 5 is not obvious. This is because the step difficulties are smaller than 1.4 logits, which implies that the rating scale can not be completely decomposed into five categories.

Table 3: Rating scale statistics for Accuracy

Category Score	Average Measure	Infit (mean square)	Step Difficulty
1	-3.96	.4	
2	-1.09	.8	-3.54
3	0.66	1.0	-0.69
4	1.71	1.2	1.60
5	2.97	0.9	2.63

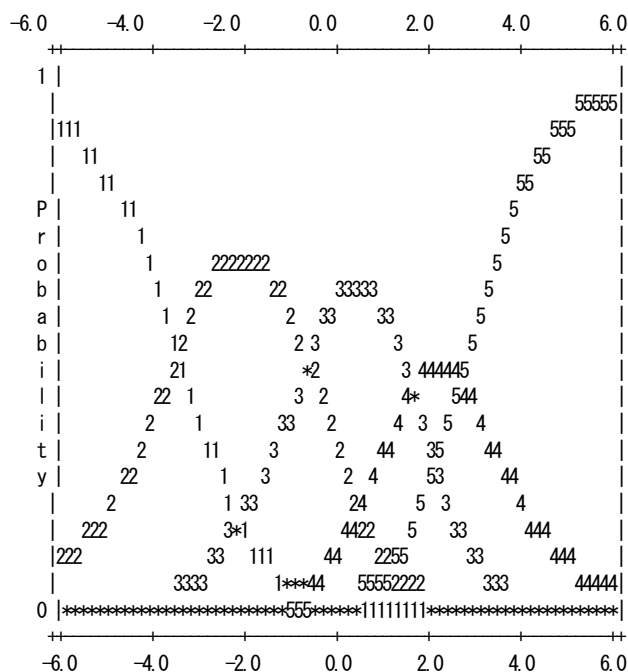


Figure 1 Probability Curves for Accuracy

Table 4 shows the rating scale statistics for communicability. Average measures advanced monotonically with each category. All outfit mean-squares are less than 2.0, which meet (2). All step difficulty increases fall within 1.4 and 5.0, which does meet (3). As a result, Figure 2 shows that the modeled category probability curves depict the expected succession of hills. In sum, when combining categories 5 and 6 of the accuracy scale, and 3 and 4 of communicability scale, the modified scale for accuracy did not conform completely to the expectations regarding their use, but the modified scale for communicability worked as intended.

Table 4: Rating scale statistics for Communicability

Category Score	Average Measure	Infit (mean square)	Step Difficulty
1	-3.00	1.5	
2	-2.03	.7	-4.42
3	0.73	.9	-.99
4	2.06	.8	1.80
5	2.23	1.6	3.61

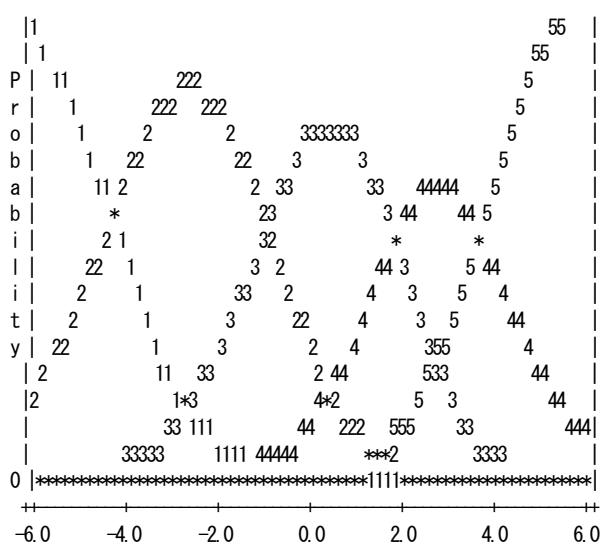
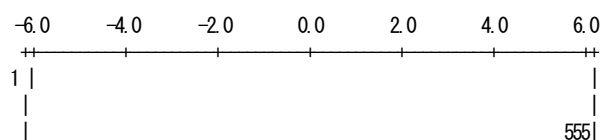


Figure 2 Probability Curves for Communicability

1.3.2. Combining categories 5 and 6 (accuracy), 4 and 5 (communicability)

Table 5: Rating scale statistics for Accuracy

Category Score	Average Measure	Infit (mean square)	Step Difficulty
1	-4.10	.4	
2	-1.16	.9	-3.66
3	0.63	.9	-0.74
4	1.80	1.2	1.65
5	3.11	.8	2.75

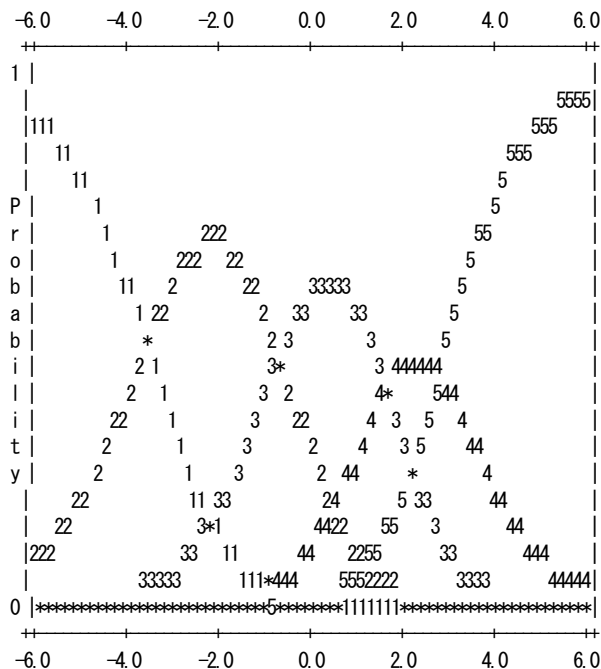


Figure 3 Probability Curves for Accuracy

Table 5 shows the rating scale statistics for accuracy. Average measures advance monotonically with each category, which does meet (1). All outfit mean-squares are less than 2.0,

meaning that each of the five categories has expected randomness in choosing categories. Most step difficulty increases fall within 1.4 and 5.0, but step difficulties between 4 and 5 rose by 1.10, which does not meet (3). Figure 3 presents the scale structure probability curves. Since the step difficulties are smaller than the criteria, one peak (category 4) is not like the range of hill and the division between 4 and 5 is not obvious as shown in Figure 1.

Table 6: Rating scale statistics for Communicability

Category Score	Average Measure	Infit (mean square)	Step Difficulty
1	-2.90	1.5	
2	-1.88	.6	-4.30
3	0.52	1.1	-.39
4	2.04	.7	0.34
5	2.51	1.2	4.36

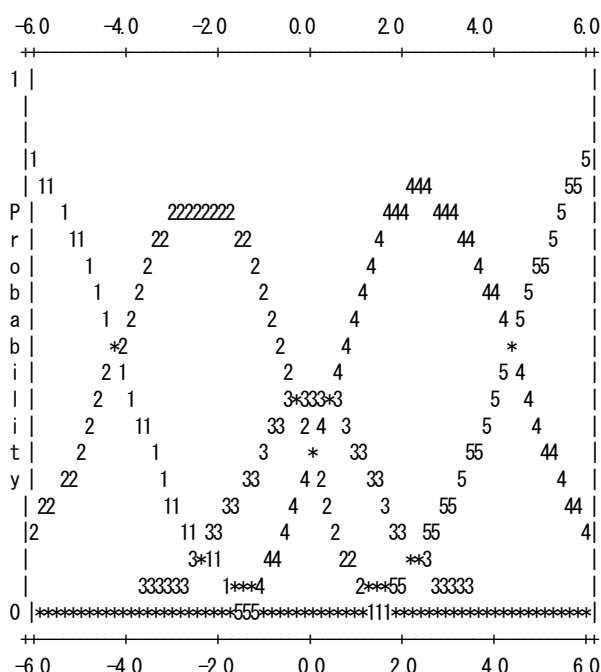


Figure 4 Probability Curves for Communicability

Table 6 shows the rating scale statistics for communicability. Average measures advanced monotonically with each category. All outfit mean-squares are less than 2.0, which meet (2). Most step difficulty increases fall within 1.4 and 5.0, but step difficulties between 3 and 4 rose by 0.73, which does not meet (3). Figure 4 indicates that neither the hill top of category 3 nor the division between 2 and 3, and 3 and 4 is obvious, meaning that the rating scale did not work properly. In sum, the modified rating scales for accuracy did not conform completely to the expectations about their use, and the modified scale for communicability hardly presents proof against its use.

As shown in the Tables 3-6 and in Figures 1-4, when combining categories 5 and 6 of the accuracy scale, and 3 and 4 of communicability scale, step difficulties between 4 and 5 on the modified scale for accuracy do not completely meet the criteria. However, the modified scale for communicability does meet the criteria. On the other hand, when combining categories 5 and 6 of the accuracy scale, and 4 and 5 of the communicability scale, the rating scales did not work properly. Therefore, the previous scales are further utilized in this study.

## 1.4. Improving rating scales

### 1.4.1. Rating scale reconstruction

In terms of improving rating scales, the fourth section of the TBWT scoring guide needs to be carefully revised. As I mentioned in the developmental procedures, the first step is to describe the constructs (i.e. accuracy and communicability) more clearly by indicating each subcomponent.

As we have seen, accuracy is defined by five descriptors, which were used as a part of the rating scale for task 1 in the pre-testing. In the construct definition, I pointed out that accuracy is comprised of organizational skills and linguistic accuracy. Organizational skills develop from 'textual competence' whose subcomponents are 'cohesion' and 'rhetorical organization.' The skills can be defined as the ability to organize logical structure which enables the content to be accurately acquired. Linguistic accuracy stems from 'grammatical competence' whose subcomponents are 'vocabulary,' 'syntax' and 'morphology.' The ability concerns errors of vocabulary, spelling, punctuation or grammar.

Communicability is also defined by four descriptors, which were used as a part of the rating scale for task 2 in the pre-testing. I have already pointed out that communicability consists of communicative quality and effect. Communicative quality develops from 'sociolinguistic competence' whose subcomponents are 'sensitivity to register' and 'sensitivity to naturalness.' It refers to the ability to communicate without causing the reader any difficulty. Communicative effect stems from 'functional competence' whose subcomponents are 'ideational function,' 'manipulative functions,' 'heuristic functions' and 'imaginative functions.' It concerns the quantity of ideas necessary to develop the response as well as the relevance of the content to the proposed task.

In order to construct a new scale, a 5-point Likert scale was adopted:

- A (5) I strongly agree to assign the above criteria
- B+(4) I partially agree to assign the above criteria
- B (3) I agree to assign the above criteria



samples (IDs: 1, 4, 9, 10) for category 2 and 1 sample (ID: 8) for category 1 as raw information, respectively. In order to examine and select the most suitable sample for each category, the following four questions were asked:

- 5) Does the written text display consistent facility in use of the language?
- 6) Is the vocabulary well-chosen to express the ideas?
- 7) Does the written text effectively address the task?
- 8) How many relevant ideas does the written text have?

These questions are also concerned with the subcomponents of *communicability*: communicative quality and effect. Finally, in order to present the five written samples which were chosen from the modified score points at each level, the description for each sample was provided as an explanatory part of the rating scales.

## 2. Conclusion

The results of the FACETS analysis for the pre-testing suggest that rating scales for task 1 and 2 should be optimized. As for the scale for task 1, by combining categories 5 and 6, a modified 5-point scale was established. For the developmental scale for task 2, there were some options to combine adjacent categories. In order to construct a sufficient discernable distance between steps in the scale, categories 3 and 4 were selectively combined to construct its modified scale. Although the modified scale for accuracy did not completely meet the criteria, the scale for communicability met all the criteria. Therefore, those scales were determined to be utilized in subsequent studies.

In order to improve the rating scales, the TBWT scoring guide also needs to be revised. The five raters in pre-testing were all experienced Japanese high school teachers of English and shared similar backgrounds in terms of qualifications of ten or more years of teaching experience. The results of the pre-testing, however, indicated an influence by errors of measurement resulting from variations in rater severity. Although the elimination of rater differences in severity is an unachievable goal as shown in the previous studies (Lumley & McNamara, 1995; Weigle, 1994), it is desperately important for raters to be self-consistent in scoring. This internal consistency will make the subjects' measurement of ability more reliable and valid. Therefore, the TBWT scoring guide needs to be revised in order to give raters a shared understanding of the construct of writing ability and to make them more self-consistent in scoring. First, the relationship among subcomponents was

reconsidered, and the rating instrument which includes clearer descriptions of each construct and a 5-point Likert scale was developed anew. Then, the most appropriate written samples for each category were chosen on the basis of the modified score points at five levels. Referring to Hyland's analytical process, eight questions related to the subcomponents of the constructs were asked to examine and select the most suitable written samples for each category. These samples along with their descriptions were used as an explanatory segment of the new modified rating scales.

In further research, the second task-based writing performance tests will be conducted to see if these modified 5-point scales can be a more reliable and valid tool in determining the best estimate of subjects' writing ability. The effects of the revised TBWT scoring guide on raters' shared understanding of the constructs will also be examined.

## References

- Alderson, J.C., Clapham, C. and Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Hyland, K. (2003). *Second language writing*. Cambridge: Cambridge University Press.
- Linacre, J. (1999). Category disordering vs. step disordering, *Rasch Measurement Transactions*, 13, (Online). Available at <http://www.rasch.org/rmt/rmt131a.htm> (accessed 18 March 2008).
- Linacre, J. (2002). Guidelines for rating scales. Mesa Research Note 2 (Online). Available at <http://www.rasch.org/rn2.htm> (accessed 18 March 2008).
- Linacre, J. (2007). *A User's guide to FACETS: Rasch-model computer program*. Chicago, IL: MESA Press.
- Linacre, J. (2008). *Facets*, version no. 3.63. Computer program. Chicago, IL: MESA Press.
- Lumley, T. & McNamara, T. F. (1995). Rater characteristics and rater bias: implications for training. *Language Testing* 12, 54-71.
- Weigle, S. (1994). Effects of training on raters of ESL compositions. *Language Testing* 11, 197-223.

# The Development and Implementation of Task-based Writing Performance Assessment for Japanese Learners of English: (4) Main Experiment 2

Yoshihito Sugita

Faculty of Nursing, Yamanashi Prefectural University

sugita@yamanashi-ken.ac.jp

## Abstract

The research examines the main data of the third task-based writing performance test in which the five junior high school teachers participated as a novice rater. In order to examine the degree of reliability and validity of the task-based writing performance test, the following are focused on, based on the second testing: raters' severity, interactions with writers' abilities and tasks difficulties, the reliability of elicitation tasks and rating scales and the measure's validity. The results of the five novice raters in this study indicated that all raters displayed acceptable levels of consistency with themselves, and that the students' ability was effectively measured using these tasks and rating scales. The FACETS analysis showed that the difficulty of the two tasks and the impressionistic scoring were considered equivalent, which provided reasonable fit to the Rasch model. There were, however, relatively small but significant differences between raters in terms of severity. The bias analyses of this study indicated all of the five raters were significantly biased towards certain types of subjects, while three of the five raters were biased in the second testing. These findings suggest that the TBWT Guide for scoring may have contributed to the reduction of biased interactions, but training for certain raters with his/her unique bias patterns might still be required.

## Keywords

writing performance, task-based assessment, scoring guide

## Introduction

The FACETS analysis for the second task-based writing performance test (Main Experiment 1) showed that the difficulty of the two tasks and the impressionistic scoring were considered equivalent, which provided reasonable fit to the Rasch model. The equivalence of task difficulty may indicate that task development based on the construct-based

processing approach could be reliable and valid to estimate students' writing ability. The modified scales associated with the five rating categories and their specific written samples were shown to be mostly comprehensible and usable by raters, and demonstrated acceptable fit. However, there is still room for argument about the reliability and validity of assessment tasks and rating scales. In addition, the question of whether new teacher raters are self-consistent in scoring the same writing samples with the rating scales must be observed and confirmed in further research. This research examines the main data of the third task-based writing performance test in which the five junior high school teachers participated as a novice rater.

## 1. The Study

### 1.1 Purposes and research questions

In order to examine the degree of reliability and validity of the task-based writing performance test, the following are focused on: raters' severity, interactions with writers' abilities and task difficulties, the reliability of elicitation tasks and rating scales, and the measure's validity.

The specific research questions are as follows:

- 1) Is student ability effectively measured?
- 2) Are teacher-raters equally severe?
- 3) How much do tasks that are designed to be equivalent actually differ in difficulty?
- 4) How well do scales conform to expectations about their use? Do raters use all parts of them, and use them consistently?
- 5) Do individual raters score a particular group of subjects more harshly or more leniently? If so, what are the sub-patterns of ratings in terms of rater-subject interaction for each rater?
- 6) Do the raters score particular tasks more harshly or more leniently than others? If so, what are the sub-patterns of ratings in terms of rater-task interaction for each rater?
- 7) To what extent, statistically, is the task-based writing test a reliable and valid measure?

The purpose of this study is to conduct a pilot test of the TBWT.

## 2.2 Procedures

Each of the forty identical scripts used in Main Experiment 1 was scored by five raters, who were all experienced Japanese junior high school teachers of English. They were all native speakers of Japanese, and they shared similar backgrounds in terms of qualifications of ten or more years of teaching experience. The rating scales and scoring guidelines are also the same as used in Main Experiment 1. Both scripts and scoring guidelines were given to the raters by mail at the beginning of August, 2008. Each of the five raters rated the entire set of forty scripts and sent them back by the end of August, 2008. They were instructed to rate the 20 scripts of Task 1 first and then to rate the 20 scripts of Task 2. Finally, they were asked to rate each of the participants' writing proficiency based on the total impression at five levels, 1-5. A questionnaire was also enclosed and sent back with the materials.

## 2.4 Data analysis

Tables 1, 2 and 3 show the descriptive statistics for the scores of the two test tasks and the impressionistic scoring. Since the average of the inter-rater coefficients for each scoring is relatively high (0.78, 0.77, 0.79), the five raters appear to be of acceptable reliability.

Table 1: Descriptive statistics of scoring Task 1

Rater	1	2	3	4	5
Mean	3.20	3.35	3.00	3.00	3.25
SD	0.92	1.10	1.14	0.70	0.88
Min.	2.0	2.0	1.0	2.0	2.0
Max.	5.0	5.0	5.0	4.0	5.0

Table 2: Descriptive statistics of scoring Task 2

Rater	1	2	3	4	5
Mean	3.25	3.20	2.80	3.20	3.35
SD	1.08	0.92	1.24	1.16	0.97
Min.	1.0	2.0	1.0	1.0	2.0
Max.	5.0	5.0	5.0	5.0	5.0

Table 3: Descriptive statistics of impression

Rater	1	2	3	4	5
Mean	3.25	3.35	3.05	3.15	3.35
SD	0.88	1.01	1.16	0.85	0.90
Min.	2.0	2.0	1.0	2.0	2.0
Max.	5.0	5.0	5.0	5.0	5.0

Table 4 reports results for each test task, the impressionistic scoring and the scores of *Criterion*, including its mean and standard deviation. The mean scores for all variables are very close, ranging from 3.16 to 3.23. The alpha coefficients for the test

tasks and the impressionistic scoring were calculated. Using Davies' cut-off (.90) as an acceptable level of internal consistency on a high-stakes test, each Cronbach's  $\alpha$  would meet the point: .9434, .9432 and .9480 for Task 1, Task 2 and impressionistic scoring, respectively.

Table 4: Descriptive statistics of different scoring

	Task 1	Task 2	Impression	Criterion TWE
N	100	100	100	20
Mean	3.01	2.92	2.98	2.40
SD	1.12	1.16	1.14	0.94
Min.	1.0	1.0	1.0	1.0
Max.	5.0	5.0	5.0	4.0

The correlation coefficients between the scores provide a preliminary estimate of the parallel-form reliability of each test task. As seen in Table 5, the correlation coefficients between each task and the impressionistic score fall in a range of .797 to .924, which are all significant at the 0.01 level. The correlation between the two test tasks (.797) is, however, slightly lower than the established estimate of reliability (.80). Table 5 also shows that the two tasks and impressionistic scoring correlate positively with the scores of *Criterion* ( $p < .01$ ). The highest correlation is between the *Criterion* score and Impression ( $r = .734$ ), followed by that between the *Criterion* score and Task 1 ( $r = .710$ ) and finally, between the *Criterion* score and Task 2 ( $r = .678$ ).

Table 5: Pearson correlation coefficients

	Task 1	Task 2	Impression
Task 2	.797**		
Impression	.924**	.884**	
Criterion	.710**	.678**	.734**

Note: \*\*all correlations significant at 0.01 level.

There is a possibility that the test data can be influenced by errors of measurement resulting from variation in rater harshness and test tasks, as well as by the nature of the rating scale used and by the range of ability of the subjects who are being assessed. Therefore, it was necessary to use statistical models which take into account all of the factors that might affect a student's final score.

The analyses for the present study were done using FACETS version 3.63 (Linacre, 2008). To examine the measurement characteristics of this testing, the data was specified as having three facets, namely, the ability of the subjects, the difficulty of tasks and the severity of raters. The partial-credit model was chosen because the scoring criteria for the rating scales were qualitatively different.

### 3. Results

#### 3.1 FACETS summary

Measure +Raters +Subjects +Tasks				S.1   S.2   S.3		
+ 7 +	+ 9 +			+ (5) + (5) + (5) +		
+ 6 +	+ +			+ — + + —		
+ 5 +	+ 18 +			+ + — + +		
+ 4 +	+ 6 +			+ 4 + + 4 +		
+ 3 +	+ 12 +			+ + 4 + +		
+ 2 +	+ 14 +			+ — + + —		
+ 1 +	+ 7 +			+ + — + +		
+ 0 +	+ 13 +			+ + + + +		
+ -1 +	+ 19 +			+ + + + +		
+ -2 +	+ 2 +			+ + + + +		
+ -3 +	+ 10 +			+ 3 + 3 + 3 +		
+ -4 +	+ 15 +			+ + + + +		
+ -5 +	+ 8 +			+ — + + —		
+ -6 +	+ 3 +			+ + + + +		
+ -7 +	+ 20 +			+ + + + +		
	+ 16 +			+ 2 + + 2 +		
	+ 5 +			+ 2 + + 2 +		
	+ 4 +			+ — + + —		
	+ 11 +			+ + + + +		
				+ (1) + (1) + (1) +		
Measure +Raters +Subjects +Tasks				S.1   S.2   S.3		

Figure 1 FACETS summary

Figure 1 shows a summary of all facets and their elements. They are positioned on a common logit scale, which appears as “measure” in the first column. The second column shows the severity variation among raters. The most severe rater (ID: 5) is at the top, and the least severe rater (ID: 3) is at the bottom. The third column shows the ability variation among the 20 subjects. The subjects are ranked with high ability at the top (ID: 9) and low ability at the bottom (ID: 11) as in Study 1. The fourth column shows equivalence of the difficulty variation among tasks. The last three columns graphically describe the three rating scales. Each of the two tasks and total impression has their own scale. The most likely scale score for each ability level is shown.

#### 3.2 FACETS analysis

##### 1) *Is student ability effectively measured?*

As shown in Figure 1, subject ability estimates range from a high of 6.96 logits to a low of -6.27 logits, indicating a spread of 13 logits in terms of

students’ ability. Subject separation value was 6.54, meaning that populations like the students in this study can be spread into about seven levels. The reliability index was .98, which demonstrates it is possible to achieve reliable ability scores.

##### 2) *Are teacher-raters equally severe?*

The severity span between the most severe rater and the most lenient rater was 1.79 and the difference, based on fair average scores, is 0.44 of one grade in the scale. The reliability of the separation index (which indicates the likelihood to which raters consistently differ from one another in overall severity) was high (.82). The chi-square of 27.8 with 4 df was significant at  $p < .00$  and, therefore, the null hypothesis that all raters were equally severe must be rejected. There was a significant difference in severity among raters. On the other hand, no raters were identified as misfitting: fit values for all raters were within the range of two standard deviations around the mean ( $0.93 \pm [0.20 \times 2]$ ). In other words, all raters behaved consistently in the scoring.

##### 3) *How much do tasks that are designed to be equivalent actually differ in difficulty?*

The analysis of the two test tasks and impressionistic scoring shows that no significant variation in difficulty exists among them. Raters are considered to be self-consistent in scoring, and the tasks do not appear to separate the subjects to a significant degree meaning that the difficulty of the two tasks and the total impression of the tasks can be considered equivalent. An estimate of the item discrimination was computed according to the “Generalized Partial Credit Model” approach. 1.0 is the expected value, but discriminations in the range 0.5 to 1.5 provide a reasonable fit with the Rasch model (Linacre, 2007, p.132).

##### 4) *How well do scales conform to expectations about their use? Do raters use all parts of them, and use them consistently?*

Linacre (1997) has proposed guidelines for a rating scale: (1) average category measures should advance monotonically with each category, (2) outfit mean-squares should be less than 2.0, and (3) the step difficulty of each scale should advance by at least 1.4 logits and by no more than 5.0 logits.

Table 6 shows the rating scale statistics for accuracy. Since higher category scores are intended to reflect higher measures, the average category measures are expected to rise. All outfit mean-squares are less than 2.0, meaning that each of the five categories has expected randomness in choosing categories. All increases in step difficulty fall within 1.4 and 5.0, which does meet (3).

Table 6: Rating scale statistics for Accuracy

Category Score	Average Measure	Outfit (mean square)	Step Difficulty
1	-6.26	.4	
2	-3.24	.9	-7.07
3	.58	1.1	-1.42
4	3.76	.8	2.28
5	6.16	1.6	6.22

Table 7: Rating scale statistics for Communicability

Category Score	Average Measure	Outfit (mean square)	Step Difficulty
1	-5.57	.6	
2	-2.99	1.1	-5.54
3	.66	1.3	-1.61
4	3.05	.6	2.26
5	5.88	1.3	4.90

Table 7 shows the rating scale statistics for communicability. All outfit mean-squares are less than 2.0, which meet (2). All step difficulty increases fall within 1.4 and 5.0, which does meet (3).

Table 8: Rating scale statistics for Impression

Category Score	Average Measure	Outfit (mean square)	Step Difficulty
1	-6.02	.4	
2	-3.57	.6	-6.81
3	.76	.6	-1.66
4	3.65	.9	2.23
5	6.61	1.0	6.23

Table 8 shows the rating scale statistics for Impression. Average measures advance monotonically with each category. All outfit mean-squares are less than 2.0. All step difficulty increases fall within 1.4 and 5.0, which does meet (3). While the step difficulty between 2 and 3 of the accuracy scale (5.56) was slightly larger than 5.0, the rating scales for communicability and total impression conformed to expectations about its use. In sum, these modified 5-point scales could be a reliable tool for novice raters in determining the estimate of subjects' writing ability.

5) *Do individual raters score a particular group of subjects more harshly or more leniently? If so, what are the sub-patterns of ratings in terms of rater-subject interaction for each rater?*

There were a total of nine significantly biased interactions among all raters. Table 9 summarizes the frequencies of rater-subject interactions that displayed a significant bias for each rater at various levels of the ability range. In Table 9, the first column shows the ability estimate range, and the second column shows the number of subjects within the particular range of ability estimate. In the range of 3.00 or higher, there were four subjects. Rater 2 and 4 harshly scored one subject for each, and Rater

3 leniently scored one subject. The total number of rater-subject bias interaction was three, which was 75% of the total number of subjects within this range ( $3/4=0.75$ ). There were eleven subjects whose ability estimate was between -2.99 and 2.99. Rater 4 harshly scored two subjects and Rater 1 and 2 leniently scored one subject for each. The total number of rater-subject bias interactions was four, which was 36% of the total number of subjects within this range ( $4/11=0.36$ ). In the range of 3.00 or lower, there were five subjects, and Rater 5 leniently scored one subject. This is the only one-rater-subject bias interaction, which was 20% of the total number of subjects within this range ( $1/5=0.20$ ).

Table 9: Frequency of rater-subject bias interaction

Ability	N	Harsh (Raters)					Lenient (Raters)				
		1	2	3	4	5	1	2	3	4	5
3.00 higher	4		1		1				1		
-2.99~2.99	11				2		1	1			
-3.00 lower	5										1

Each rater had the following unique rater-subject bias pattern.

- Rater 1: There was a more leniently scored subject than expected for Rater 1. The leniently scored subject was of middle range ability (between -2.99 and 2.99).
- Rater 2: There were both more harshly and leniently scored subjects than expected for Rater 2. The harshly scored subject was a high ability subject (3.00 or higher) and the leniently scored subject was of middle range ability.
- Rater 3: As in the case of Rater 1, there was a more leniently scored subject than expected for Rater 3. Unlike Rater 1, the leniently scored subject was one with high ability.
- Rater 4: There were more harshly scored subjects than expected for Rater 4. The harshly scored subjects included one subject with high ability and two subjects with middle range ability.
- Rater 5: As in the case of Rater 1 and Rater 3, there was a more leniently scored subject than expected for Rater 5. Unlike the two raters, the leniently scored subject had low ability.

6) *Do the raters score particular tasks more harshly or more leniently than others? If so, what are the sub-patterns of ratings in terms of rater-task interaction for each rater?*

The results of the bias analysis in terms of the interaction between raters and tasks indicate that no raters show significantly biased rater-task interactions. The outfit mean value by Rater 4 on

‘communicability’ indicates that Rater 4 did not consistently evaluate the task in the identified patterns of bias across all subjects.

7) *To what extent, statistically, is the task-based writing test a reliable and valid measure?*

#### (1) Reliability

All raters displayed acceptable levels of consistency with themselves, but there were significant differences among raters in terms of severity. The difference, based on fair average scores, is 0.44 of one grade in the scale, suggesting that there would be no impact on scores awarded in an operational setting. The analyses of the two tasks and the impressionistic scoring show that no significant difference occurs between the tasks and the impressionistic scoring. The scoring forms do not appear to separate the subjects to a significant degree. This means that in normal operations the three scoring forms can be considered equivalent.

#### (2) Validity

An estimate of the item discrimination was computed according to a “Generalized Partial Credit Model” approach. 1.0 is the expected value, but discriminations in the range 0.5 to 1.5 provide a reasonable fit to the Rasch model (Linacre, 2007, p.132). All the estimates fall in this range (1.04, 1.00, 1.24), which indicates that the randomness in the three sets of data fit the Rasch model. The two tasks and the impressionistic scoring were, therefore, of relevance to dependent data acquisition. There is also evidence that detracts from the measure’s validity. Table 10 shows the resulting correlation coefficients for the relationship between each of three raters’ scores and the *Criterion* score, and they were statistically significant ( $p < .01$ ) for Task 1, Task 2 and impressionistic scoring. This result supports the validity of the task-based writing test including these three scores.

Table 10 Inter-rater correlation coefficients between raters’ scores and the *Criterion* score

Rater	1	2	3	4	5	Avg.
T1	.67	.74	.78	.72	.68	.72
T2	.79	.67	.67	.64	.70	.70
IS	.68	.75	.81	.68	.76	.74

Note. T1=task1; T2=task 2; IS=impressionistic scoring

## 4. Discussion

The five junior high school teachers in this study were all novice raters in this task-based writing assessment. The interrater correlation coefficients between pairs of raters were relatively high, and the five raters appeared to be of acceptable reliability. The FACETS analysis showed that the raters displayed acceptable levels of consistency with

themselves. There were, however, significant differences between raters in terms of severity as we have seen in the pre-testing and the second testing. The bias analyses indicated that all raters were significantly biased towards certain types of subjects, and their bias patterns were unique. Moreover, it must be said that one rater-subject interaction and one rater-task interaction were identified as misfitting, so these raters were not consistent in the identified patterns of bias across the subjects or tasks. The FACETS analysis also showed that the difficulty of the two tasks and the impressionistic scoring were equivalent, indicating that the test tasks did not separate the subjects to a significant degree and thus the tasks measured the unified writing ability to be assumed as a construct. Since the 5-point scales demonstrated acceptable fit, the five categories and their specific written samples were mostly comprehensible and usable by raters. Therefore, it is quite likely that the assessment tasks and rating scales were reliable in determining an estimate of students’ writing ability.

The findings of this study suggest that the second edition of the TBWT Guide used in Main Experiment 1 may have effectively given novice raters a shared understanding of the construct of writing ability defined by the test writer and may have contributed to the consistency in scoring and the reduction in the biased interactions with tasks. It is, therefore, reasonable to suppose that the TBWT Guide for scoring may possibly reduce the differences or biases caused by variation among raters. In order to confirm this, a questionnaire was also administered to the five teacher-raters. The results of the questionnaire show that the five teacher-raters felt that the TBWT Guide was fairly useful. Whereas there seems to be an admission of improvement in the scale of accuracy and definition of communicability, the guidelines are supposed to lead the raters to self-consistency and reduction of biased interactions with tasks.

This present study, however, indicates that there were significant biased interactions with subjects’ ability among all five raters. Each rater was found to be self-consistent in scoring 20 subjects’ writing performances, but all of the raters had a unique bias pattern toward a certain type of subject. In the second testing, on the contrary, the number of the raters with a significant bias was three of the five (60%). In considering this difference, it must be assumed that the raters in the second testing were the same participants as in the pre-testing. The raters are, therefore, assumed to increase their internal consistency in assigning ratings as they gained experience. This point leads us to the comparative study of the second and the third testing, which suggests whether there are

significant differences between experienced teacher raters and novice teacher raters in scoring the same writing samples with the rating scales.

## 5. Conclusion

The results of the five novice raters in this study indicated that all raters displayed acceptable levels of consistency with themselves, and that the students' ability was effectively measured using these tasks and rating scales. The FACETS analysis showed that the difficulty of the two tasks and the impressionistic scoring were considered equivalent, which provided reasonable fit to the Rasch model. The modified scales associated with the five rating categories and their specific written samples were shown to be mostly comprehensible and usable by raters, and demonstrated acceptable fit. This is because the TBWT Guide for scoring has given novice raters a shared understanding of the construct of writing ability, and has contributed to the consistency in scoring as seen in the responses to the questionnaire.

There were, however, relatively small but significant differences between raters in terms of severity. The bias analyses of this study indicated all of the five raters were significantly biased towards certain types of subjects, while three of the five raters were biased in the second testing. When considering this difference, it must be assumed that the raters in the second testing were the same participants as in the pre-testing. The raters are, therefore, assumed to increase their internal consistency in assigning ratings as they gained experience. These findings suggest that the TBWT Guide for scoring may have contributed to the reduction of biased interactions, but training for certain raters with his/her unique bias patterns might still be required. This point leads us to the comparative study of the second and the third testing, which suggests there are significant differences between experienced teacher raters and novice teacher raters in scoring the same writing samples with the rating scales. This issue will be examined further in the comparison of the two main experiments

## References

- Alderson, J.C., Clapham, C. and Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Bachman, L.F. and Palmer, A.S. (1996). *Language testing in practice: designing and developing useful language tests*. Oxford: Oxford University Press.
- Bachman, L.F. (2002). Some reflections on task-based language performance assessment. *Language Testing* 19, 453-76.
- Brown, J.D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall Regents.
- Connor, U. and Carrell, P. (1993). The interpretation of tasks by writers and readers in holistically rated direct assessment of writing. In J.G. Carson & I. Leki (Eds.), *Reading in the composition classroom* (pp. 141-160). Boston, MA: Heinle and Heinle.
- Cumming, A. (1989). Writing expertise and second language proficiency. *Language Learning* 39, 81-141.
- Cumming, A. (1990). Expertise in evaluating second language composition. *Language Testing* 7, 31-51.
- Davies, A. (1990). *A Principles of language testing*. Basil Blackwell.
- Hyland, K. (2003). *Second language writing*. Cambridge: Cambridge University Press.
- Linacre, J. (2002). Guidelines for rating scales. Mesa Research Note 2 (Online). Available at <http://www.rasch.org/rn2.htm> (accessed 18 March 2008).
- Linacre, J. (2008). *Facets*, version no. 3.63. Computer program. Chicago, IL: MESA Press.
- Lunz, M., Wright, B. and Linacre, J. (1990). Measuring the impact of judge severity on examination scores. *Applied Measurement in Education* 3, 331-45.
- Pollitt, A. (1990). Response to Charles Alderson's paper: 'Bands and scores.' Alderson, J. C. (1991). In J. C. Alderson and B. North (Eds.), *Language Testing in the 1990s*. (pp. 87-91). London: Macmillan.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.
- Shohamy, E., Gordon, C. and Kraemer, R. (1992). The effect of raters' background and training on the reliability of direct writing tests. *Modern Language Journal* 76, 27-33.
- Sugita, Y. (2008). *Task-based performance assessment of Japanese second language writing*, Paper presented at 42<sup>nd</sup> International Annual IATEFL Conference, Exeter, the UK, April.
- Weigle, S. (1994). Effects of training on raters of ESL compositions. *Language Testing* 11, 197-223.
- Weigle, S. (1998). Using FACETS to model rater training effects. *Language Testing* 15, 263-87.
- Weigle, S. (2002). *Assessing writing*. Cambridge: Cambridge University Press.

# The Development and Implementation of Task-based Writing Performance Assessment for Japanese Learners of English: (5) Comparison of Two Main Experiments

Yoshihito Sugita

Faculty of Nursing, Yamanashi Prefectural University

sugita@yamanashi-ken.ac.jp

## Abstract

This study describes the analysis of composite ratings given by five high school teachers in Main Experiment 1 (OLD raters) and by five junior high school teachers in Main Experiment 2 (NEW raters). The purpose of this comparative study is to investigate the degree of difference in raters' severity, consistency and biased interactions between experienced and inexperienced raters. The analysis of comparative data reveals that all raters as a group differ significantly from one another in terms of severity, and most of the NEW raters were more severe than the OLD raters. Despite having previous experience with the TBWT, all raters displayed acceptable levels of consistency with themselves. The NEW raters are supposed to be less consistent, but a clear distinction between OLD and NEW raters can not be made as a group. According to the bias analyses, there was no difference in the number of rater-subject bias interactions between OLD and NEW rater groups, which shows the biases are caused by variation among raters, not by variation between the groups of the OLD and NEW raters. These findings suggest that the TBWT scoring guide significantly succeeded in giving raters a shared understanding of a definition of the ability being measured by the test, so that a distinction between OLD and NEW raters was not made despite the group difference in the rating experience.

## Keywords

task-based assessment, rating experience, group difference

## Introduction

The study described in this article involves the analysis of composite ratings given by five high school teachers in Main Experiment 1 (OLD raters) and by five junior high school teachers in Main Experiment 2 (NEW raters). The purpose of this comparative study is to investigate the degree of

difference in raters' severity, consistency and biased interactions between experienced and inexperienced raters. There were two reasons for comparing these two groups. First, the groups represent the two major backgrounds of typical teacher-raters in Japanese secondary schools, and thus constitute a sample that is representative of the population about which the study is intended to generalize. The second reason is to investigate any differences between inexperienced and experienced raters in terms of how they approach the ratings of each elicitation task.

## 2. The Study

### 2.1 Purposes and research questions

The purpose of this comparative study is to investigate the degree of difference in raters' severity, consistency and biased interactions between experienced and inexperienced raters. The following research questions were addressed:

- 1) To what extent do experienced and inexperienced raters differ in the spread of rater severity?
- 2) To what extent do experienced and inexperienced raters differ in consistency in their judgment?
- 3) Are experienced and inexperienced raters biased differently with respect to rater-subject interaction?
- 4) Are experienced and inexperienced raters biased differently with respect to rater-task interaction?

### 2.2 Participants

The participants in the research consisted of 10 Japanese teachers of English, 5 from each of the following groups:

#### *1 Experienced (Old) raters*

Raters who had rated TBWT using accuracy and communicability rating scales in the pre-testing (Pilot Experiment) had prior experience with the rating scales and the elicitation tasks (hereinafter referred to as OLD). This group consisted of five

high school teachers, all of whom shared similar backgrounds in terms of qualifications of ten or more years of teaching experience. They had also displayed acceptable levels of consistency with themselves in the pre-testing. Four of these raters were male and one was female. All were native speakers of Japanese.

### 2 Inexperienced (New) raters

Raters who had never rated TBWT, and thus had never been exposed to the rating scales and the elicitation tasks (hereinafter referred to as NEW) consisted of five junior high school teachers, all of whom shared similar backgrounds in terms of qualifications of ten or more years of teaching experience. Four of these raters were female and one was male. All were native speakers of Japanese.

## 2.3 Data analysis

Weigle (1998) conducted a study using FACETS to model rater training effects. She explored differences in rater severity and consistency among inexperienced and experienced raters both before and after rater training in an administration of UCLA's English as a Second Language Placement Examination (ESLPE). The FACETS analysis for the pre- and post-data were performed separately. In each analysis, four facets were used: examinee, raters, prompts, and scales. Although the purpose of her research is different from this comparative study, her analytical method is seen to be of great use for us in estimating the differences between experienced and inexperienced raters. Therefore, ratings in this study were also analyzed using FACETS, which provides comparative characteristic.

First, rater behavior both in Main Experiment 1 and Main Experiment 2 was modeled using FACETS, which provides estimates of the ability of 20 subjects, the difficulty of tasks and the severity of 10 raters on a common logit scale. In addition to providing logit estimates of the ability, severity, or difficulty of each element of each facet, FACETS also provides statistics indicating the relative spread of these estimates within each facet. In other words, the analysis provides information about the significance of any differences that may exist among elements of a facet, for example, differences in severity among raters or ability among subjects.

Another important feature of the FACETS analysis is that it provides fit statistics for each element, which provide an indication of the degree to which each element is behaving in a way that is predicted by the model. In the case of raters, the fit statistics are indicators of rater consistency. Thus, a detailed picture of the behavior of each rater in terms of both severity and consistency can be formed. In the FACETS analysis for the

comparative data, three facets were used: Subjects (n=20), Raters (five OLD, five NEW) and Tasks (accuracy, communicability and impression). The partial-credit model was chosen as well as in Main Experiments 1 and 2.

## 3. Results

### 3.1 FACETS summary

Measure Raters	+Subjects	+Tasks	S1   S2   S3
+ 7+	+	+	+ (5) + (5) + (5) +
+ 6+	9	+	+ + + +
+ 5+	+	+	+ + + +
+ 4+	18	+	+ + + +
+ 3+	12	+	+ + + +
+ 2+	6	+	+ + + +
+ 1+	14	7	+ + + +
+ 0	13	+	+ + + +
+ -1	2	+	+ + + +
+ -2	1	+	+ + + +
+ -3	10	+	+ + + +
+ -4	19	+	+ + + +
+ -5	1	+	+ + + +
+ -6	17	+	+ + + +
+ -7	15	+	+ + + +
+ -8	8	+	+ + + +
+ -9	3	+	+ + + +
+ -10	20	+	+ + + +
+ -11	5	+	+ + + +
+ -12	16	+	+ + + +
+ -13	4	+	+ + + +
+ -14	11	+	+ + + +
+ -15	+	+	+ (1) + (1) + (1) +
Measure Raters	+Subjects	+Tasks	S1   S2   S3

Figure 1 FACETS summary

Figure 1 shows a summary of all facets and their elements. They are positioned on a common logit scale which appears as "measure" in the first column. The second column shows the severity variation among raters. The most severe rater (ID: NEW5) is at the top, and the least severe rater (ID: OLD5) is at the bottom. The third column shows the ability variation among the 20 subjects. The subjects are ranked with high ability at the top (ID: 9) and low ability at the bottom (ID: 11). The fourth column shows that the difficulty variation among the tasks was equivalent. The last three columns graphically describe the three rating scales. Each of the two tasks and total impression has their own scale. The most likely scale score for each ability level is shown.

As the figure indicates, subject estimates range from a high of about 6 logits to a low of -6 logits,

indicating a spread of 12 logits in terms of students' ability. Subject separation value was 9.44, meaning that populations like these students in this study can be spread throughout about 9 levels. The reliability index was .99, which demonstrates it is possible to achieve reliable ability scores.

The column for raters shows that the NEW raters tend to cluster around the mean (0), and two OLD raters are more lenient than the rest. Thus, the figure indicates that the raters are not at the same level of severity.

### 3.2 FACETS analysis

1) *To what extent do experienced and inexperienced raters differ in the spread of rater severity?*

Table 1: FACETS analysis of rater characteristics

Rater	Fair-M average	Severity (logits)	Error	Infit (mean square)
NEW5	3.34	1.45	.25	.84
NEW2	3.25	1.14	.25	.86
OLD3	3.21	1.02	.25	.96
NEW1	3.18	.90	.25	.55
OLD2	3.09	.54	.25	1.05
NEW4	3.07	.48	.25	1.22
OLD1	3.04	.36	.25	.88
NEW3	2.92	-.11	.25	.72
OLD4	2.76	-.71	.25	.82
OLD5	2.55	-1.45	.25	1.18
Mean-All	3.30	.36	.25	.91
SD	.23	.85	.00	.19
Mean-OLD	2.93	-.05	.25	.98
SD	.24	.90	.00	.13
Mean-NEW	3.15	.77	.25	.84
SD	.14	.54	.00	.22

*All raters:* Separation: 3.29 Reliability=.92; fixed (all same) chi-square: 115.7, df: 9; significance: .00  
*OLD raters:* Separation: 3.50 Reliability=.92; fixed (all same) chi-square: 65.4, df: 4; significance: .00  
*NEW raters:* Separation: 1.96 Reliability=.79; fixed (all same) chi-square: 24.1, df: 4; significance: .00

Table 1 provides information on the characteristics of raters. From the left, each column shows rater IDs, fair average scores, rater severity, error and infit mean square values. Raters are presented in descending order of severity. The severity span between the most severe rater (ID: NEW5) and the most lenient rater (ID: OLD5) was 2.90; the difference, based on fair average scores in the first column, is 0.79 of one grade in the scale. Summary statistics for the entire sample of raters as well as for the two groups of raters (OLD and NEW) are also provided. In Table 1, the NEW raters are found to be more severe than the OLD raters. This suggests that the NEW raters tend to apply stricter standards overall to the written samples than the OLD raters. This tendency of the NEW raters is

also reflected in the mean severity estimates (.77 logits for NEW raters, -.05 logits for old raters).

The FACETS analysis provides a number of indications of the differences in severity among raters. These are the Separation Index, Reliability, and the Fixed (all identical) chi-square. The Separation Index of Table 1 is the ratio of the true standard deviation of raters to the root mean-square standard error. If the raters are equally severe, the standard deviation of the rater difficulty estimates should be equal to or smaller than the mean estimation error of the entire data set. However, the Rater Separation Index for the entire sample of raters is 3.29, indicating that the variance among raters is about three times as large as the error of estimates.

The Reliability statistics provided by the FACETS analysis indicate the degree to which the analysis reliably distinguishes between different levels of difficulty or severity among the different raters. For raters, low reliability is desirable, since ideally the different raters would be equally severe. In this case, however, the reliability is .92 for all raters, indicating that the analysis is fairly reliably separating raters into different levels of severity. Finally, the Fixed chi-square tests the null hypothesis that all the elements of the facet are equal. The chi-square of 115.7 with 9 d.f. is significant at  $p=.00$ , indicating that the null hypothesis must be rejected; in other words, the raters are not equally severe.

Comparing the range of severities for the two groups of raters, Table 1 makes clear the fact that the OLD raters as a group vary much more in severity than the NEW raters, with a standard deviation of .90 logits, compared with .54 for the NEW raters. This variability is also reflected in the Separation Indices for the two groups: 3.50 for the OLD raters, and 1.96 for the NEW raters. It should be noted that the NEW raters are more similar in severity despite their inexperience in using the rating scales and the elicitation tasks, indicating that the TBWT has the potential for reducing the spread of rater severity estimates among experienced and inexperienced raters, although both OLD and NEW raters differ significantly in their severity.

2) *To what extent do experienced and inexperienced raters differ in rater consistency in their judgement?*

The FACETS analysis also provides two measures of fit, or consistency: the infit and the outfit. The infit is the weighted mean-squared residual which is sensitive to unexpected responses near the point where decisions are being made, while the outfit is the unweighted mean-squared residual and is sensitive to extreme scores. For the purpose of this

study, only the infit statistics will be reported. Fit statistics of  $M+2SD$  or greater indicate too much unpredictability in raters' scores, while a value of  $M-2SD$  or less indicate overfit, or not enough variation in scores. Applying these standards to Table 1, all raters fall within the range of two standard deviations around the mean ( $0.91 \pm [2 \times 0.19]$ ), indicating that no raters were identified as misfitting. In other words, both OLD and NEW raters behaved consistently in scoring.

In terms of consistency, the infit statistics also show slight group differences. The standardized infit has an expected mean of 0 and a SD of 1. As for the OLD raters, its mean is .98 and the SD is .13, while these figures for the NEW raters are .84 and .22, respectively. Thus the NEW raters are supposed to be less consistent, but a clear distinction between OLD and NEW raters can not be made as a group.

### 3) *Are experienced and inexperienced raters biased differently with respect to rater-subject interaction?*

According to the results of the bias analysis in terms of interaction between rater severity and subject ability, there were six significantly biased interactions among OLD and NEW raters, respectively. Table 2 summarizes the frequencies of rater-subject interactions that displayed a significant bias for each rater group at various levels of the ability range.

Table 2: Frequency of rater-subject bias interaction

Ability	N	Harsh (Raters)		Lenient (Raters)	
		OLD	NEW	OLD	NEW
3.00 higher	4	1	2	2	
-2.99 ~2.99	11		2	2	2
-3.00 lower	5			1	

In Table 2, the first column shows the ability estimate range and the second column shows the number of subjects within the particular range of ability estimate. In the range of 3.00 or higher, the total number of subjects was four, but five biased interactions were observed. Four of the interactions were biased with the same subject. One OLD rater and two NEW raters harshly scored the subject, but two OLD raters leniently scored the subject. In the range between -2.99 and 2.99, there were six biased interactions. Two OLD and NEW raters leniently scored the subjects, respectively. Two NEW raters also harshly scored the subjects. One thing to be noticed is that one NEW rater (NEW4) was significantly biased with the three subjects in this range. In the range of -3.00 or lower, there was only one biased interaction among OLD raters, who harshly scored Subject 4.

The OLD raters leniently scored the subjects in five cases, which was 83% of the six rater-subject bias interactions ( $5/6=0.83$ ). On the contrary, the NEW raters harshly scored the subjects in four cases, which was 67% of the six rater-subject bias interactions ( $4/6=0.67$ ). These results indicate that there was no difference in the number of rater-subject bias interactions between OLD and NEW rater groups, but NEW raters might be more biased with the various types of subjects than the OLD raters, which show the NEW raters' inconsistency in scoring. Another finding is that the OLD raters were more lenient, and the NEW raters were more severe, as some previous research has shown the tendency of many of the NEW raters to judge more severely than the OLD raters (Weigle, 1994, 1998).

### 4) *Are experienced and inexperienced raters biased differently with respect to rater-task interaction?*

The results of the bias analysis in terms of the interaction between raters and tasks show that there were thirty rater-task interactions (10 raters  $\times$  3 tasks) in all, but only two interactions among OLD raters were significantly biased, OLD5  $\times$  Communicability and OLD3  $\times$  Accuracy. The bias analysis indicates that the two raters tended to award lenient ratings to all the subjects on each task. This may be a consequence of the tendency of many of the OLD raters to judge more leniently than the NEW raters, and thus have ended up consistently evaluating the task in the identified patterns of bias across all subjects. The value by NEW4 on 'communicability' and the values by OLD2 and OLD5 on 'accuracy' indicate that these raters were not consistent in evaluating the tasks in the identified patterns of bias across all subjects.

## 4. Discussion

### 4.1 Implications for judging severity

The ten raters in this study were all experienced Japanese secondary school teachers of English and shared similar backgrounds in terms of qualifications of ten or more years of teaching experience. The results of the present study, however, indicate an influence by errors of measurement resulting from variation in rater severity. According to McNamara (1996), "raters typically show highly reliable differences in severity, indicating that elimination of rater differences in severity is an unachievable goal" (p.140). This statement is supported by previous empirical studies that have shown that relative severity of individual raters could not be reduced to acceptable levels even after rigorous rater training or among thoroughly trained raters (Luntz et al., 1990; Weigle, 1998). These findings justify the use

of FACETS analysis, which assumes, makes use of and compensates for inter-rater differences.

The rater-subject bias analysis shows particular raters' tendency toward excessive severity or leniency towards particular subjects. The OLD raters exhibited a tendency towards leniency, while the NEW raters awarded unexpectedly severe ratings to the most subjects. The one NEW and three OLD raters strongly disagreed on Subject 12. One NEW rater had three biased interactions with the subjects in the range between -2.99 and 2.99. These results suggest that the acceptable overall fit values for both OLD and NEW raters (Table 1), despite the possibly problematic rater behavior suggested by bias analysis, are plausible results as bias analysis looks at individual raters, not the differences between the groups of the OLD and NEW raters. Therefore, as confirmed in the previous testing (Main Experiments 1 and 2), if the raters behave consistently, the TBWT scoring guide significantly succeeds in giving raters a clear understanding of a definition of the ability being measured by the test. Although the scoring guide may have limitations in reducing the differences or biases caused by variation among raters, this view of the function of the scoring guide addresses the concern that a shared understanding of the construct of writing ability could effectively be promoted by the scoring guide.

#### 4.2 Implications for task difficulty

A task was one source of score variance in the TBWT, but the variance was negligible in terms of difficulty. This finding was supported by good evidence of fit to the Rasch model for all elements of this facet. Tasks in the form of five rating categories were found to show acceptable fit, and the scales associated with the five rating categories were shown to be mostly comprehensible and usable by raters. However, it must be said that rater and scale reliability deserve a statistical treatment reflecting the extended observation. Given the limited amount of data available in this study, it is difficult to speculate about what may occur in a given situation. Many more similar analyses will be necessary to see what is likely to occur.

Of the two instances of biased rater-task interactions, one occurred with Task 2 (Communicability), which was scored by OLD5. As it was observed in Main Experiment 1, OLD5 (Rater 5) was the only rater that displayed a biased interaction with the tasks and tended to become more lenient when the script displayed adequate communicative effect. This suggests that the effect is insufficiently reliable to indicate systematic bias of tasks, but that tasks are more likely to affect individual raters differently, rather than at the group

level.

These somewhat inconclusive results suggest that further exploration is warranted, although the expectation is higher that these findings will be confirmed rather than the discovery of major differences. This study offers further evidence that task development based on the construct-based processing approach could be a reasonably solid basis to estimate students' writing ability, and those tasks may draw valid inferences to their writing performance.

#### 5. Conclusion

The analysis of comparative data reveals that all raters as a group differ significantly from one another in terms of severity, and most of the NEW raters were more severe than the OLD raters. Despite having previous experience with the TBWT, all raters displayed acceptable levels of consistency with themselves. The NEW raters are supposed to be less consistent, but a clear distinction between OLD and NEW raters can not be made as a group. According to the bias analyses, there was no difference in the number of rater-subject bias interactions between OLD and NEW rater groups, which shows the biases are caused by variation among raters, not by variation between the groups of the OLD and NEW raters. These findings suggest that the TBWT scoring guide significantly succeeded in giving raters a shared understanding of a definition of the ability being measured by the test, so that a distinction between OLD and NEW raters was not made despite the group difference in the rating experience.

Tasks in the form of five rating categories were found to show acceptable fit to the Rasch model, and the scales associated with the five rating categories were shown to be mostly comprehensible and usable by raters. However, it must be said that rater and scale reliability deserves a statistical treatment reflecting the extended observation. Given the limited amount of data available in this study, it is difficult to speculate about what may occur in a given situation. Many more similar analyses will be necessary to see what is likely to occur.

#### References

- Alderson, J.C., Clapham, C. and Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Bachman, L.F. and Palmer, A.S. (1996). *Language testing in practice: designing and developing useful language tests*. Oxford: Oxford University Press.

- Bachman, L.F. (2002). Some reflections on task-based language performance assessment. *Language Testing* 19, 453-76.
- Brown, J.D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall Regents.
- Connor, U. and Carrell, P. (1993). The interpretation of tasks by writers and readers in holistically rated direct assessment of writing. In J.G. Carson & I. Leki (Eds.), *Reading in the composition classroom* (pp. 141-160). Boston, MA: Heinle and Heinle.
- Cumming, A. (1989). Writing expertise and second language proficiency. *Language Learning* 39, 81-141.
- Cumming, A. (1990). Expertise in evaluating second language composition. *Language Testing* 7, 31-51.
- Davies, A. (1990). *A Principles of language testing*. Basil Blackwell.
- Hyland, K. (2003). *Second language writing*. Cambridge: Cambridge University Press.
- Linacre, J. (2002). Guidelines for rating scales. Mesa Research Note 2 (Online). Available at <http://www.rasch.org/rn2.htm> (accessed 18 March 2008).
- Linacre, J. (2008). *Facets*, version no. 3.63. Computer program. Chicago, IL: MESA Press.
- Lunz, M., Wright, B. and Linacre, J. (1990). Measuring the impact of judge severity on examination scores. *Applied Measurement in Education* 3, 331-45.
- Pollitt, A. (1990). Response to Charles Alderson's paper: 'Bands and scores.' Alderson, J. C. (1991). In J. C. Alderson and B. North (Eds.), *Language Testing in the 1990s*. (pp. 87-91). London: Macmillan.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.
- Shohamy, E., Gordon, C. and Kraemer, R. (1992). The effect of raters' background and training on the reliability of direct writing tests. *Modern Language Journal* 76, 27-33.
- Sugita, Y. (2008). *Task-based performance assessment of Japanese second language writing*, Paper presented at 42<sup>nd</sup> International Annual IATEFL Conference, Exeter, the UK, April.
- Weigle, S. (1994). Effects of training on raters of ESL compositions. *Language Testing* 11, 197-223.
- Weigle, S. (1998). Using FACETS to model rater training effects. *Language Testing* 15, 263-87.
- Weigle, S. (2002). *Assessing writing*. Cambridge: Cambridge University Press.

# A Qualitative Look at Japanese and Native Speaker Teachers' Views on the University EFL Classroom

Richard Silver, Satoko Ito, and William Baber

Language Education Center (BKC), Ritsumeikan University

richinwit@hotmail.com, s-ito@fc.ritsumeai.ac.jp, willbaber@yahoo.com

## Abstract

In this paper we undertake to provide a snapshot of current beliefs of native speaking and non-native speaking teachers of English as a foreign language in universities in Japan. This qualitative study provides a current look into this web of issues. The study material is based on interviews gathered by the research team, transcribed, summarized, and coded according to themes developed from current literature on teacher beliefs, knowledge and cognition. Nineteen interviews, eight of Japanese speaking and eleven of English speaking teachers, form a body from which themes are identified and developed. Particular attention is given to teacher roles, teacher perceptions of institutions and learners, and the impact of teacher experience on their views and actions. Findings include the attunement of teachers to student in-class needs, the perceived importance of the teacher as motivator, sophistication about teacher roles vis-a-vis institutions and students, and a high level of cultural sensitivity towards the student experience and context from non-Japanese teachers.

## Keywords

Classroom management; teacher roles; culture; teacher beliefs; TEFL; EFL; Japan; Higher education; teacher beliefs; teacher cognition; acculturation

## Introduction

The ongoing debate within Second Language Acquisition over the roles of the teacher in English as a Foreign Language (EFL) classrooms has resulted in considerable research internationally, though little has been conducted in Japan.

This research recognizes that EFL teaching is a continually evolving process, as new generations of teachers enter the profession bringing new or rediscovered techniques and

beliefs with them. Through interviews with Japanese non-native speaking (NNS) and native speaking (NS) teachers of EFL in Japanese universities, this research attempts to capture the present moment in Japan.

Given that teacher beliefs and behaviors in the classroom often differ, particular attention was given to the interplay of beliefs and actions. Further attention is directed to the experiences and training of the educator in Japan and abroad. The comments generated will lead to a survey aiming to tease out detail regarding particular systems or practices teachers adopt in the EFL classroom in order to produce their desired teaching goals and manage classroom behavior.

## 1 Background and literature review

### 1.1 Teacher classroom behaviors and beliefs, knowledge, and cognition

There has been a general understanding that “what teachers do is a reflection of what they know and believe, and that teacher knowledge and ‘teacher thinking’ provide the underlying framework or schema which guides the teacher’s classroom actions” (Richards & Lockhart, 1996: 29). Further, the links between the teachers’ classroom behaviors and their beliefs and assumptions about learners and teaching have been investigated by many preceding research efforts. Due to the complexity of constituents that form a teacher’s mental constructs, research in this field has been adopting increasingly holistic perspectives when looking at teachers’ mental constructs. Accordingly, numerous terms such as *teacher beliefs* or *teacher knowledge* have been used with different definitions describing similar concepts, inevitably bringing about a certain amount of confusion in the research field.

Reviewing and outlining the developments of a wide range of teacher cognition research in the past several decades, S. Borg writes that “the predominant focus today is on understanding

teacher knowledge (used as an umbrella term for a range of psychological constructs), its growth and use” (S. Borg, 2006: 35), and adopts a collective term *teacher cognition* instead of teacher beliefs or teacher knowledge to refer to a concept that embraces various mental constructs of teachers, which can be “characterized as an often tacit, personally-held, practical system of mental constructs held by teachers and which are dynamic – i.e., defined and refined on the basis of educational and professional experiences throughout teachers’ lives” (S. Borg, 2006: 35). Although efforts have been made to distinguish the concept of belief and that of knowledge,<sup>1</sup> given that any clear distinction between the two is very delicate, the current study follows S. Borg’s concept of teacher cognition to discuss the various mental constituents labeled as teacher beliefs or teacher knowledge.

## 1.2 Socio-cultural factors

Among the constituents that affect teachers’ cognition, the current study has given special attention to socio-cultural factors due to its investigation into the Japanese university context. A variety of research seems to agree that Japanese linguistically and culturally is distant from English and English-speaking culture: “In addition to the difficulties posed by great grammatical, lexical, and phonetic disparity, Japanese speakers’ attitudes to language in general are heavily coloured by aspects of their own language....Given these striking differences between Japanese and English attitudes to language, it takes the student a good while to tune in” (Thompson, 2001: 296-97). Sociolinguists approach this phenomenon by employing the dominant frameworks provided by Hofstede (1986) or Hall (1966), most of which agree that Japanese and Japanese culture stand at almost opposite positions from English and English speaking culture in the dimensions provided such as *individualism vs. collectivism*, *masculine culture vs. feminine culture*, the degrees and extent to accept *power distance*, and those of *uncertainty avoidance* (Brown, 2000: 176-200). Given that teaching contexts have an important influence on teachers’ decisions about their classroom management, the current study considers NS and NNS teacher’s experiences of acculturation and their attitudes toward and

expectations of Japanese EFL learners, and further their actual and assumed teacher roles in the classroom (Richards & Lockhart, 1996: 107-108; Scharle & Szabó, 2000: 5-7).

## 1.3 Teacher roles

To investigate the influence of teacher cognition on classroom management, the current study looks into the links between actual or assumed teacher roles and the systems teachers employed to manage and control their classrooms based on the framework provided by Richards and Lockhart (1996). They argue that teachers’ assumptions about their teacher roles in the classrooms influence how they respond to: classroom management and organization; teacher control; curriculum, content, and planning; instructional strategies; motivational techniques; assessment philosophy (Richards & Lockhart, 1996: 106). Considering factors in these dimensions, they set eight different types of general teacher roles as examples:

**Planner:** The teacher sees planning and structuring of learning activities as fundamental to success in teaching and learning.

**Manager:** The teacher’s role is to organize and manage the classroom environment and student behavior in a way that will maximize learning.

**Quality controller:** A central task for the teacher is to maintain the quality of language use in the classroom. Correct language use should be reinforced and incorrect use discouraged.

**Group organizer:** The teacher’s role is to develop an environment in which students work cooperatively on group tasks.

**Facilitator:** The teacher’s role is to help students discover their own ways of learning and to work independently.

**Empowerer:** The teacher tries to take as little control or direction over the lesson as possible and lets the students make decisions about what they want to learn and how they want to learn it.

**Motivator:** The teacher seeks to improve students’ confidence and interest in learning and to build a classroom climate that will motivate students.

**Team member:** The teacher and all the students in the class constitute a team and should interact like members of a team.

(Richards & Lockhart, 1996: 105-106)

These eight roles do not represent a complete

<sup>1</sup> M. Borg (2001), for example, distinguishes belief and knowledge as the latter being justified as true in some external sense.

list of teacher roles that emerge in actual classrooms and this study. Although these roles are expected to overlap with each other depending on the shifting situations and contexts of any given classroom, the current study considers teacher responses to each role to capture some general views of the roles and background factors that may affect the choices of teachers.

This research attempts to characterize the following broad question with consideration of culture, acculturation, teaching experience, and learning experience, and teacher roles: What are NNS and NS teacher views of issues and actors in university EFL classrooms?

## **2 Methodology**

### **2.1 Participants**

In total, nineteen university English instructors were interviewed. Interviews lasted from about forty-five minutes up to an hour and a half. The researchers sought to interview a broad range of teachers, thus the participants consisted of both NS and NNS teachers of English. Ages ranged from 28 to 60 and length of experience teaching in Japan from 4 to over 30 years. The teachers interviewed were at the time working at numerous universities in Kansai and held positions as part-time teachers, contract teachers and full time faculty members, some in management roles.

### **2.2 Ethics**

All participants were informed as to the nature and purpose of the research and gave their express permission to be recorded. None of the teachers approached refused. Anonymity was assured for all participants except among the investigators where it was necessary to avoid approaching the same person. No names were attached to any of the recordings, notes, or transcripts.

### **2.3 Data collection and analysis**

All interviews were recorded using a portable digital device that was visible to the participants throughout the interview. NS teachers were recorded in English and NNS teachers in Japanese. The recordings were then transcribed or summarized and coded. All coded material was shared and reviewed by the researchers. In addition, English summaries of the Japanese interviews were prepared and shared among the researchers.

### **2.4 Procedure**

Richards and Lockhart's framework for teacher's roles (Richards & Lockhart, 1996) was used in the interviews as a comparison with established research. The questions were agreed upon and tested among the researchers before the first interview. While the same instrument (English and Japanese versions) was used by all interviewers, interviewers used the responses of past interviewees to gain a fuller general picture of teacher's views through follow-up and clarification questions. In addition, the views of the other interviewees were put to new interviewees to try to identify general and unique themes. While this study cannot claim to be a full grounded theory study, it does approach the subject of investigation using many of the tools and techniques as described by Strauss and Corbin (1998), and Senior (2006). The research question and supporting questions were posed without a hypothesis and interviews were semi-open ended though focused on the areas that the researchers sought to examine. Moreover, while a grounded theory study aims to create theory from data, when accepting the time limits on this preliminary study, the researchers hoped to generate further avenues for investigation within the context of the Japanese university classroom and teacher's beliefs. During discussions of the data, researchers compared their experiences of the interviews, their interpretations of answers and their knowledge of the subjects to examine and cross examine the possibilities presented therein. During the process the codes were pared down and possible follow-up questions were formulated in the event of second interviews.

## **3 Analyses and discussion**

### **3.1 Classroom roles**

#### **3.1.1 Teacher classroom roles accepted**

Generally the interviewees accepted several overlapping roles in the classroom. Among all teachers the most commonly cited role was that of motivator. The general outlook was of motivator as a role they felt was a leading, though not a sole role. Commonly selected supporting roles included planner, manager, and facilitator, though some NNS teachers particularly downplayed the role of planner. Some interviewees also expressed a desire to be an empowerer, though others rejected it. Some chose facilitator with caution, one saying, "Wish

I could be more.” Another would have chosen quality controller if the level of their students were higher.

### 3.1.2 Teacher classroom roles rejected

Commonly rejected roles included lecturer, team member, “Sounds nice...not here!”, and quality controller, “That sounds a bit heavy-handed to me.” The reasons cited were that these roles were inapplicable or even undesirable in the Japanese university classroom.

### 3.1.3 Teacher as a representative of the university

Teachers in general appear to accept their position as being part of a greater institution. For example, comments such as “I cannot be demotivatingly tough” and “I must respect that the university needs happy students and parents” were found.

### 3.1.4 Teacher as a negotiator of curriculum

In dealing with the curriculum most interviewees accepted the need for an overall structure to their classes to be imposed from above both for students and for teachers. While not unhappy, most found it necessary to adjust that curriculum. Many posited the curriculum as necessarily being a flexible space in which the institution must leave enough room for teachers to express their character. Most teachers in managerial roles commented that the system must leave room for teachers to make adjustments for assessment and in the class.

In general teachers felt that standardization was appropriate. However, more recent teachers commonly expressed a desire to take more control over their classes once they had discovered the level of their students, but with guidance from a department. While older teachers cited the lack of unity in the past, in particular with teachers choosing their own textbooks, younger teachers who may only have experienced choices being made for them, were less satisfied with uniformity in the curriculum or institution but still recognized the need for balance. (See 3.2.1)

There appeared to be a generational divide among NS teachers with older teachers comfortable with curricular uniformity applied by the institution whereas younger NS teachers were less so. Among NNS teachers, younger ones were somewhat less accepting of institutional uniformity, though all ages saw a

balance between institutional guidelines and space for maneuver in the classroom as preferable.

## 3.2 Teacher roles beyond the classroom

### 3.2.1 Textbook choosers

Interviewees not in management roles liked the textbook to be chosen by someone with knowledge of the students, but repeated criticism surfaced of inappropriate textbooks, “If you’re given a textbook obviously the negative is ...if it’s not very good you’re stuck with it. I would ... unambiguously like to choose the textbook for the second semester.”

### 3.2.2 Curriculum designers

Managers were accepting of curriculum design, despite some caution about criteria imposed from above. Meanwhile teachers not in management roles adapted to the curriculum situations presented to them. Nonetheless there were desires evident for varying degrees of freedom (See 3.1.4).

### 3.2.3 Goal creators

Some but not all generally rejected the need to take on the role of setting goals for students beyond the immediate classroom.

### 3.2.4 Content decision maker

Often teachers indicated a desire for less control in their first years at an institution, with more control coming thereafter. Additionally, many teachers did not want complete responsibility, as one interviewee, related another teacher’s experience of being told, “Do what you want,” and responding, “Well, I don’t want that!” Others specifically wanted to build experience with the student body, as indicated here, “[In] my first year, I am...testing how students respond...my experiences in my first semester...will be a great influence in terms of how I structure things in the next semester.”

## 3.3 Views on current Japanese university education

### 3.3.1 View of students as customers

While teachers did not reject the term outright, in general they did not view students as retail customers but as “a type” of customer in varying respects. However, some teachers are aware of fees, so the teacher tries to satisfy the distant parental customer while representing the

institution. This awareness caused some teachers to feel a professional responsibility or accountability toward private or publicly funded institutions or students. Comments included: “They are paying for it,” “Their parents are paying for it,” and “They are paying for opportunity.”

### 3.3.2 View of students’ academic ability and attitude

Some NNS in comparison to their own experience described current students as lacking passion, or *don’yokusa* (食欲さ). However, others suggested that learners had not fundamentally changed from past decades.

All NS made it clear that there was a strong difference in the experience of Japanese students and their experience as students in the “west” regarding academic rigor and behavior. Nonetheless, NS teachers saw this disparity in a more complex cultural context which allows them to integrate prevailing student attitudes. To paraphrase one, “a lot of people (NS) say it’s a passive atmosphere, but...not passive, a different approach.”

NS teachers generally found students weak in critical thinking and their learning superficial. The interviews did not reflect high opinions of undergraduate ability, focus, and output. Nonetheless, NS teacher views of students were not negative, but sympathetic and often based on some understanding of the cultural experience of students in Japan.

### 3.3.3 View of students as well rounded individuals

Though keen not to overstep the boundaries of their teacher roles, teachers did perceive students as developing characters. Some teachers acknowledged the changing circumstances of their students, for example living alone and greater freedom, and could respond to that change in their classes by incorporating more group work and communicative activities. Generally, teachers take actions in class to engage and support students on a humane level.

### 3.3.4 View of institution

NS teachers in general were unwilling to criticize institutions outright and commented on the difficulty of making comparisons with their own experiences as students. (See 3.3.2) Some also cited language barriers and a lack of

knowledge about their students’ majors. With these caveats, general criticisms included: critical thinking and deep understanding not fostered by institutions; too much academic freedom; and “downright lax.” Another NS said, “Clearly there are some deeply entrenched problems...in my opinion they are hard to fix.” However this teacher continued, preferring that *daigaku* and universities not be held equal or even compared, but that *daigaku* be understood in their own context.

Similarly, NNS teachers commented on the low expectations of university from society. These teachers were able to understand the Japanese university experience within the context of Japan. Their comments included: expectations set on students are lower than the expectations that may be placed on western universities by western societies; therefore it is difficult for Japanese students and universities to change.

## 3.4 Interplay of beliefs and actions

### 3.4.1 The classroom space

Teachers view the classroom as a space wherein teachers and students interact on a human level. (See 3.1.1)

### 3.4.2 Teacher boundaries

Teachers see a line between acceptable teaching of behavior (in terms of those skills related to language learning) and unacceptable teaching (doing things that are outside of their mandate). However the line was different for different teachers. Both NS teachers and NNS teachers appear unwilling to take a commanding role in terms of students’ development or a peer role (team member) beyond the scope of their teaching mandate. NS teachers avoided these roles due to uncertainty about Japanese culture.

### 3.4.3 Students as stimulus for action

For all teachers interviewed, interaction with students in class is the main influence on action and adjustments by teachers. (See 3.5.1)

## 3.5 Educational training and experience

### 3.5.1 Influence of experience

The most often cited influence for changing the way a class was taught was students in the classroom. (See also 3.4.3)

### 3.5.2 Influence of teacher’s language learning

experience

The interviews revealed that NS teachers generally felt influenced by their own learning experience and sought to incorporate their experiences in classes. As one said, "I recall what worked or did not work for me and I use or avoid accordingly." Quite the opposite, many NNS teachers said their own learning experience was not very helpful in their current teaching because "curriculum and environment were very different then compared to now" or because "English classes I had were for English (or relevant) majors."

### 3.5.3 Influence of teaching theory

Teachers said they referred to teaching theory but that it was not in the foreground of their thoughts when they taught. Some specifically said it influenced them not day to day but semester to semester/over longer periods. Nonetheless, most teachers identified easily with the teacher roles presented in the interview instrument. Thus it appears that theory is present in a significant way, though not in a manner identifiable to the interviewees.

### 3.6 Influence of cultural background and length of time in Japan

Some native speakers showed a high level of understanding of Japanese students and institutions connected in a non-simplified cultural context.

### Conclusion

The following conclusions refer to university level teachers of English in Japan, NS teachers, and NNS teachers, the target of this investigation. These conclusions are generalizations that may have limited application to the broader population of English teachers in Japan. The extent to which these conclusions can be generalized and more deeply analyzed will be the focus of a survey to be conducted by the authors:

1. NS teachers see students and institutions in a culturally complex context rather than as simplified or culturally un-integrated actors.
2. University English teachers in Japan have clear self images as educators and about their classroom roles. This group is sophisticated in their understanding of their work and behaviors.
3. Teachers prefer a certain level of freedom in the classroom but recognize the need for

institutional or curriculum level uniformity and guidance.

4. Teachers see motivation as a primary issue, one that they, not the institution, must resolve in the classroom.
5. Teachers generally have systems that they employ in the classroom.
6. Teachers adjust their actions and approaches based most of all upon the students in any given class.
7. Some teachers rejected the role of setting goals for students beyond the immediate classroom. On the other hand, others felt that students could and should learn skills and behaviors applicable to post-university life.

### References and bibliography

- Borg, M. (2001). Teachers' beliefs. *ELT Journal*, 55, 2, 186-188.
- Borg, S. (2006). *Teacher Cognition and Language Education*. NY: Continuum.
- Brown, H. G. (2000). *Principles of Language Learning and Teaching* (4th ed.). NY: Longman.
- Hall, E. T. (1966). *The Hidden Dimension*. NY: Doubleday.
- Hofstede, G. (1986). Cultural differences in teaching and learning. *International Journal of Intercultural Relations*, 10, 301-320.
- Richards, J. & Lockhart, C. (1996). *Reflective Teaching in Second Language Classroom*. Cambridge: CUP.
- Scharle, Á. & Szabó, A. (2000). *Learner Autonomy*. Cambridge: CUP.
- Senior, R. M. (2006). *The Experience of Language Teaching*. Cambridge: CUP.
- Strauss, A. & Corbin, J. (1998). *Basics of Qualitative Research* (2nd ed.). Thousand Oaks: Sage Publisher.
- Thompson, I. (2001). Japanese speakers. In M. Swan & B. Smith (Eds.), *Learner English* (2nd ed.), 296-309. Cambridge: CUP.

# Consonant clusters by Japanese learners of English

Aya Kitagawa

School of Education, Waseda University

aya@fuji.waseda.jp

## Abstract

The present study aims to investigate the consonant cluster produced by Japanese learners of English. The speakers who had never taken formal training in pronunciation participated in this experiment in order to reveal which consonant clusters are naturally acquired and which are not without explicit instruction. The results showed that the difficulty depended on the component of the consonant cluster.

## Keywords

consonant clusters, acoustic analysis

## 1 Introduction

### 1.1 Consonant clusters produced by Japanese learners of English

The consonant cluster is a sequence of the consonants within the onset or the coda, and both the number of consonants and the component allowed in the cluster are language-dependent. For instance, the template of the consonant cluster in English and Japanese can be shown as follows;  $C_0^3VC_0^4$  for English and  $C_0^2V$  for Japanese. Furthermore, the types of the consonant cluster permitted in Japanese are only geminates and contracted sounds, such as /katta/ (won) and /kjou/ (today), whereas English holds more variable types of the consonant cluster.

Due to this difference of the consonant cluster among languages, language learners are generally believed to have some difficulty in pronouncing the sequence of consonants which does not exist in their native languages. The same is true of Japanese learners of English since, as mentioned previously, the consonant clusters allowed in Japanese are far more limited than those in English. What is more, many of the textbooks to instruct the English pronunciation for the native speakers of Japanese point out that they tend to insert the vowels between the consonants of the cluster on the ground that there is a difference in the basic syllable structure between English and Japanese. (Igarashi, 1981; Tori & Kaneko, 1962).

### 1.2 Possible errors by the non-natives

One of the previous studies which suggest how the non-native speakers repair the consonant clusters was carried out by Davidson (2006). She conducted research on pseudo-Czech consonant clusters spoken by native speakers of English, using the clusters whose first segment was /s/, /f/, /z/ or /v/ and whose second segment was an obstruent or a nasal. In her study, she categorized the language learners' responses to the target consonant clusters into the six types: correct, deletion, prothesis, vowel insertion, segment change and others. "Others" held the targets which were not produced, had more than one error, or was impossible to recognize. As a result, she found that the most frequent error was vowel insertion and it occurred to all types of the consonant clusters, most frequently to the initial-/v/ consonant clusters. The second largest error type was prothesis, and the third was segment change.

### 1.3 Articulation of the consonant cluster

Some of the consonant clusters are articulated in a particular way, which is different from when the consonant is solely pronounced. They include incomplete release, nasal release, lateral release, similitude and epenthesis. Firstly, incomplete release is related with the plosive plus plosive consonant cluster, whose first segment is not audibly released. To take an example for this, "act" is pronounced with the release of /k/ inaudible. Secondly, nasal release happens to the /tn/ and /dn/ consonant clusters, where the plosive of the first segment is released during /n/. Thirdly, lateral release is observed in the production of the /tl/ and /dl/ consonant clusters and is particularly common in British English. Similar to nasal release, the plosive is released during /l/. Fourthly, similitude is the phenomena that the feature of one segment influences that of the other. The example of this is the voiceless plosive plus voiced approximant consonant cluster, like /tr/ in "tree." For this articulation, the voicing of the first segment affects that of the second, which makes the approximant of the second segment devoiced. Finally, epenthesis occurs to the /ns/ consonant cluster at the end of the word, as seen in "dance." Since /n/ is a stop and /s/

is homorganic to /n/, it is natural that /t/ is inserted between /n/ and /s/ when the closure at the alveolar ridge is released.

#### 1.4 Research questions

The ultimate goal of this study is to develop a way to teach English pronunciation effectively and efficiently to Japanese learners of English. Then, the present study deals with the consonant cluster, addressing two research questions. The first is which types of consonant clusters are difficult for Japanese learners of English and which are easy. The second is what acoustic characteristics the consonant clusters produced by Japanese learners of English have.

## 2 Methodology

### 2.1 Subjects

Twenty-five high school students in the elementary English class participated in this experiment. They had never lived in a foreign country for a long period of time. None of them also had received formal training in English pronunciation.

### 2.2 Materials

One of the phonetically balanced passages, "Arthur the Rat," was used in this experiment, and the consonant clusters were picked up from the passage. The tested tokens were as follows; asked [1], looked [1], cooped [1], stay [3], stood [3], stamping [1], scouts [2], snug [1], joists [2], coughed [1], loft [2], rafters [2], exactly [1] and next [1]. The number within the square brackets refers to how many times the tokens were repeated in the passage.

The consonant clusters which accompanied nasal release, lateral release, similitude and epenthesis were also investigated, but are not reported in this paper. This is because Japanese learners of English have difficulties in articulating these tokens more at the segmental level than at the cluster level: the confusion of /l/ and /r/ and the substitution of nasalized vowels for nasals in a certain context. Further study is, therefore, needed to reveal the details of the production of these consonant clusters. Thus, only the tokens above are dealt with in this study.

### 2.3 Recording

The subjects first obtained instructions to comprehend the vocabulary and the context of the passage, and then were provided time, from a month to two months, to practice reading it aloud.

The speech was recorded using a digital recorder, Roland R-09, and a condenser microphone, SONY ECM-MS957. The recording was carried out in a small room, which was not soundproof but quiet.

## 2.4 Data analysis

First of all, using the acoustic analysis software, Praat, the segmentation was provided to each speech with the spectrogram and the waveform. Next, the tested tokens were categorized into the six types; correct, deletion, prothesis, vowel insertion, segment change and others. Then, when the error was identified with vowel insertion or prothesis, the duration and the F1 and F2 of the added vowels were measured. Lastly, using SPSS, the statistical tests, a one-way repeated-measured ANOVA and a paired samples t-test, were performed.

Several tested tokens were repeated more than once in the speech. In the case that the subjects repaired these consonant clusters in a different way as the tokens were produced, the most frequent strategy of repairing was chosen for the analysis. Moreover, the tokens produced correctly even once were considered to be correct, even if the subjects produced the tokens incorrectly in another time. It must be also noted that the categorization of the target tokens was based on the consonant clusters not on the word. Hence, the token was regarded to be correct as long as the target consonant clusters were produced correctly if it had no or one error at the word level. Consequently, the token "scouts" pronounced as [skouts], was, for example, divided into "correct" when /sk/ was accurate.

Besides, the addition of a sound after the word-end consonant, or paragoge, happened to the many subjects, but this type of errors was not analyzed here because this response is common to the consonant at the end of the word, not specific to the consonant cluster. Therefore, paragoge is not mentioned in this paper if it is not needed.

In addition, the values of the duration and the F1 and F2 were normalized in order to make it possible to compare the data across the subjects. The duration of the inserted vowels was calculated by dividing it by the duration of the word. As for the normalization of the F1 and F2 values, this study applied the formula suggested by Labanov (1971):  $F_i^N = (F_i - M_i) / \sigma_i$ .

## 3 Results and Discussion

### 3.1 The error types

#### 3.1.1 The plosive + plosive cluster

Table 1 summarizes the results of how the subjects produced the plosive plus plosive consonant clusters. As argued in the introduction, the native speakers generally release the first plosive inaudibly; therefore, the category "correct" is made up of two subcategories: incomplete release and complete release. When the subjects produced this type of the consonant clusters without errors but with audible release, the tokens were classified into the category "correct but complete release."

Table 1: The plosive + plosive clusters

	C		D	Pr	In	SC	O
	IR	CR					
asked	0	6 (2)	4	0	11	11	1
looked	0	9 (4)	0	0	6	14	2
cooped	0	10 (2)	0	0	7	8	4

Note: C, D, Pr, In, SC, O, IR and CR stand for correct, deletion, prothesis, vowel insertion, segment change, others, incomplete release and complete release respectively. The number within the parenthesis shows how many tokens paragogue happened to.

As illustrated in Table 1, none of the subjects released the plosive of the first segment inaudibly. Although four tokens for “asked,” five for “looked” and eight for “cooped” were produced correctly without any critical errors, their first segments were audibly released. Additionally, according to the statistical test, there was no significant difference in the correct rate among the three tokens,  $F(1.82, 34.55) = 1.95, p > .05$ .

As far as the errors are concerned, the most frequent way to repair the plosive plus plosive consonant clusters was the segment change for all the tokens, and the second was the vowel insertion. The deletion was the third largest error type for “asked,” which occurred to the second segment for all of the four related tokens. More noteworthy is, however, that the subjects often repaired the consonant clusters using more than one strategy. This is why some consonant clusters were repaired by changing a segment and inserting a vowel between the first and second segments of the consonant clusters at the same time, as indicated in Table 2. In contrast, deletion did not accompany another type of errors.

Table 2: The simultaneity of the errors

	asked	looked	cooped
In	3	0	2
SC	3	8	3
SC + In	8	6	5
D	4	0	0

Furthermore, there are two points to be noticed about the segment change for these consonant clusters. Firstly, the segment change in this type of consonant cluster happened mostly to the second segment. Only one token for “asked,” only two for “cooped” and no tokens for “looked” changed their first segments. Secondly, to what segment the target clusters changed was dependent on the context. While the tokens “asked” and “looked” both have the /kt/ consonant cluster, the former was most frequently changed to the /k/ plus a flap; the latter was most likely to be changed to the /k/ plus an approximant. Taking into account that “looked” and “asked” were followed by the word starting with an

approximant, “wise,” and a fricative, “his,” respectively, one of the possible reasons why even the same consonant clusters changed into the different segment was that the second segment of the consonant cluster was affected by the initial segment of the following word to some degree.

### 3.1.2 The /s/ + stop cluster

Table 3 presents the results of the production of the /s/ plus stop clusters. Surprisingly, this type of the consonant clusters was not difficult for the subjects, despite the fact that many of the pronunciation textbooks encourage the learners to practice these consonant clusters. Japanese learners of English seem to be able to acquire them naturally with no explicit instructions. The five tokens of “snug” were, meanwhile, produced incorrectly and all of them were pronounced as [sɯŋ]. In addition, a vowel was inserted for the two tokens of “snug,” but it is probably because the subjects produced this token long-duration, failing to recognize this word. In spite of this, the statistical test confirmed that the tokens did not significantly affect the correct rate of the /s/ plus stop consonant cluster production,  $F(1.74, 33.14) = 1.48, p > .05$ . It suggests that the subjects were able to produce all of this type of the consonant clusters correctly, including “snug.”

Table 3: The /s/ + stop clusters

	C	D	Pr	In	SC	O
snug	18	0	0	2	0	5
stay	25	0	0	0	0	0
stood	25	0	0	0	0	0
stamping	24	1	0	0	0	0
scouts	25	0	0	0	0	0
asked	25	0	0	0	0	0
joists	24	0	0	0	1	0

### 3.1.3 The fricative + fricative cluster

Similar to the /s/ plus stop clusters, the fricative plus fricative clusters were not demanding for the subjects regarding the tokens “loft” and “rafters,” as shown in Table 4. One token classified into “segment change” changed the second segment to a voiceless affricate, /ts/. The statistical test, moreover, proved that there was no significant difference in the correct rate in producing these two tokens,  $t(24) = -1.00, p > .05$ .

On the other hand, the result of the token “coughed” needs examining carefully. Most of the tokens which were categorized into “others” in this experiment were pronounced as [kɔ:t]. One possible reason is that the subjects were distracted by the orthography of the token because they did not recognize or know this word. However, another reason is that they deleted the first segment of the consonant cluster. Although the former would be

more likely considering that the subjects were all in the elementary class, it is impossible to define precisely which category these errors should be divided into, “deletion” or “others.” In this paper, therefore, the word pronounced as [kɔ:t] was categorized into “deletion” tentatively whereas the word pronounced as [kɔ:f] into “deletion.” At any rate, there is a possibility that the recognition rate of this token would be much lower than that of the other two tokens, so this token “coughed” was excluded from the statistical test.

Table 4: The fricative + fricative clusters

	C	D	Pr	In	SC	O
coughed	8	2	0	0	0	13
loft	24	0	0	0	1	0
rafters	25	0	0	0	0	0

### 3.1.4 The plosive + fricative cluster

The results demonstrated in Table 5 represent that the plosive plus fricative consonant clusters, /gz/ and /ks/, were not easy for the Japanese learners of English. Furthermore, the statistical test showed the two tokens were not significantly affected by the voicing,  $t(21) = -1.56, p > .05$ .

Most of the subjects changed the first segment into either a fricative or an approximant. Regarding “exactly,” /g/ was changed to the voiced fricative for the four tokens and to the approximant for the eleven tokens. Concerning “next,” /k/ was changed to the voiceless fricative for the eleven tokens and to the approximant for the two tokens. In other words, the subjects reduced the stricture to produce this type of the consonant clusters. The reason why /g/ tended to be produced by reducing the stricture up to the degree of the approximant, compared with /k/, would be that voiced fricatives require more energy than voiceless ones because of the vocal fold vibration, which promoted the learners to change the manner of /g/ more radically for producing it.

Table 5: The plosive + fricative clusters

	C	D	Pr	In	Pa	SC	O
exactly	4	0	0	4	0	15	3
next	9	1	0	2	0	13	0

### 3.2 The acoustic features

In Figure 1, the values of F1 and F2 are plotted on the y axis and the x axis respectively. These values are from the inserted vowels for “asked,” “looked,” “cooped,” “exactly” and “next. As a reference, the F1 and F2 values of vowels added to the word end, paragon, are also plotted in the same figure. As far as the visual observation goes, there appears a general tendency that the quality of the vowels inserted between the first and second segment was less variable, compared with that of the vowels

added after the word-final consonant. Therefore, it would imply that the former type of the vowels had fewer factors that affected their segmental features than the latter type of the vowels.

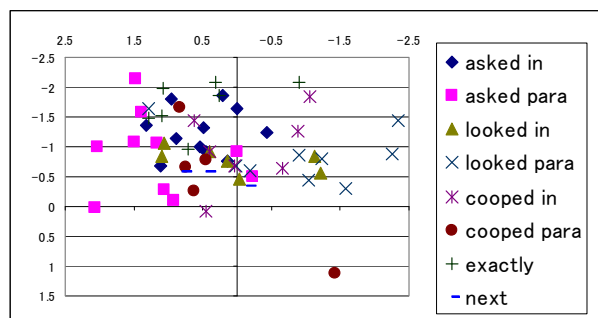


Figure 1: The F1 and F2 of the inserted vowels

Table 6 shows the duration of the inserted vowels for the five tokens above. The simple comparison is not appropriate since the token “exactly” for the plosive plus fricative cluster has eight segments whereas the others, four or five segments. Hence, this paper simply reports that the inserted vowels were approximately 100 ms long.

Table 6: The duration of the inserted vowels

	Plo+Plo	Plo+Fri
Mean	0.136	0.076
S.D.	0.091	0.035

Note: Plo stands for plosive and Fri for fricative.

## 4 Conclusion

To conclude, this study revealed not all types of the consonant clusters were difficult to Japanese learners of English. In the present study, it was obvious that the /s/ plus stop and fricative plus fricative clusters were produced accurately without vowel insertion or critical segment change; the consonant clusters containing the plosive were problematic, though the difficulty slightly varies from token to token. Taking these findings into account, the effective and efficient instruction of the consonant cluster is hoped to be provided.

## References

- Davidson, L. (2006). Phonology, phonetics, or frequency: Influences on the production of non-native sequences. *Journal of Phonetics*, 34, 104-137.
- Igarashi, S. (1981). *English - Its vocal expression*. Tokyo: Nanundo
- Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *Journal of the acoustical society of America*, 49, 606-608.
- Tori, T. & Kaneko, N. (1962). *The pronunciation of English*. Tokyo: Taishukan.

# Using, not knowing traditional Chinese characters affect Kanji recognition by Japanese: A MEG case study

Yoko Okita

International Exchange Center, Tokyo Medical and Dental University  
yokita.isc@tmd.ac.jp

## Abstract

We scanned brain activities while Japanese, Chinese, Taiwanese, and Korean native speakers passively watched the five types of characters using magnetoencephalography (MEG). Participants graduate or undergraduate students living in the Tokyo are. The MEG wave pattern of a female Japanese participant who has learned Chinese and had lived in Taiwan for two years was different from that of other Japanese participants who have not studied Chinese. The MEG wave patterns of the female participant of complex Japanese Kanji and traditional Chinese characters are the almost same, while those of other Japanese participants are different. The female participant with a history of living in Taiwan for two years recognized the traditional Chinese characters and the complex Japanese characters in the same way, while Japanese participants without a history of living Taiwan recognized differently. Using Chinese characters does affect recognition pattern, but not knowing Chinese characters. This result presents a neurological evidence to show the importance of using, not knowing of language, in language processing and learning.

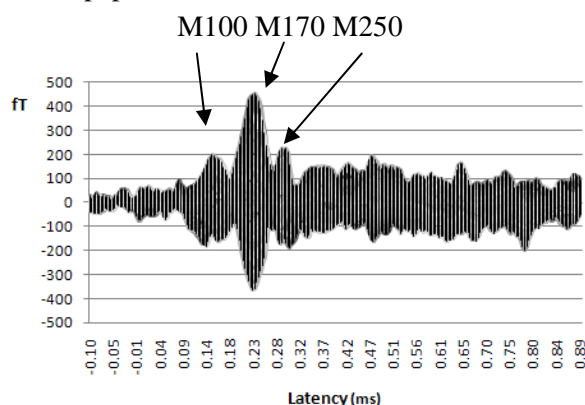
## Keywords

Learning experience, Kanji recognition, MEG

## Introduction

According to previous English word recognition studies using MEG (e.g., Pylkkaen and Marant, 2003), three magnetic fields, M100, M170, and M250 were observed when English words are presented visually, M100 appears after around 100 ms onset. M100, M170, and M260 appear after around 100 ms, 170ms, and 250ms onset respectively. It has been supposed that M100 reflects primary graphic analysis of a character, M170 reflects more complicated graphic analysis and M250 reflects graphic to phonological analysis. Figure 1 shows a representative MEG when Kanji stimuli were passively presented to a male Chinese native speaker. The three major responses, M100, M170, and M250 were observed around the time

window 100-300ms post-stimulus. This example replicated what Pylkkaen and Marant (2003) found in English word recognition. M100 was not found in all participants, whereas M170 was found in all participants. Therefore, the ratios of M170 in magnitude of M170 were compared in this paper.



*Figure 1: A representative MEG from a male Chinese native speaker who saw Kanji stimuli.*

Although, in East Asian regions, Japan, China, Taiwan and Korea, Chinese characters have been the main writing system for long time, Chinese characters have been simplified and used very differently in each region in the past 60 years. In terms of simplification, traditional characters are used in Taiwan and Korea. Japan and China use different simplified characters. For example, 廣 広 广 are the same one Chinese character which means “wide”. 廣 is the traditional character used in Taiwan and Korea. 広 is Kanji character which is the Japanese simplified version, and 广 is the simplified version in China. The way of using Chinese characters are also different in each region. In China and Taiwan only Chinese characters are used. In Japan, besides Chinese characters, Japanese original syllabograms, Hiragana and Katakana are used together. That is three different writing systems are used simultaneously in Japanese. In Korea, Korean characters, Hangeul (한글) is the main writing system. Chinese characters are used very limited

occasions such as in personal names, in a cornerstone, or in the headline of newspapers. A Hangeul character represents one Korean syllable. Korean people study traditional Chinese characters up to high school. Therefore, they know Chinese characters as knowledge, however, they do not use Chinese characters. Especially Korean people do not write Chinese characters so often in daily life. In Japan most of Japanese learn traditional Chinese characters when they study Japanese and Chinese classics in secondary schools. However, traditional Chinese characters are not often used in daily life in Japan. Traditional Chinese characters are occasionally used in personal nouns, such as family names. Japanese people might occasionally see traditional Chinese characters, but they seldom write like Korean. Japanese know traditional Chinese characters as knowledge; however, they do not use them daily.

This paper reports the ratios of M170 magnitude of the female Japanese participant who had studied Chinese and had lived in Taiwan for two years and one Taiwan-Japanese bilingual and discusses how the use of Chinese characters in daily life affects the pattern of kanji recognition.

## 1. Method

### 1.1 Participants

The participants are 16 (9 male, 7 female) Japanese university students who are native speaker of Japanese and one female Taiwan-Japanese bilingual university student. The mean ages of Japanese participants were 24.3(SD=2.20). The Taiwan-Japanese bilingual participant was 23 years old at the time of the experiment. The father and mother of Taiwan-Japanese bilingual participant is Japanese and Taiwanese respectively. Although the Taiwan-Japanese bilingual participant was born, graduated from high school in Taiwan and was studying at a university in Tokyo, she travelled back and forth between Taiwan and Japan. All participants were right handed and had normal or corrected normal vision.

### 1.2 Stimuli

Visual stimuli consist of five groups, Kanji 1 (山, simple Chinese character in Japanese), Kanji 2 (講, complex Chinese character in Japanese), Hangeul characters 마, simplified Chinese 广, and traditional Chinese (廣). The average

number of stroke order of Kanji 1 was 7.93. The average number of stroke order of Kanji2 and traditional Chinese were 11.3 and 10.9 respectively. Some simple Chinese characters are the same in the three systems, simplified Chinese (Chinese), traditional Chinese (Taiwan) and Japanese (Kanji), such as 山 (mountain). Kanji 1 consisted of the characters which are the same in the three systems. Some simplified Chinese and Kanji are the same, whereas the counterpart in traditional characters is different. For example, a character meaning “study” is written as 学 in Chinese and Kanji, whereas the same character is written 學 in Taiwan. The simplified character stimuli were equivalent for the traditional character stimuli characters. Further, there were no equivalent characters for those stimuli in Kanji. Thus the 80 Chinese stimuli are unique in Chinese and the 80 Taiwan stimuli are unique in traditional Chinese characters. It is easily suspected that Chinese character recognition processes are affected by the number of stroke. Therefore, we made the numbers of stroke of stimuli in Taiwan and Kanji 2 almost the same.

### 1.3 Procedure

Each stimulus (black letter with white background) was presented for 100ms followed by blank (white background) for around 1 second. Each stimuli group had 80 items (total 400 items). Four hundred items were divided into 40 blocks. Each block consists of randomly arranged 10 stimuli. The counter traditional character of a simplified Chinese was not included in the same block. Namely, 处 (Chinese) and 處 (Taiwanese) were not in the same block. There was a 2 second resting time between blocks. The participants were instructed to blink during the resting time. Neuromagnetic data were recorded with Bti Magnes 2500, a whole-head MEG system with 148 channels at a sampling rate 254.31 H with high pass filter 1.0Hz, and bandwidth 50Hz. After removing trials in which amplitudes exceeded 4000ft and filtered at 50Hz, the data were averaged by stimulus groups.

### 1.4. Analysis

MEG wave data were transformed to RMS (root mean square) wave forms. And then average The areas for M170 (around 120-200ms) were computed based on RMS waves. Further, the ratios of Kanji2, Hangeul, Chinese and Taiwan

to Kanji 1 in M170 were calculated in each participant.

## 2 Result and Discussion

Figure 2 shows ratios of the five stimuli in the 16 Japanese participants. It appears that individual differences are large in the reaction to the different characters.

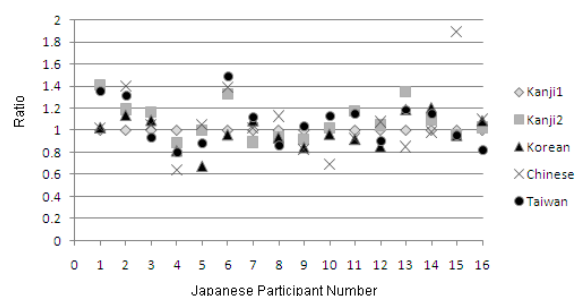


Figure 2: Ratios of the five stimuli in the 16 Japanese participants

An interesting result was found in the Japanese participant 1 who had studied Chinese and lived in Taiwan. The Japanese participant 1 seemed to classify the five groups into two, one group consists of Kanji 2 and traditional Chinese, and the other group consists of Kanji 1, simplified Chinese and Korean characters. Other Japanese participants without the experience of learning Chinese and living in Taiwan did not classify the group into two as clear as the Japanese participant 1.

M170 of Kanji 2 and traditional Chinese characters

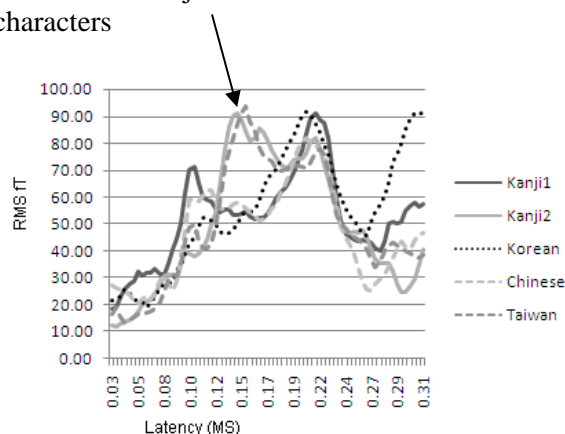


Figure 3: RMS wave forms of the five stimuli of the Japanese participant 1 from 30-310 ms onset.

Figure 3 shows the RMS waves of the Japanese participant 1. The RMS wave pattern also showed that Kanji 2 and traditional Chinese

were processed differently from Kanji 1, simplified Chinese, and Korean characters. Kanji 2 and traditional Chinese showed clear M170 components, whereas other groups did not. Kanji 2 and traditional Chinese characters required more energy to process Chinese character forms, whereas the other groups required little energy. It is also seemed that Korean characters were slightly differently processed from Kanji 1 and simplified Chinese. The numbers of stroke of simplified Chinese and Kanji 1 were not the same. Therefore, the number of stroke is less likely the key for determining the difference in kanji recognition process. It seemed that the Japanese participant 1 processed Kanji 2 and traditional Chinese characters in a very similar way. This result suggests that learning and using experience of traditional Chinese characters might affect kanji recognition process.

Figure 4 shows the ratios of the five stimuli of female Japanese participant 1 and the female Taiwan-Chinese bilingual participant. Unlike the Japanese participant 1, the Taiwan-Japanese bilingual participant processed only Kanji 2 differently from the other four groups.

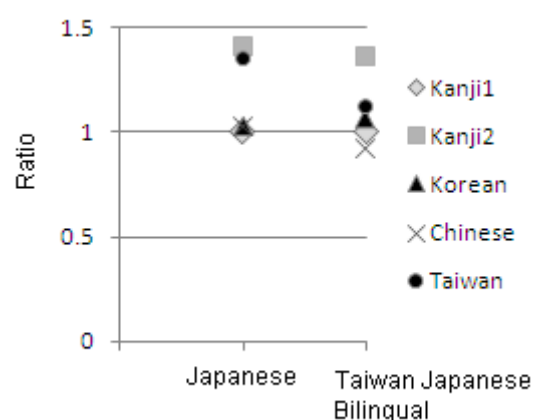


Figure 4: Ratios of the five stimuli of Japanese participant 1 and the Taiwan-Japanese bilingual participant.

Figure 5 shows RMS wave forms of the five stimuli of the Taiwan-Japanese bilingual participant. The magnitude of M170 of Kanji 2 is quite larger than the others. It is clearly showed that only Kanji 2 was processed differently and required much more energy to analyze character forms compared with the other four groups. Although Kanji 2 and traditional Chinese character are almost same in the numbers of stroke, only Kanji 2 was required much more energy to process the character

forms. It appears that for the Taiwan-Japanese bilingual participant, only Kanji 2 were foreign characters. Probably for the Taiwan-Japanese participant, traditional Chinese characters are the first “character” and Japanese kanji is the second “character”. Kanji 1 might be a part of the first “character”. Simplified Chinese characters might be foreign to the participant. However, the structure of simplified Chinese characters is quite simple and the numbers of stroke are also smaller than those of traditional Chinese characters and Kanji 2. Therefore, the magnitude of only Kanji 2 was predominant.

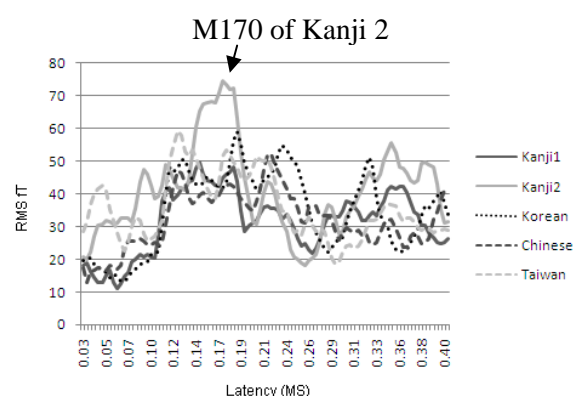


Figure 5: RMS wave forms of the five stimuli of the Taiwan-Japanese bilingual participant from 30 -400 ms onset

This case suggests that Kanji 2 is foreign to Taiwanese, even if the numbers of stroke Kanji 2 and traditional Chinese characters are almost the same.

### 3 Conclusion

A Japanese participant, who had experience of learning and living in Taiwan, processed complex Kanji and traditional Chinese character in the same way. Other Japanese participants who have not studied Chinese and lived in Taiwan processed Kanji 2 and traditional Chinese characters in different ways. Most of Japanese had learned traditional Chinese characters during secondary school days. Therefore, Japanese university students know traditional Chinese characters as a knowledge, even if use of traditional Chinese characters are not common. A Taiwan-Japanese bilingual participant seemed to process Kanji 2 as foreign characters. It appears that using, not knowing, Chinese characters do affect kanji recognition process.

### References

- Pylkkänen, L., & Marantz, A. (2003). Tracking the time course of word recognition with MEG, *Trends in Cognitive Sciences*, 7(5), 187-189

# **Investigating the Effects of English E-Tutoring in a Junior High Afterschool Program**

Hung-Hsuan Kao

Jinwen University of Science & Technology

alex@alexkao.net

## **Abstract**

With the advent of technology, applying E-learning platforms in instruction has become a trend. This study investigated the effects of applying an E-learning platform for English tutoring in an afterschool program in a junior high school in Taiwan. The sample of the study came from a junior high school in a rural area in northern Taiwan. A survey method was used to gather the data for analysis. It analyzed the participants' attitudes towards using a new technology system to learn English and their learning results. The results of this study add to the literature of applying computer for assisting language learning.

## **Keywords**

Computer-assisted language learning (CALL), E-learning platform, Joinnet

## **1. Introduction**

The 21st century is termed the Information Age. As it's coming, the information technology and Internet have brought great influence on our life and education. Applying E-learning platforms in instruction has been a trend. Recently it has also been used in one-to-one tutoring in afterschool programs, especially for rural

schools. In 2006, the Ministry of Education (MOE), Taiwan initiated projects aimed at providing afterschool online tutoring programs for elementary and junior high schools in rural areas. As students in rural areas have fewer resources for learning and the public transportation for them to go downtown is inconvenient. Thus, school authorities and students view the afterschool program as a great opportunity for students to improve their academic work. During the past three years, a great number of students have benefited from the online synchronous one-to-one tutoring in afterschool programs. More and more schools are interested in joining the project to help their students to gain more education opportunities.

This study reports one of the projects in the MOE online tutoring programs conducted by the researcher in northern Taiwan. It investigated the effects of one-to-one English tutoring in an afterschool program in a junior high school in the researcher's county by applying an E-learning platform, JoinNet.

## **2. Theoretical framework**

Synchronous communication has been increasingly incorporated into second

language instruction to expand learners' exposure to the target language through real-time interaction. When learners engage in interaction, they receive input, feedback, and opportunities to produce modified output. All of the above facilitate the development of learners' new language (Long & Robinson, 1998; Swain & Lapkin, 1995). Providing more opportunities for learners to practice using the new language helps them achieve better learning results.

Recently, with the advent of technology, E-learning platforms have been created to provide not only synchronous audio interaction but also synchronous face-to-face interaction. For example, JoinNet is the platform used widely for online courses. The face-to-face function offers opportunities for learners to have real interactions to practice their English skills in more authentic and communicative setting. It allows for real-time oral communication among students and between students and tutors (Hampel & Hauck, 2004).

JoinNet has been used in Taiwan for online tutoring afterschool programs in elementary schools and junior high schools in rural areas. What are the online tutoring skills? According to Shepherd (1999), there are many potential roles for the tutor. Three of the most important are as a subject expert, a coach, and an assessor. As a subject expert, the tutor is to supplement self-study materials, which can fill in any gaps, clarify any misunderstandings, and point learners to sources of information. As a coach, the tutor is responsible for helping the learner to achieve their learning goals by challenging,

encouraging and providing constructive feedback. As an assessor, the tutor checks to ensure that learners have achieved the learning objectives. The tutors in this project hope to play the role as recommended by Shepherd (1999).

### 3. The project

Twenty tutees participated in this afterschool tutoring program. They took this chance to learn English. Some of them volunteered to join the program, while some of them were recommended by their teachers. This means that not all of the students fell behind or were low achievers. Students attended the class for two hours once a week.

Twenty tutors were recruited from the researcher's department. They had taken the researcher's computer-assisted language learning in the previous semester. Four workshops on how to use the equipment and the platform, how to provide assistance for English learners, how to deliver English lessons with supplementary materials, and guidelines for dealing with frustrating teaching experiences were arranged. Besides, parties for getting to know about each other were arranged, one at the tutors' school and one at the tutees' school.

As this is a one-to-one synchronous tutoring program. One tutor and one tutee are in a party. The equipment for the person in each side includes a computer, a headphone, a microphone, a webcam, and a sketch pad. Both the school systems need to provide the Internet connection. Figure 1 shows the diagram of the equipment needed

in this project.

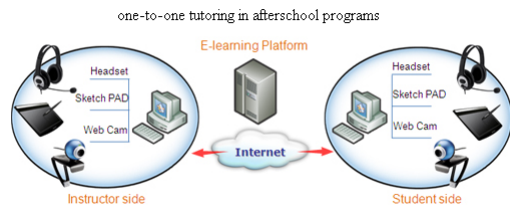


Figure 1: The Equipment for the One-to-One Tutoring Afterschool Program

In this project, an E-learning platform, JoinNet, is used as the platform for tutors and tutees to communicate. Figure 2 shows the interface of JoinNet. The Ministry of Education provides JoinNet platform for both schools running and joining the project. JoinNet is the HomeMeeting's client software, connecting to the HomeMeeting Server, for video matting, white-board presentation, synchronized web browsing, desktop and application sharing, remote PC control, and recoding file playback.

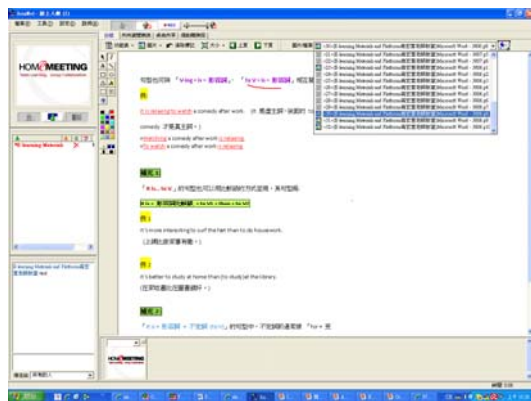


Figure 2: The Interface of JoinNet

The program lasted for 15 weeks in a semester. At the end of the program, a questionnaire was administered to the tutee to gain their perspectives of the effects of the online tutoring. In addition, the results

of the tutee's English midterm and final exams were collected for analysis.

## 4. The results

### 4.1 Student demographics

The students' demographics are presented in Table 1. Eight male students and twelve female students participated in this project. They are eighth graders, the second year in the junior high school. One student is thirteen years old, thirteen are fourteen years old, and six are fifteen years old.

Table 1: Students' Background Information

Gender	Numer	%
Male	8	40
Female	12	60
Age		
13 Years old	1	5
14 Years old	13	65
15 Years old	6	30

### 4.2 Students' learning results

How have students performed in their academic work? Data were collected from students' two school-held midterm academic tests in March and May. Figure 3 shows the comparison of the two test result scores (Spring 2009-1 and Spring 2009-2).

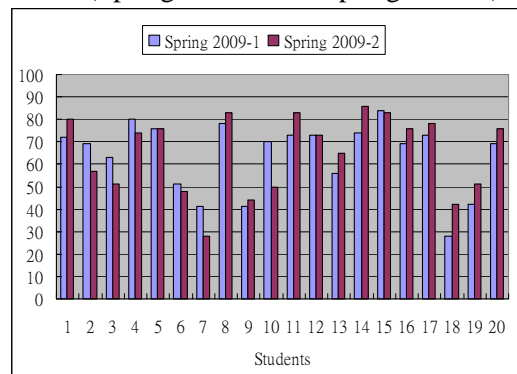


Figure 3: Comparing Students' Test Scores

Comparing the two test results, the researcher finds that students have made progresses. Among the twenty students, thirteen students scored higher in the second time than the first time. Overall, it shows that students have benefited from this tutoring program.

### 4.3 Students' attitude analysis

The results of the data from the questionnaire are shown in Table 2. See Appendix A for the items in the questionnaire.

Table2: Result of Students' Attitudes towards E-learning Platforms.

Item	<i>N</i>	<i>M</i>	<i>SD</i>
1	20	4.60	.50262
2	20	3.90	.96791
3	20	4.15	.67082
4	20	4.35	.58714
5	20	3.55	1.09904
6	20	4.35	.58714
7	20	4.15	.74516
8	20	4.45	.68633
9	20	4.35	.74516
10	20	4.35	.67082
11	20	4.40	.68056
12	20	4.35	.58714
13	20	4.55	.68633
14	20	4.40	.59824
15	20	4.60	.68056

The questionnaire was designed by the researcher. It consisted of 15 items

measuring the students' attitudes towards using the platform for English learning, using a 5-point Likert scale (5: strongly agree, 4: agree, 3: neutral, 2: disagree, 1: strongly disagree). Responses to the fifteen items yielded high internal consistency (Cronbach's  $\alpha=0.893$ ). The means of thirteen items are higher than 4. One item, Item 2, has a mean of 3.9, which is close to the statement of agree. One item (Item 5) has the mean of 3.55. In sum, students show positive attitudes towards the tutoring program.

Analyzing the items, the study finds that first students feel excited and happy attending this online after-school tutoring program (see Item 8, 9 and 10). Second, the online after-school tutoring program is helpful for students to study English and they think their English has been improved (see Item 2, 3, and 12). Third, as to the functions of the tutors, one to one tutoring can help students have more interactions with the teacher, make them feel safe to ask questions, get the teacher's response instantly, and help them understand the parts that she/he didn't understand. (see Item 1, 11, 13, and 14).

Fourth, about the ease of using the platform, students agree that the platform is easy to understand and use (see Item 6 and 7). Lastly, students express that they hope to participate in the school's next online after-school tutoring program again (see Item 15). As literature indicates, when learners engage in interaction, they receive input, feedback, and opportunities to produce modified output (Long & Robinson,

1998; Swain & Lapkin, 1995). This study finds that this online tutoring program facilitates the development of learners new language learning.

Items 5 did not have a mean higher than 4 to show participants agreed with this statement. This item asks, "Using JoinNet in the tutoring program, I can hear the teacher and see the images clearly." The reasons why this item did not achieve a higher mean is due to the technology problem—the computers at the students' school are not very new and efficient, and the bandwidth of Internet connection is not fast enough. The above problems have been reported and discussed by their school teachers and the researcher. The junior high school needs to update their computer system. Thus, after this summer, the Government of Taipei County will update the computers for this school. It hopes to provide tutors and tutees with better equipment for participating in this program.

## 5. Conclusion

This study finds that students have improved their English in academic performance and they show positive attitudes toward using E-learning platforms for their tutoring. This online tutoring program provides students in rural schools have an opportunity to have a scenario which is similar to a face-to-face learning context. It also gives students the opportunity to have interaction with able learners, the tutor.

This study has demonstrated an approach to implement technology in education to those EFL learners who need

the most assistance in learning. It adds to the literature of applying computer-assisted language learning. It hopes to offer pedagogical implications for researchers and educators who are interested in researching this topic.

## 6. References and appendices

### 6.1 References

- Hampel, R., & Hauck, M. (2004). Towards an effective of audio conferencing in distance language courses, *Language Learning & Technology*, 8 (1), 66-82. Retrieved on June 20, 2009, from <http://llt.msu.edu/vol8num1/hampel/default.html>
- Lee, L. (2008). Focus-on-form through collaborative scaffolding in expert-to-novice online interaction. *Language Learning & Technology*, 12(3), 53-72.
- Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in second language acquisition* (pp. 15-41). Cambridge: Cambridge University Press.
- Shepherd, C. (1999). *Online tutoring skills*. Tactix. Retrieved on June 20, 2009, from <http://www.fastrak-consulting.co.uk/tactix/features/tutoring/tutoring.htm>
- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16(3), 371-391.

## 6.2 Appendices

### Appendix A. The Questionnaire

Item	Questions
1	I think the instruction mode of one-to-one can help me have more interactions with the teacher.
2	I think the online after-school tutoring program is helpful for me to study English.
3	After participating in the online after-school tutoring program, I think my English has been improved.
4	I think the instruction mode of one-to-one in the online after-school tutoring program is helpful for my English learning.
5	6. Using JoinNet in the tutoring program, I can hear the teacher and see the images clearly.
6	I think the JoinNet platform is easy to use.
7	I think the interface of the JoinNet platform is easy to understand.
8	When I participate in the online after-school tutoring program, I feel very happy.
9	When I participate in the online after-school tutoring program, I feel very excited.
10	When I participate in the online after-school tutoring program, I feel I have a lot of expectations.
11	When I use the whiteboard on the platform, I can get the teacher's response instantly.
12	I think the online after-school tutoring program is helpful for me.

13	In the one-to-one online after-school tutoring program, I feel safe to ask questions.
14	My teacher in the tutoring program can help me better understand the parts that I don't understand.
15	I hope I can participate in the school's online after-school tutoring program again.

# The Acquisition of Restrictive Relative Clauses by Japanese and Korean Learners of English

Hiromasa Ohba<sup>1</sup>, Kenichi Yamakawa<sup>2</sup>, Naoki Sugino<sup>3</sup>, Yuko Shimizu<sup>4</sup>, Michiko Nakano<sup>5</sup>

<sup>1</sup>Joetsu University of Education, <sup>2</sup>Yasuda Women's University, <sup>3,4</sup>Ritsumeikan University,

<sup>5</sup>Waseda University

hohba@jien.ac.jp, kyamakaw@yasuda-u.ac.jp, gwisno@is.ritsumeai.ac.jp, yuko  
s@ec.ritsumeai.ac.jp, nakanom@waseda.jp

## Abstract

The present study examines the acquisition of English restrictive relative clauses by Japanese and Korean learners of English. Previous studies have shown that Japanese learners of English can acquire *wh*-movement in relative clause formation, which is not involved in relative clause formation in Japanese, and be sensitive to a constraint in *wh*-movement in English. It is also said that there is no *wh*-movement involved in restrictive relative clauses in Korean. A grammatical judgment task was administered to university-level Japanese and Korean learners of English in order to explore their knowledge of restrictive relative clauses. They were asked to read the test sentences and rate their grammaticality on the 5-point scale indicated. Then, 34 Japanese and 34 Korean learners (intermediate and high-intermediate proficiency groups of English) were selected on the basis of the scores of the test aiming to measure their proficiency of the English grammar. Their task performances were compared in terms of (1) whether there is any difference in the grammaticality judgment on the surface morphological properties of restrictive relative clause constructions between Japanese and Korean learners of English, and (2) whether there is any difference in sensitivity to the Subjacency violations in restrictive relative clause constructions between Japanese and Korean learners of English.

## Keywords

restrictive relative clauses, *wh*-movement, surface morphological properties, Subjacency

## Introduction

One of the research topics in the field of second language (L2) acquisition is to explain how adult L2 learners develop their knowledge of syntax (Hawkins, 2001). The present study focuses on the acquisition of English restrictive relative clauses,

which are considered one of the most difficult grammatical constructions for Japanese learners of English as a foreign language (EFL) to learn. One of the studies which investigated the acquisition of restrictive relative clauses in English showed that Japanese can acquire *wh*-movement in restrictive relative clause formation, which is not involved in relative clause constructions in Japanese, and be sensitive to a constraint in *wh*-movement in English (Ohba et al., 2006). It is also said that there is no *wh*-movement involved in relative clause formation in Korean.

Therefore, this study examines the acquisition of restrictive relative clauses by Japanese and Korean EFL learners in terms of similarities and differences in their task performances on these grammatical constructions in English.

## 1. Theoretical background

In English, relative clause constructions are formed by a relative operator being extracted from the relative clause domain and being moved to the position after the relative head which is being modified (i.e., *man*), as in (1).

(1) The man [*who*(*m*)<sub>i</sub> [I saw *t*<sub>i</sub> yesterday]] is John.

Within the Minimalist Program (Chomsky, 1995; 1998), overt movement is only allowed when it is motivated by the presence of a strong formal feature. English has the feature [+R] in Complementizer (C), which drives relative-operator movement, as in (1) (Takeda, 1999).

Since in the case of operator-oriented relative clause as in English relative clause constructions, the relation between the relative head and its associated relative clause is established by binding of the relative pronoun by the relative head, they are subject to the Subjacency condition, a constraint on *wh*-movement, as in (2). This constraint specifies that *wh*-movement cannot cross more than

one bounding node, where bounding nodes are inflectional phrase (IP) and noun phrase (NP).

- (2) \*a gentlemen [who<sub>i</sub> the suit [that t<sub>i</sub> is wearing] is dirty]

On the other hand, in Japanese, such a relative operator movement is not included in relative clause constructions due to the lack of relative operators and the relative head modified is put after the relative clause, showing no Subjacency condition. It is also stated that there is no *wh*-movement involved in relative clauses in Korean, indicating no Subjacency condition.

## 2. The study

The aim of the present study is to investigate the acquisition of restrictive relative clauses by Japanese and Korean EFL learners. The research questions addressed here are as follows:

- (1) Is there any difference in the grammaticality judgment on the surface morphological properties of restrictive relative clause constructions between Japanese and Korean EFL learners?
- (2) Is there any difference in sensitivity to the Subjacency violations in restrictive relative clause constructions between Japanese and Korean EFL learners?

### 2.1 Participants

Experimental participants were adult native speakers of Japanese and Korean (university students). On the basis of the scores of the test called MEG (Measure of English Test) (Shimizu et al., 2003, 2006), which we developed to measure learners' proficiency of the English grammar, 34 Japanese and 34 Korean learners were selected and classified into 2 proficiency groups of English, respectively (J1 and K1 are intermediate and J2 and K2 are high-intermediate groups). Details of the number of participants and the scores on the MEG in each group are summarized in Table 1.

Table 1: Details of Participants

	<i>N</i>	<i>M</i>	<i>SD</i>
J1	18	16.83	3.47
J2	16	25.50	3.50
K1	15	17.13	3.70
K2	19	25.37	2.27

### 2.2 Materials and Procedure

In order to investigate the acquisition of restrictive relative clause constructions, especially focusing on whether Japanese and Korean EFL learners can acquire *wh*-movement or not, a written

grammaticality judgment task was designed. This test includes 32 target and 13 filler sentences. The sentences in the task fell into the following 3 groups, as in (1) to (3) (with 10 categories) (see Appendix).

- (1) The sentences involving the following 3 types of grammatical relative clauses
  - (a) Category 1: Relative clauses with a *wh*-operator (7 items)
  - (b) Category 2: Relative clauses with a complementizer *that* (4 items)
  - (c) Category 3: Relative clauses with a null operator and a null complementizer (3 items)
- (2) The sentences involving the following 2 types of ungrammatical relative clauses
  - (a) Category 4: Relative clauses with a doubly-filled complementizer (*who(m) that/which that*) (4 items)
  - (b) Category 5: Relative clauses with a resumptive pronoun (4 items)
- (3) The sentences violating Subjacency conditions in the following 5 relative clause construction types
  - (a) Category 6: Relative clauses with an extraction from a relative clause (2 items)
  - (b) Category 7: Relative clauses with an extraction from a sentential subject (two items)
  - (c) Category 8: Relative clauses with an extraction from an adjunct island (2 items)
  - (d) Category 9: Relative clauses with an extraction from an embedded question (*wh*-island) (2 items)
  - (e) Category 10: Relative clauses with an extraction from a complex NP (2 items)

Test items were randomized, and the participants were asked to read the test sentences and rate their grammaticality on a 5-point scale indicated. They were told that 5 meant that the sentence was 'completely grammatical', 1 that it was 'completely ungrammatical', and 2, 3, and 4 were gradations used if they thought the sentence was more or less grammatical. Detailed instructions were given on the use of the scale before testing, and there were initial practice items for information before the test began.

In scoring the responses from the participants, we calculated the 'distance' of a response from the correct answer. In other words, if there were any discrepancy between the response and the correct answer, that discrepancy was reduced from the full mark of 4. For example, if a participant judged a grammatical sentence as "5", s/he would gain 4 points. If an ungrammatical sentence was judged "5", that response was converted to 0 points. We believe this way of scoring would be more

informative than the binary scoring.

### 3. Results and discussion

The results of the grammatical relative clauses with *wh*-operator (Category 1), complementizer *that* (Category 2), null operator or complementizer (Category 3), the ungrammatical relative clauses with *who(m) that* or *which that* (Category 4) and resumptive pronoun (Category 5) are presented in Tables 2 and Figure 1, which compares the mean scores for the four English proficiency groups. In both grammatical and ungrammatical relative clauses, participants' mean scores should approach 4 (maximum score) if they judge correctly, and their mean scores should approach 0 (minimum score) if they judge incorrectly.

The results show that there are no significant differences in mean scores between Japanese (J1) and Korean (K1) intermediate groups and between Japanese (J2) and Korean (K2) high-intermediate groups in all the grammatical and ungrammatical relative clauses.

Table 2: Grammaticality Judgment Scores on Relative Clause Constructions

		Cat. 1	Cat. 2	Cat. 3	Cat. 4	Cat. 5
J1	M	2.23	2.37	2.19	2.30	2.09
	SD	0.68	0.78	0.83	0.65	0.78
J2	M	2.55	2.75	2.31	2.45	2.23
	SD	0.61	0.75	0.73	0.59	0.97
K1	M	2.44	2.36	2.35	2.34	2.31
	SD	0.69	0.64	0.61	0.70	0.86
K2	M	2.67	2.93	2.52	2.38	2.68
	SD	0.53	0.46	0.58	0.83	0.83

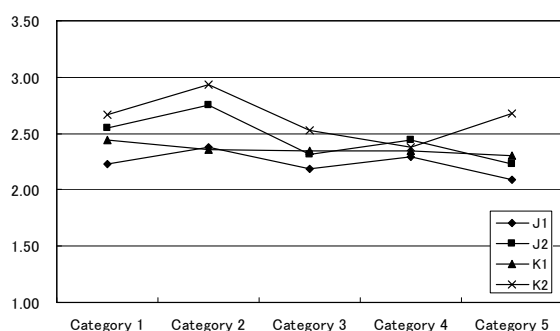


Figure 1: Grammaticality Judgment Scores on Relative Clause Constructions

Table 3 and Figure 2 give the results of Subjacency violations, which are *wh*-movements out of relative clauses (Category 6), sentential subjects (Category 7), adjunct islands (Category 8), embedded questions (Category 9) and complex NPs (Category 10), rated by the four English proficiency groups.

The results indicate that Korean (K2) high-intermediate proficiency group performed better than the other three proficiency groups.

Interestingly, there is no statistically significant difference between Japanese high-intermediate group (J2) and Korean intermediate group (K1). IN the case of Subjacency violation types, there no statistically significant difference among them.

Table 3: Grammaticality Judgment Scores on Subjacency Violations

		Cat. 6	Cat. 7	Cat. 8	Cat. 9	Cat. 10
J1	M	1.88	1.82	1.84	2.16	2.02
	SD	0.63	0.54	0.64	0.70	0.75
J2	M	2.12	2.34	2.30	2.61	2.37
	SD	0.86	0.70	0.62	0.64	0.71
K1	M	2.35	2.22	2.27	2.47	2.32
	SD	0.76	0.89	0.83	0.83	0.76
K2	M	2.86	2.75	2.88	2.97	2.69
	SD	0.71	0.69	0.54	0.49	0.55

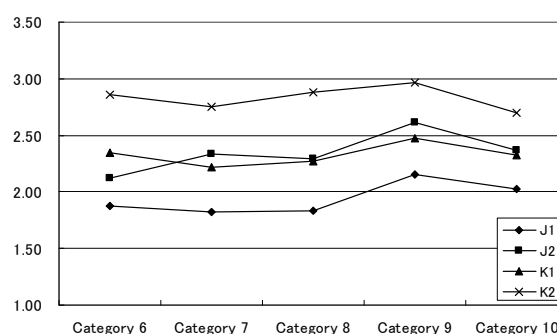


Figure 2: Grammaticality Judgment Scores on Subjacency Violations

As Korean EFL learners approach the higher level of English proficiency, they become more sensitive to Subjacency violation than Japanese EFL learners with equivalent English proficiency, which means that Korean EFL learners can easily acquire *wh*-movement in restrictive relative clause constructions.

### 4. Conclusion

The findings of this study are that there was a proficiency-related increase in possible correct judgement on Subjacency violations, but there is a difference in their rating Subjacency violations between adult Japanese and Korean EFL learners with nearly equal proficiency of English (both intermediate and high-intermediate groups). It calls for further consideration.

### References

- Chomsky, N. (1995). *The minimalist program*. Cambridge, Mass.: MIT Press.
- Chomsky, N. (1998). *Minimalist inquiries: The framework*. MIT occasional papers in linguistics. No. 15.
- Hawkins, R. (2001). *Second language syntax: A generative introduction*. Malden, MA: Blackwell.

- Ohba, H., Sugino, N., Nakano, M., Yamakawa, K., Shimizu, Y., & Kimura, S. (2006). The Development of grammatical competence of Japanese EFL learners: Focusing on relative clause constructions. *Journal of Pan-Pacific Association of Applied Linguistics*. 10, 73–87.
- Shimizu, Y., Kimura, S., Sugino, N., Yamakawa, K., Ohba, H., & Nakano, M. (2003). Eibunpou nouryoku tesuto no datousei, shinraisei no kenshou to shin eigo bunpou nouryoku Test Measure of English Grammar (MEG) [The validity and reliability of standardized tests and compilation of “Measure of English Grammar (MEG)”]. *Policy Science*. 10, 59–68.
- Shimizu, Y., Yamakawa, K., Sugino, N., Ohba, H., & Nakano, M. (2006). Developing a sample-free grammatical proficiency test for SLA research. *Proceedings of the 10th Conference of Pan-Pacific Association of Applied Linguistics*. 227-235.
- Takeda, K. (1999). *Multiple headed structures*. Unpublished doctoral dissertation. University of California, Irvine.
- 24.\*This is the soup which Mary visited a restaurant which served.
- (7) *Wh*-movement out of a sentential subject (subject island)
- 25.\*This is the meeting which that Bill attended shocked his parents.
- 26.\*This is the ghost which a picture of frightened the children.
- (8) *Wh*-movement out of an adjunct island
- 27.\*This is the homework which Ann went to school before she did.
- 28.\*This is the girl who(m) the bell rang while I was thinking of.
- (9) *Wh*-movement out of an embedded question (*wh*-island)
- 29.\*This is the house which Peter knows when Tom bought.
- 30.\*This is the story which Rebecca wondered who believed.
- (10) *Wh*-movement out of a complex NP
- 31.This is the car which we heard the news that Dick bought.
- 32.\*This is the boy who(m) Jack described the way that Bill hit.

#### Appendix. Sentences used in the Grammaticality Judgement Task

- (1) Grammatical relative clauses involving a *wh*-operator
1. The woman who helped me with my homework is Patty.
  2. The man who(m) I employed as my assistant works hard.
  3. The girl for whom I have bought a computer is my sister.
  4. The boy to whom I talked yesterday seemed very nervous.
  5. The job which I wanted to apply for was very popular.
  6. The girl whose handbag was stolen is suffering from shock.
  7. The man used a word whose meaning I don't know at all.
- (2) Grammatical relative clauses involving *that*
8. The student that has written this letter must be crazy.
  9. The car that you can see over there caused this accident.
  10. The student that you gave a present to looked very happy.
  11. The picture that he is looking at was painted by Picasso.
- (3) Grammatical relative clauses involving a null operator and a null complementiser
12. The house you can see on the corner was built ten years ago.
  13. The friend they lent money to bought a very big house.
  14. The magazine they are always talking about is very useful.
- (4) Ungrammatical relative clauses involving *who/whom/which that*
- 15.\*The woman who that is singing on the stage is my aunt.
  - 16.\*The glasses which that Judy broke were very expensive.
  - 17.\*The dogs which that I gave the milk to were very small.
  - 18.\*The woman whom that we talked with was our teacher.
- (5) Ungrammatical relative clauses involving resumptive pronouns
- 19.\*The building that it stands near the lake is our school.
  - 20.\*The classmate that you don't like him is very unkind.
  - 21.\*The friend that I lent the book to her studied very hard.
  - 22.\*The town that my mother came from it is far from here.
- (6) *Wh*-movement out of a relative clause
- 23.\*This is the book which John interviewed the man who criticized.

# A Study on Learner Factors in Acquiring English Pronunciation

Yuko Tominaga

Graduate Student, University of Tokyo

yukotominana@yahoo.co.jp

## Abstract

The objective of this research is to detect effective factors in English pronunciation teaching from the viewpoint of learner factors. This research is based on the data of the Successful Pronunciation Learners (SPL) and Poor Pronunciation Learners (PPL) selected by the raters, Japanese and native speakers of English (NSE), in Tominaga (2005). The participants' learning history, learning strategy, motivation, and so on were analyzed by the questionnaire and interview. The holistic procedure was almost based on that of Tanabe and Koyama (1998), and the questions in the questionnaire were selected based on Baba (2003). In this paper, in order to examine whether the results of Tominaga (2005) were a coincidence or not, statistic analysis was conducted focusing on learners' factors that may be contributive to their pronunciation learning. The results show that significant factors were not found between SPL and PPL of the junior high school, on the other hand, between SPL and PPL of the senior high school, several significant factors were found. This indicates that the senior high school SPL show a remarkable contrast to other learners in such cognitive and meta-cognitive areas of learning English as attitude, motivation, and persistence.

## Keywords

Learner Factors, Learner's Autonomy, SPL, PPL

## 1 Introduction

In EFL classrooms, many learners are struggling with learning English, and also many teachers are struggling with teaching English. However, it is true that there still exist Successful Foreign Language Learners (SFLL) who are excellent in the four skills of English: reading, listening, speaking, and writing in the standardized English qualification tests. How, then, have they acquired such skills and maintained the skills? In order to detect some clues to and make suggestions for better teaching English to learners in EFL settings, this paper, in the perspective of pronunciation, attempts to analyze the

questionnaires and the interviews for SFLL and PFL (Poor Foreign Language Learners), focusing on their study history, attitudes, and individual strategies that may have affected their learning. In this study, by tracing the learning experiences of junior and senior high school students, it is expected that significant awareness to teachers — when and how they have their eyes opened to the autonomous learning — would be suggested for better teaching.

## 2 Method

### 2.1 General description

This study statistically analyzed the contributive factors of pronunciation teaching/learning that Tominaga (2005) suggested because such factors may have stemmed from her intuition cultivated by her teaching experience. In Tominaga (2005), she examined to what extent SPL and PPL exist among the junior and senior high school students participated in her study, and analyzed their learning history, learning strategy, and motivation by the questionnaire and interview. The holistic procedure was almost based on that of Tanabe and Koyama (1998), and the questions in the questionnaire were selected based on Baba (2003), i.e., they were chosen under the conditions that: 1) the questions meet the situation of the classroom setting, and 2) by referring to the "self-consciousness scale" by Sugawara (1984) the questions can be factor-analyzed from the viewpoint of pronunciation learning and psychology, especially self-consciousness. Therefore, in this paper, in order to examine whether the results of Tominaga (2005) were a coincidence or not, statistic analysis was conducted focusing on learners' factors that may be contributive to their pronunciation learning.

### 2.2 Participants

The participants were junior and senior high schools students selected as SPL or PPL in the research by Tominaga (2005). From among the 232 junior high school students (first-year 101, second-year 60, and third-year 71), 24 SPL (10.3%)

and 39 PPL (17%) were selected. From among the 339 first-year senior high school students, 8 SPL (2.4%) and 25 PPL (7%) were selected. The selection was based both on the students' academic aptitude and their performance in class activities, which were evaluated by JTE (Japanese Teachers of English), and on the three-step evaluations of their reading aloud by JTE and ALT (Assistant Language Teachers). The students who had lived in the country for more than one year where English is spoken and the students whose parent (either mother or father) was a native speaker of English were excluded from the selection in advance.

### 2.3 Procedure

First, a survey by questionnaire was conducted on the participants, both SPL and PPL. The questionnaire consists of 22 items (9 Yes-No questions and 13 multiple choice questions) regarding their learning history, learning strategy, and motivation. Next, in order to obtain a more detailed factors for growing up to be SLL, individual interviews (20-30 minutes) about their answers to the questionnaire were conducted on the SPL. The interviewees talked about their own experiences independently following the topics the author gave<sup>1)</sup>, and each talk was tape-recorded. Then, in order to examine which factor contributes to the difference between the SPL and PPL in the result of the questionnaire, two tests were conducted: chi-square test on the Yes-No questions and t-test on the multiple-choice questions. Finally, regarding the items in the questionnaire that the two tests statistically indicated significant correlations, the results of the interviews of the SPL were examined in order to gain a detailed understanding of how these factors contributed to the ability of the SPL.

### 3 Results

Table 1 indicates the results of the chi-square tests. Three items are identified as significant factors for the difference between SPL and PPL among junior and high school students respectively. The significant factors between SPL and PPL of the junior high school are 1) received informal instruction, 2) be fond of English learning, and 3) have a role model. On the other hand, between SPL and PPL of senior high school, the significant factors are 1) be fond of English learning, 2) do independent pronunciation practice, and 3) have a role model. Two factors ("be fond of English learning"; "have a role model"), are common between them. Additionally, one notable result is that no senior high school SPL have received informal instruction before entering junior high school, whereas junior high school SPL received it.

Table 1 The results of the chi-square tests on Yes-No questions: %

Questions /Yes	Junior		Senior	
	SPL	PPL	SPL	PPL
Have lived abroad	0	8	0	0
Have a chance to speak English daily	17	5	0	0
Have friends talking with you in English	17	5	0	16
Receive informal instruction	100	33	*	36
Be fond of English learning	83	33	*	34
Taught pronunciation skills enough	83	49	33	24
Learned a learning method outside school	17	15	33	12
Have an independent practice	67	10	100	16
Have a role model	83	34	*	44

\* p < .05 \*\* p < .01 \*\*\* p = .00

When the numerical value of the chi-tests was under 0.5, it means a significant factor. It is shown by asterisked mark.

Table 2 indicates the results of the t-tests. No significant item was identified among the junior high school students. On the other hand, in the comparison between the senior high school SPL and PPL, six items are identified as to be relevant to their motivations and their teachers' intervention: 1) have bad impression of JTE's pronunciation, 2) JTE often corrected pronunciation, 3) JTE taught stress, intonation, and rhythm, 4) strongly want to be an SPL, 5) be fond of reading aloud and repetition, and 6) imitate native speakers.

Table 2 The results of the t-tests on five- multiple choice questions: mean (SD)

Questions	Junior		Senior	
	SPL	PPL	SPL	PPL
Frequency of using English by JTE	3.0(0.8)	2.5(0.7)	3.3(0.4)	3.0(0.5)
Bad impression of JTE's pronunciation	1.7(0.9)	0.9(1.0)	1.0(0.7)	2.1(0.9)
JTE taught detailed pronunciation	1.2(0.4)	1.2(0.6)	1.3(0.4)	0.7(0.6)
JTE often corrected pronunciation	0.3(0.5)	0.8(0.7)	1.5(0.5)	0.4(0.5)
JTE taught stress, intonation & rhythm	1.3(0.9)	1.4(1.2)	2.0(0.0)	1.1(0.9)
JTE taught the goals of pronunciation	1.2(0.9)	1.3(0.9)	1.5(0.5)	1.0(0.6)
ALT taught detailed pronunciation	1.2(0.9)	1.1(0.7)	0.8(0.8)	0.7(0.6)
ALT often corrected pronunciation	1.0(0.8)	1.1(0.7)	0.8(0.8)	0.9(0.7)
ALT taught stress, intonation & rhythm	1.0(1.2)	1.2(1.1)	1.0(1.2)	1.5(1.0)
Strongly want to be an SPL	2.5(1.1)	1.9(1.1)	3.5(0.4)	2.3(1.0)
Be fond of reading aloud & repetition	2.3(0.9)	1.6(1.1)	3.8(0.5)	1.9(0.8)
Imitate native speakers	1.7(0.9)	1.4(0.9)	2.5(0.5)	1.2(0.9)
Have confidence in pronunciation	0.7(0.5)	0.5(0.5)	0.5(0.5)	0.2(0.4)

\* p < .05 \*\* p < .01 \*\*\* p = .00

When the numerical value of the t-tests was under 0.5, it means a significant factor. It is shown by asterisked mark.

Generally, the above-mentioned items that can be considered to have statistically significant difference between the SPL and the PPL are related to strategies, motivations, and teachers' intervention.

From the interviews, three common factors were observed that might affect the junior high school SPL. On the other hand, seven common factors were identified among senior high school SPL.

## 4 Conclusion

The results of this study clearly indicate that the senior high school SPL show a remarkable contrast to other learners in such cognitive and meta-cognitive areas of learning English as attitude, motivation, and persistence. To be more concrete, the SPL excelled others in 'successfully' making use of elements such as choice, interest, relevance, expectancy, and outcomes. Thus self-engaged and self-invested, the SPL appear to have heightened their level of interest, involvement, and responsibility in their learning, which results in the achievement that otherwise could not have easily been produced in an EFL setting. In order to produce more SFL in formal education, it would be important for teachers to set a class atmosphere in which learners are encouraged and supported to learn English autonomously. The atmosphere should be created by teachers when the learners start learning English, in the case of Japan when they enter junior high school, and should be continuously renewed.

External stimuli such as films and music can be considered as a strong tool for having the students become more interested in English learning, since the SPL state in the interviews that such stimuli encouraged them to learn English positively. Moreover, the SPL make it clear that the encounter with respectable teachers is another very positive stimulus. In fact, it must never be forgotten that teachers can be an external stimulus to learners. In addition, guidance to lead learners in finding out their original strategies should be conducted in class. For example, appropriate assignments by and feedback from teachers should be continuously offered to learners.

It would be ideal that these various supports should be given to the learners in junior high schools regardless of their school evaluations. The evaluations should not be regarded as their real ability of English. It might be said that junior high school time is "in-put" period, and senior high school time is "in-take" period. As the learners grow up, they select and accept the appropriate learning way, and finally they would open their eyes to their original learning way to be SFL. Learners tend to build their foundation to be autonomous learners at the level of senior high school.

In conclusion, it is considered that what controls the learning effectiveness is meta-cognitive learning, and what supports the meta-cognitive learning is the learners' autonomy. Yoshijima and Hasegawa (2007) mentioned about the gap of the methodology and purpose of English between elementary education and secondary education. He states that elementary English education is

"affective", and secondary English education is "cognitive." Considering Japanese school system that includes the introduction of foreign language activity (practically, English) to elementary school in 2011, it might be said that the focus of teaching English in elementary schools is "affective", in junior high school is "cognitive", and in senior high school is "meta-cognitive." Moreover, it might be important that the learners soft-land between these three aspects. However, as many researchers discuss it, teaching English in elementary schools still has a number of problems. At any rate, in the long run, without sticking to the instant results and evaluations, the integrated study from the viewpoint of whole formal instruction in Japan would be necessary in order to generate more SFL.

## 5 References and Appendix

### 5.1 References

- Baba, T. (Ed). (2003). *Eigo Supikingu Ron* [Theoretical study of English speaking]. Tokyo: Kagensha.
- Celce-Murcia, M., Brinton, D., & Goodwin, J. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge: Cambridge University Press.
- Celce-Murcia, M. (Ed.). (2001). *Teaching English as a second or foreign language* (3rd ed.). MA: Heinle and Heinle Publishers.
- Cohen, A.D.(1998). *Strategies in learning and using a second language*. London: Longman.
- Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University
- Gallego, J. C. (1990). The intelligibility of three nonnative English-speaking teaching assistants: An analysis of student-reported communication breakdowns. *Applied Linguistics*, 1(2), 219-237.
- Gardner, R. & Lambert, W. (1972). *Attitude and motivation in second language learning*. Rowley, MA: Newbury House.
- Goodwin, J. (2001). Teaching pronunciation. In M. Celce-Murcia, (Ed.), *Teaching English as a second or foreign language* (3rd ed.), (pp. 117-137). MA: Heinle and Heinle Publishers.
- Hatori, H & Matsuhata, K. (1989). *Gakushusha chushin no Eigo kyouiku* [ELT: A learner-centered approach]. Tokyo: Taishukan Shoten.
- John, W. (1983). *Issues In Language Testing Research*. Rowley, MA: Newbury House.
- Maeda, T. (2002). Language learning strategies used by Japanese high school learners and achievement. *Language Education & Technology*, 39,137-148

- Naiman, N., Frohlich, M., Stern, H., & Toedesco, A. (1978). *The good language learner*. Toronto: Ontario Institute for Studies in Education.
- Oller, J., & Richards, J. (Eds.). (1973). *Focus on the learner*. Newbury House Publishers: Rowley, MA.
- Oxford, R. (1990). *Language learning strategies: What every teacher should know*. Cambridge: Newbury House.
- Rubin, J. (1975). What the 'good language learner' can teach us? *TESOL Quarterly*, 9: 1, 41-51.
- Rubin, J. & Thompson, I. (1994). *How to be a more successful language learner*. New York: Heinle & Heinle.
- Shirahata, T. (2007). Gengo Shutoku kara Mita Sho Chu Renkei [Coordination between ES and JHS from a viewpoint of language acquisition]. In R. Matsukawa & K. Kinoshita (Eds.), *Shogakko Eigo to Chugakko Eigo wo Musubu [Linking ES and JHS English programs]* (pp. 64-76). Tokyo: Koryosha.
- Shirai, Y. (2004). *Gaikokugo Gakusyu ni Seikosuru-hito, Shinai-hito. [Successful Foreign Language Learners and Poor Foreign Language Learners]*. Tokyo: Iwanami Science Library.
- Stern, H. (1975). What can we learn from the good language learner? *Canadian Modern Language Review*, 31, 304-318.
- Sugawara, K. (1984). Ji-ishiki shakudo. [Self-consciousness scale]. *The Study of Psychology*, 55, 184-188.
- Takada, T. (2003b). Learner characteristics of early starters and late starters of English language learning: Anxiety, motivation, and aptitude. *JALT Journal*, 25, 5-30.
- Takeuchi, O. (2003). *Yoriyoi Gaikokugo Gakushuho Wo Motomete. [Research on better language learning strategies]*. Tokyo: Shohakusha
- Takeuchi, O. & Wakamoto, N. (2001). *Language learning strategies used by Japanese college learners of English: A synthesis of hour empirical studies. Language Education & Techonology*, 38, 21-43.
- Tanabe, Y., & Koyama, K. (1998). An experimental study on teaching English pronunciation: from input to intake. *English Phonetics*, 2, 356-376.
- Tanabe, Y., Koyama, K., & Nakano, K. (1998a). 'A Constructive Examination of English Pronunciation Work: The Case of Junior High School Textbooks.' a paper to be presented at *LLA National Conference* in August, 1998.
- Tanabe, Y., Koyama, K., & Nakano, K. (1998b). 'A Constructive Examination of English Pronunciation Work: The Case of Senior High School Textbooks.' a paper to be presented at *Zenkoku Eigo Kyoiku Gakkai* in August, 1998.
- Tanabe, Y. (2003). English pronunciation instruction in Japan: Current state of affairs. *Senshu Daigaku Gaikokugo Kyouiku Ronshu [A Journal of the Senshu University Research Society]*, 31, 77-101.
- Tominaga, Y. (2005). *An Analysis Successful Pronunciation Learners: What They Can Tell Us about Pronunciation Instruction*. MA thesis submitted to Waseda Univerisity.
- Yoshijima, S. & Hasegawa, H. (Eds). (2007). *Gaikokugo Kyoiku IV : Sho-gakko kara Chu-gakko e. [Foreign Language Teaching □: Elementary School to Secondary School]*. Tokyo: Asahi-shuppansha

## 5.2 Appendix

### Appendix : A Part of the Questionnaire in English

#### Questionnaire on learning English

(This question has nothing to do with your grade.)

grade( ) class( ) ID No.( ) Name( )  
(Notice) As for the questions in **bold type**, choose either one and circle it.

- No. 1 Have you ever lived abroad? **Yes • No**  
\* If your answer is, "Yes"  
Where have you lived?  
How long did you live there?
- No. 2 Do you ever have a chance to speak in English in your daily life? **Yes • No**  
\* If your answer is, "Yes,"  
How often do you speak English per week?  
About( ) per week  
How long do you speak English at one time?  
About( ) minutes
- No. 3 Have you had any friends talking with you in English? **Yes • No**  
\* If your answer is, "Yes,"  
Where does he/she come from?  
(If you have more than one friends, please write down about all of them.)  
How long have you known her/him?
- No. 4 Have you had any chance to speak or practice English before you entered the junior high school? **Yes • No**  
\* If your answer is "yes", where was it?  
(e.g.: cram school, English conversation school, church and etc.)
- No. 5 Have you been fond of learning English through your school life? **Yes • No**
- No. 6 Do you think your previous English teachers have taught you pronunciation and listening skills enough?  
**I think so. • I do not think so.**  
\* If your answer is, "I think so"  
Can you give some details ?  
(procedures, techniques and etc.)  
\* If your answer is, "I don't think so",  
Can you give some details ? )

Continued to No. 15.

# Examination of Foreign Language Anxiety Constructs between Japanese and Korean College Students Learning English

Nami Iwaki<sup>1</sup> and Hyun Jin Kim<sup>2</sup>

<sup>1</sup>Nagoya University, <sup>2</sup>Cheongju National University of Education

iwaki@ecis.nagoya-u.ac.jp, 37hjkim@hanmail.net

## Abstract

The purpose of this study is to investigate and compare the anxiety constructs of Japanese colleges students and Korean college students learning English. The first part of this study will examine anxiety construct of Japanese and Korean students using FLCAS (Horwitz, Horwitz, and Cope, 1986). It will examine Japanese and Korean students' anxiety levels and how they are constructed. The second part of this study will compare the anxiety levels of Japanese and Korean learners and see if there are differences where and how they feel anxiety.

## Keywords

Foreign language anxiety, anxiety construct, classroom English

## Introduction

Numbers of studies show that for many students, language courses are the most anxiety-provoking class that they take (Horwitz, Horwitz, and Cope, 1986; MacIntyre and Gardner, 1999), and that anxiety is common among foreign language learners (Aida, 1999). This phenomenon especially applies to Asian students, whom in many studies show that they tend to feel much anxiety when learning a foreign language. Iwaki's (2007) study compared and examined FLCAS results for Asian and Western learners who are learning English and found that the average of FLCAS was over 100, whereas Western students' average were all below 100. Many studies suggest that something unique about Asian students may play a role in explaining the differences in English confidence levels between Asian students and that of other nationalities. Although studies have shown that Asian students seem anxious towards learning a foreign language, only few studies have compared the difference within Asian students. This study will examine the foreign language anxiety difference between Japanese and Korean students.

## 1 Research Questions

Research Question 1: What is the anxiety construct of Japanese students? What is the anxiety construct of Korean students?

Research Question 2: Are there anxiety differences between Japanese and Korean students? If so, which group feels more anxiety, in what area?

## 2 Methods

### 2.1 Participants

The participants were 110 Japanese and 102 Korean college students learning English. The average age for Japanese students was 19.1, consisting of first year students and Korean students consisting from third and fourth year, with the average of 22.2. Japanese students' majors were Economics, Business and Foreign language (mix of German, French and Spanish) and English for Korean students.

### 2.2 Instrumentation

The FLCAS by Horwitz, Horwitz and Cope (1986) was used in this study. The FLCAS is a self-report measure of language learners' feeling of anxiety in the foreign language classroom, consisting from 33 statements. Participants answered on a six-point Likert scale, ranging from (a) Strongly agree, (b) mostly agree, (c) somewhat agree to (f) strongly disagree. For each statement, the highest degree of anxiety was six points, and the lowest was one point, anxiety scores ranging from 33 to 198. The word "foreign language" in the original scale was replaced with "English", for example, "I keep thinking that the other students are better at foreign language than I am" was "I keep thinking that the other students are better at English than I am".

### 2.3 Procedure

The data collection was done in April 2009. A brief description of the present study was introduced, along with informed consent form. The entire survey was carried out during English class time.

### 3 Results

#### 3.1 Reliability

Cronbach's alpha for the FLCAS for both Japanese and Korean students were  $\alpha = .91$ .

#### 3.2 FLCAS construct of Japanese and Korean students

Factor analyses for Japanese student extracted a three-factor solution as shown in Table 1, and Korean students' FLCAS extracted a four-factor solution as shown in Table 2. In interpreting the results, a factor loading of .40 was employed as a cutoff for inclusion of a variable. Consequently, 6 items for Japanese and 5 items for Korean students did not load on any factor.

##### Japanese students

For Japanese students, the total variance accounted for was 40.1%. Factor 1 consisted of 12 items, and was labeled *Self-consciousness towards English*, exemplified by "I am afraid that the other students will laugh at me when I speak English", and "I feel very self-conscious about speaking English in front of other students". These and other statements are related to being self-conscious in the English class such as being laughed at if they say something wrong or strange in the class. Factor 2 consisted of 7 items, accounting for 12.2% of variance. Items included "I start to panic when I have to speak without preparation in English class", and "I always feel that the other students speak English better than I do". This factor dealt with students being overly worried about speaking in the class. Therefore, Factor 2 was labeled *Speaking in the classroom anxiety*. The third factor accounted for 11.2% of variance with 8 items. These items consisted of positive attitudes toward English class, such as "I am usually at ease during tests in my English class", and "I don't feel pressure to prepare very well for English class". Therefore, Factor 3 was named *Comfortableness towards English class*. The Alpha coefficients were .88, .82, and .76, respectively for three factors.

Table 1: Factor analysis for Japanese students

Factors and items	Factor			$h^2$
	F1	F2	F3	
Factor 1 Self-consciousness towards...				
F19	.68			.62
F27	.59			.66
F31	.58			.64
F21	.56			.49
F25	.55			.56
F30	.54			.60
F10	.51			.54
F26	.49			.70
F33	.47			.59

F24	.46		.63
F20	.45		.68
F29	.45		.60
Factor 2 Speaking in the classroom ...			
F9	.73		.65
F4	.70		.56
F16	.56		.72
F23	.51		.65
F11	.48		.59
F12	.46		.61
F15	.44		.51
Factor 3 Comfortableness towards...			
F18		.69	.64
F14		.66	.60
F2		.61	.62
F8		.46	.51
F32		.44	.52
F28		.43	.61
F13		.42	.61
F22		.40	.48
% of variance			
% of variance accounted for by the solution	14.2	12.2	11.2
			40.1
Items not included in factor solution			
F5	-.11	-.16	.29
F1	.17	.35	.25
F6	-.06	-.08	.04
F3	.39	.36	.33
F7	.38	.34	.16
F17	.36	-.02	.00

##### Korean students

For Korean students, the total variance accounted for was 43.0%. Factor 1 consisted of 12 items, which included "I never feel quite sure of myself when I am speaking in my English class" and "I start to panic when I have to speak without preparation in English class". Therefore, factor 1 was named *Speaking anxiety*. Factor 2 consisted of 7 items, accounting for 10.5 variance. Items included "It embarrasses me to volunteer answers in my English class" and "I often feel like not going to my English class". Factor 2 was labeled *English classroom anxiety*. The third factor accounted for 7.9% of variance with items such as "I get nervous when I don't understand what the teacher is saying in English" and "I get upset when I don't understand what the English teacher is correcting". Therefore, it was labeled *English teacher anxiety*. The fourth factor accounted for 6.4% of variance with four items. Items included were "I am usually at ease during tests in my English class" (negatively worded) and "I worry about the consequences of failing my English class" and "The more study for an English test, the more confused I get". This factor dealt with students worrying about the test, so it was labeled *Test anxiety*. The Alpha coefficients were .90 .78 .76 and .62, respectively for four factors.

Table 2: Factor analysis for Korean students

Factors and items	Factor				$h^2$
	F1	F2	F3	F4	
Factor 1 Speaking anxiety					
F1	.78				.74
F9	.73				.74
F3	.65				.65
F18	.64				.59
F33	.64				.68
F27	.63				.71
F2	.63				.63
F14	.60				.62
F20	.53				.68
F26	.51				.67
F11	.51				.52
F28	.48				.56
Factor 2 English classroom ...					
F13		.66			.66
F31		.62			.78
F6		.55			.59
F19		.52			.45
F17		.49			.62
F30		.48			.73
F23		.43			.68
Factor 3 English teacher anxiety					
F29			.72		.67
F15			.65		.61
F4			.54		.67
F25			.47		.57
F22			.44		.57
Factor 4 Test anxiety					
F8				.66	.61
F7				.50	.77
F10				.49	.43
F21				.48	.49
% of variance					
% of variance accounted for by the solution	18.1	10.5	7.9	6.4	43.0
Items not included in factor solution					
F5	.12	.16	.12	.00	
F12	.37	.31	.16	.13	
F32	.10	.04	.07	.10	
F16	.36	.29	.37	.20	
F24	.51	.46	-.14	.11	

### 3.3 FLCAS results between Japanese and Korean students

In order to examine the second research question, an independent t-test was conducted to compare the FLCAS scores for Japanese and Korean students (Table 3). First, the total score of FLCAS and the average of 33 items were compared. The results show that there is a significant difference between the two groups. Japanese students expressed much higher anxiety than Korean students. Next, in order to examine FLCAS items individually, independent t-tests were conducted to all 33 items (Table 4). The results found a significant difference for 21 items, mostly Japanese students showing higher anxiety than Korean students except for item 25.

Table 3: FLCAS comparison between Japanese and Korean students (total score)

	Japan ( $n=110$ )		Korea ( $n=102$ )		$t$ (201)
	$M$	$SD$	$M$	$SD$	
FLCAS average	3.91	.65	3.56	.61	3.93***
FLCAS total	129.04	21.41	117.64	20.18	3.93***

\*\*\* $p < .001$

Table 4: FLCAS comparison between Japanese and Korean students (individual items)

	Japan ( $n=139$ )		Korea ( $n=102$ )		$t$ (239)
	$M$	$SD$	$M$	$SD$	
FLCAS 1	5.22	.94	3.94	1.14	8.84***
FLCAS 3	4.55	1.34	3.77	1.33	4.21***
FLCAS 4	4.45	1.25	3.32	1.26	6.49***
FLCAS 5	2.92	1.60	2.44	1.23	2.43**
FLCAS 6	3.99	1.29	2.49	1.01	9.17***
FLCAS 10	3.96	1.43	3.53	1.18	2.41*
FLCAS 12	4.22	1.29	3.30	1.15	5.46***
FLCAS 13	4.35	1.22	3.30	1.18	6.32***
FLCAS 14	4.41	1.25	3.86	1.29	3.11**
FLCAS 15	3.99	1.20	3.25	1.14	4.49***
FLCAS 17	3.61	1.42	2.49	1.41	5.76***
FLCAS 18	4.52	1.07	3.99	.99	3.27***
FLCAS 19	3.31	1.22	2.47	1.15	5.14***
FLCAS 20	4.35	3.67	3.67	1.25	3.97***
FLCAS 22	3.42	1.37	4.40	1.22	5.52***
FLCAS 24	4.35	1.27	3.71	1.14	3.91***
FLCAS 25	3.21	1.34	3.77	1.09	3.36**
FLCAS 26	3.72	1.49	3.08	1.31	3.31**
FLCAS 28	3.35	1.33	3.80	.96	2.87**
FLCAS 31	3.35	1.27	2.91	1.21	2.54*
FLCAS 32	3.99	3.13	3.33	1.30	3.89***

\*\*\* $p < .001$ , \*\* $p < .01$ , \* $p < .05$

## 4 Discussion and future suggestion

In the past, many studies labeled Asian students as one group in foreign language anxiety research stating that they show higher anxiety than the Western students (Truitt, 1995). However, this study showed that there is a significant difference between Japanese and Korean learners and that it may be unsafe to unify students as one group labeled "Asian".

Research Question 1: First research question found that Japanese and Korean students show different constructs of English anxiety. This Factor analyses indicated that different set of groups show different anxiety constructs. The results suggest that there may be a difference in constructs between Japanese and Korean students. The series of factor analyses showed that factors vary according to the group investigated. Nonetheless, speaking in the foreign language, which is the leading anxiety factor pointed out by numerous researchers, (Burden, 2004; Kim, 2002), was also the principle anxiety factor for students learning English for this study.

**Research Question 2:** The results examined FLCAS items individually, and results indicated that Japanese students showed much higher anxiety than the Korean students. This result suggests that within Asian students, there are differences to the level in what students feel anxiety. However, the following must be considered before further interpretation of the results.

**Major:** Where as Japanese students' major were Business, Economics and other foreign language, Korean students' major was English. This perhaps is the best explanation thought for the anxiety difference between the two groups. The Japanese students in this group studies English as subsidiary subject, where as Korean students take English as their major subject.

**Grade:** Japanese students participated were all first year students just out of high school. On the other hand, Korean students participated were third and fourth year. These differences also probably affected the anxiety score.

**Gender:** The male-female ratio of Japanese students was 55:55, and 19:83 for Korean students. Some studies have indicated that female students are better at languages.

**English conversation class:** Only 22.7% of Japanese students have the experience of attending English conversation school whereas 98% of Korean students have. This reason also may play a role in the anxiety difference.

For the future research, it is important that the subjects of the two countries have the similar background, such as same major, same grade, balanced size of gender and so on. However, despite these different factors affecting students, the anxiety difference is still quite large between the two groups. Does Junior and High school education system in both country affect the level of anxiety? Are there any differences in teaching styles between Japanese and Korean teachers? It is necessary to look at such factors suggested as well in order to understand the differences found in this study.

## 5 References

- Aida, Y. (1990). Examination of Horwitz, Horwitz, and Cope's construct of foreign language anxiety: The case of students of Japanese. *The Modern Language Journal*, 78, 155-167.
- Burden, P. (2004). The teacher as facilitator: Reducing anxiety in the EFL university classroom. *JALT Hokkaido Journal*, 8, 3-18.
- Horwitz, E., Horwitz, M., & Cope, J. A. (1986). Foreign language classroom anxiety. *The Modern Language Journal* 70, 125-132.
- Iwaki, I. (2007). A study of foreign language anxiety and its constructs of Japanese college students learning English. *Proceedings of the 12th Conference of Pan-Pacific Association of Applied Linguistics*, 284-287.
- Kim, Y. (2002). Construction of a theoretical mode of foreign language anxiety and development of a measure of the construct: Korean university EFL learners' case. *Unpublished doctoral dissertation, Indiana University at Bloomington*.
- Truitt, S. (1995). Anxiety and beliefs about language learning: A study of Korean university students learning English. *Unpublished doctoral dissertation, The University of Texas at Austin*.

## Appendix A. Foreign Language Classroom Anxiety Scale (FLCAS)

1. I never feel quite sure of myself when I am speaking in my English class.
2. I don't worry about making mistakes in English class.
3. I tremble when I know that I'm going to be called on in English class.
4. It frightens me when I don't understand what the teacher is saying in English.
5. It wouldn't bother me at all to take more foreign language classes.
6. During English class, I find myself thinking about things that have nothing to do with the course.
7. I keep thinking that the other students are better at English than I am.
8. I am usually at ease during tests in my English class.
9. I start to panic when I have to speak without preparation in English class.
10. I worry about the consequences of failing my English class.
11. I don't understand why some people get so upset over English classes.
12. In English class, I can get so nervous I forget things I know.
13. It embarrasses me to volunteer answers in my English class.
14. I would not be nervous speaking the English with native speakers.
15. I get upset when I don't understand what the English teacher is correcting.
16. Even if I am well prepared for English class, I feel anxious about it.
17. I often feel like not going to my English class.
18. I feel confident when I speak in the English class.
19. I am afraid that my English teacher is ready to correct every mistake I made.
20. I can feel my heart pounding when I'm going to be called on in English class.
21. The more I study for an English test, the more confused I get.
22. I don't feel pressure to prepare very well for English class.
23. I always feel that the other students speak English better than I do.
24. I feel very self-conscious about speaking English in front of other students.
25. English moves so quickly I worry about getting left behind.
26. I feel more tense and nervous in my English class than in my other classes.
27. I get nervous and confused when I am speaking in my English class.
28. When I'm on my way to English class, I feel very sure and relaxed.
29. I get nervous when I don't understand every word the English teacher says.
30. I feel overwhelmed that the other students will laugh at me when I speak English.
31. I am afraid that the other students will laugh at me when I speak English.
32. I would probably feel comfortable around native speakers of English.
33. I get nervous when the English teacher asks questions that I haven't prepared in advance.

# Supporting and Assessing L2 Learners' Self-Regulated Learning

Eiichiro Tsutsui<sup>1</sup>, Kazuharu Owada<sup>2</sup>, Norifumi Ueda<sup>3</sup> and Michiko Nakano<sup>4</sup>

<sup>1</sup>International Center, Hiroshima International University,

<sup>2</sup>Tokyo College of Music

<sup>3</sup>Faculty of Foreign Language Studies, Mejiro University

<sup>4</sup>School of Education, Waseda University

tsutsui@ic.hirokoku-u.ac.jp, qwq04229@nifty.com,

uedan@mejiro.ac.jp, nakanom@waseda.ac.jp

## Abstract

As information and communication technologies evolve over the recent years, L2 learners have more opportunities to study English in the e-learning environment. Many schools and universities have started to use and apply the CMS (Course Manage System) to the EFL (English as a Foreign Language) context. Networking technologies have enabled learners to use synchronous texting and oral chatting for their cross-cultural learning. Against this background, learners need to study or learn English more independently and interdependently. However, because learners (especially, basic-level learners) tend to depend too much on teachers' guidance, we need to support not only their self-regulated learning outside the classroom but also their effective use of newly-introduced learning methods. In doing so, Tsutsui et al. (2007, 2008) established an English Course Navigation System. This system is designed to introduce a wide variety of language learning courses as well as learning methods and learner strategies. According to their preferences, strategy use, anxiety, and motivational directions, the learners are given individualized feedback and are encouraged to take some language courses that might suit their needs. In addition, we can adopt some ways of encouraging learners in class to increase their self-regulatory activities.

## Keywords

Self-Regulated Learning, Self-assessment, Self-reflection, Information and Communication Technologies

## 1 Introduction

The purpose of this study is two-fold: to describe an

on-line feedback system called English Course Navigation System; and to find some ways of supporting self-regulated learning for learners in the e-learning environment.

Since many universities are making extensive use of information and communication technologies for their educational purposes, L2 learners have more opportunities to study English in the e-learning environment. One of the inevitable consequences for this approach is that L2 learners must study more or less independently. Many researchers have so far claimed that the use of e-learning can lead to learners' independent learning, and they focused on the importance of self-paced learning and student-centered approaches. However, some learners who tend to rely so much on teachers' guidance not ready for such environmental changes. Therefore, the motivation of this study arises from the need to raise awareness of learners' independence in order to encourage and support learners' self-regulated learning in class.

With a view to raising the awareness of independent learning, Tsutsui et al. (2007, 2008) created an English Course Navigation System. Firstly, this system was designed to introduce a wide variety of learning methods and learner strategies, because the wise use of various strategies is indispensable for successful language learning (See also Oxford, 1989; Vandergrift, 2003). Secondly, the system focuses on learners' anxiety and procrastinations. It shows us how much the learners are influenced by these negative affects as well as how well they are prepared to cope with their anxiety and procrastinations. These learners' meta-cognitive analyses are also important for their future accomplishments (See also O'Malley and Chamot, 1990). Finally, the feedback about one's

motivational directions and intensity is given in this system. The importance of self-regulation and self-determination cannot be emphasized enough because learners should be motivated to learn and be responsible for their own learning process.

To sum up, in terms of their individual characteristics such as strategy use, anxiety, and motivational directions, the diagnostic feedbacks are available in this system. Thus, the system users are given individualized feedbacks and are encouraged to take some language courses that might suit their needs.

## 2 Strategic Control Loop

The important aspect of self-regulated learning is that learners need their behavioral control as well as affective and cognitive controls (Zimmerman and Shunk, 2008). Zimmerman (1989) categorized the mechanism of self-regulation into three; covert, behavioral, and environmental regulations. Therefore, learners need to try hard to change their mind, behavior and environment, by using a wide variety of strategies and obtaining positive or negative feedback of the strategies they have used. That is, self-regulated learners can examine whether or not the strategy use is effective and successful in order to achieve learning goals. Then, they adjust the strategies or try other strategies. This ever-going control loop can be self-regulation. In the process of self-regulation, learners may use 'bad' and 'ineffective' strategies. However, the most important point here is that they should carefully monitor their strategy use and adjust their strategies accordingly in order to cope with various deficiencies that they experience as well as their own self-efficacy. The 'loop' of strategy use, self-monitoring and strategy use is named "strategic control loop" by Zimmerman (1990). He explains that the use of a given strategy should accompany self-regulatory activities, such as self-observations and self-evaluations. This may be the reason why simply knowing and using a lot of strategies does not necessarily lead to self-regulated learning. Learners should be more analytic about their strategy use, and language teachers need to incorporate the mechanism in practical teaching.

One drawback of our English Course Navigation System is that the system provides one-shot feedback rather than long-term feedback, even though "the loop" is necessary for self-regulated learning. Therefore, we created a self-checklist from the need to find a way to facilitate the "feedback loop" in the practice of teaching.

## 3 Self-checklist

### 3.1 Data Analysis 1

In order to facilitate learners' self-regulation, we

created a self-checklist (See Appendix A for more details).

In the e-learning environment, we asked basic-level learners to answer this self-checklist on the scale of four, ranging from '1: strongly disagree' to '4: strongly agree.' The survey was conducted at the end of each class. The weekly class lasts 90 minutes. We used the self-checklist only during regular classes; however, we did not use this self-checklist during the weeks before and after the exam because the classes held in these weeks are different from those in the other weeks. Therefore, the data for the nine weeks' classes are subject to our analyses.

We obtained and factor-analyzed 232 responses. The results are shown in Table 1. Against our initial speculations, two factors were extracted. Factor 1 included three items; self-control, and highlighting, note-taking. Factor 2 consisted of four items; preparation, speaking-aloud, inferencing, dictionary use and self-reflection. Highlighting and note-taking in English classes may be closely related to affective regulations because some learners have reported that these behaviors are useful to fight off their sleepiness. However, after obtaining more respondents, we conducted a factor analysis again and found that only one factor was extracted.

Table 1: Results of factor analysis

	Components		h <sup>2</sup>
	1	2	
1.Preparation		.554	.411
2.Speaking		.870	.627
4.Self-Control	.450		.449
5.Highlighting	.843		.594
6.Note-Taking	.869		.750
7.Inferencing		.352	.400
8.Dictionary use (comprehension-checking)		.491	.218
9.Self-Reflection		.575	.429

In order to find some relations between learners' test performance and self-regulatory behaviors in class, we examined whether or not learners' test scores are correlated with perceived self-regulatory assessment scores.

82 respondents reported the scores of 4 different standardized tests, (1) vocabulary, (2) grammar, (3) reading, (4) listening tests. All tests were paper-based, and conducted at the end of the term.

It should be noted that none of the items are negatively correlated with proficiency tests, although high correlation coefficients are not yielded. Grammar and reading tests are more highly correlated with self-regulatory behaviors in class

than vocabulary and listening tests.

Table 2: Pearson's correlation coefficients  
(Checklist items vs. standardized tests)

		Item 1	Item 2	Item 4	Item 5
V	r	.192	.076	.035	.191
G	r	<b>.383**</b>	<b>.270*</b>	<b>.283*</b>	<b>.368**</b>
R	r	<b>.265*</b>	.132	<b>.229*</b>	<b>.230*</b>
L	r	.212	.165	.21	.18

		Item 6	Item 7	Item 8	Item 9
V	r	.181	<b>.249*</b>	.109	.140
G	r	<b>.376**</b>	<b>.335**</b>	.108	<b>.222*</b>
R	r	<b>.322**</b>	<b>.293**</b>	.146	.143
L	r	<b>.220*</b>	<b>.232*</b>	.027	.094

\*\* $p < .01$ , \* $p < .05$ ,  $N = 82$ .

In our data, self-regulated behaviors might lead to higher scores in grammar and reading questions. Across all tests, item 7 (inferencing the phrases) may be the most effective self-regulatory control. The second most effective behavior is note-taking (item 6). Item 8 (dictionary use) did not show high correlations with any proficiency tests. Just because they use a dictionary many times does not mean they are self-regulated learners. The point is that self-regulated learners can make effective use of the dictionary. Therefore, the wording of item 8 can be a little problematic. In version 2 which changed from version 1, item 8' asks whether or not they use a dictionary as a means of comprehension checking.

### 3.2 Data analysis 2

The research questions of data analysis 2 are two-fold:

RQ1: Is there any difference between two different proficiency groups in terms of self-regulatory behaviors?

RQ2: Is there an upward trend in perceived self-regulation scores as the sessions go by?

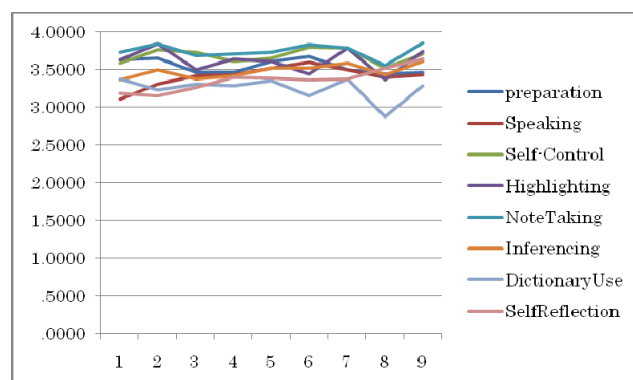


Figure 1: Mean score movement: Group A ( $N = 28$ )

Table 3: Mean score movement: Group A ( $N = 28$ )

Session	Item1	2	4	5	6	7	8	9
1	3.63	3.11	3.59	3.63	3.74	3.37	3.37	3.19
2	3.65	3.31	3.77	3.85	3.85	3.50	3.23	3.15
3	3.46	3.42	3.73	3.50	3.69	3.38	3.31	3.27
4	3.46	3.43	3.61	3.64	3.71	3.43	3.29	3.39
5	3.61	3.52	3.65	3.61	3.74	3.52	3.35	3.39
6	3.68	3.60	3.80	3.44	3.84	3.52	3.16	3.36
7	3.50	3.50	3.79	3.79	3.79	3.58	3.38	3.38
8	3.44	3.40	3.52	3.36	3.56	3.44	2.88	3.52
9	3.46	3.43	3.71	3.75	3.86	3.61	3.29	3.64

#### 3.2.1 RQ1

In order to investigate RQ1, Groups A and B are compared and contrasted. Group A includes 28 learners with higher proficiency, and Group B consists of 31 learners with lower proficiency. In both classes, the quality of teaching is equal. That is, the same teacher taught them based on the same syllabus, textbooks and activities (reading-based activities). Table 4 shows the differences of vocabulary, grammar, reading and listening test scores as well as total scores. The results of t-tests indicate statistical significance in all tests ( $p < .01$ ).

Table 4: Test results of Groups A and B

TESTS	Group	N	M (perfect scores)	SD
Total	A	28	195.3 (300)	31.3
	B	31	146.2(300)	23.2
V	A	28	34.3(50)	6.3
	B	31	28.0(50)	6.1
G	A	28	30.8(50)	6.1
	B	31	22.8(50)	5.3
R	A	28	60.9(100)	15.4
	B	31	45.2(100)	12.2
L	A	28	69.4(100)	12.4
	B	31	50.1(100)	10.9

By comparing Group A (Figure 1 and Table 3) with Group B (Figure 2 and Table 5), it can be noted that high-proficiency learners tend to be more perceptive to their own self-regulatory control in the classroom.

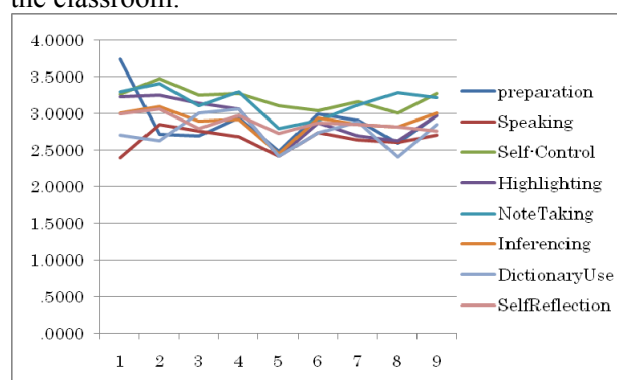


Figure 2: Mean score movement: Group B ( $N = 31$ )

Table 5: Mean score movement: Group B (N=31)

Session	Item1	2	4	5	6	7	8	9
1	3.74	2.39	3.26	3.23	3.29	3.00	2.71	3.00
2	2.72	2.84	3.47	3.25	3.41	3.09	2.63	3.06
3	2.69	2.76	3.24	3.14	3.10	2.90	3.00	2.79
4	2.94	2.68	3.26	3.06	3.29	2.91	3.06	2.97
5	2.48	2.41	3.10	2.41	2.79	2.45	2.41	2.72
6	3.00	2.73	3.03	2.87	2.90	2.93	2.73	2.87
7	2.91	2.64	3.15	2.70	3.12	2.85	2.88	2.85
8	2.59	2.59	3.00	2.63	3.28	2.81	2.41	2.81
9	2.97	2.70	3.27	2.97	3.21	3.00	2.85	2.76

### 3.2.2 RQ2

There is a slight upward trend in inferencing and self-reflection (Items 6 and 9). However, almost all the items fluctuate.

By turning our attention to sessions 6 and 8 in Group A and sessions 5 and 8 in Group B, it can be found that learners forget to use strategies in class. One thing in common among these sessions is that a time lag of and at least three weeks have occurred since the previous session. It is noteworthy that if they haven't used strategies in class for a while, their self-regulation may not work effectively.

## 4 Findings

Overall, our results supported Zimmerman's claims. (1) Self-regulated learners tend to be successful learners and receive higher proficiency. (2) Strategic control loop (Feedback loop) is very important. With intervening blanks, basic-level learners tend to forget to use strategies or become less perceptive to one's self-regulation. Thus we, as language teachers, need to keep in mind that learners need to sustain their effort to use a wide variety of strategies in order to achieve learning goals.

## 5 References

- O'Malley, J.M. & Chamot, A.U. (1990). *Learning Strategies in Second Language Acquisition*.  
 Oxford, R.L. (1990). *Language Learning Strategies: What Every Teacher Should Know*.  
 Tsutsui, E. et al. (2007). Why Do We Need to Teach Communication Strategies to Japanese EFL Learners? *Proceedings of the 13th Conference of Pan-Pacific Association of Applied Linguistics*, 192-195.  
 Tsutsui, E. et al. (2008). Creating an online feedback System: L2 Learners' Strategy Use. *Proceedings of the 13th Conference of Pan-Pacific Association of Applied Linguistics*, 166-167.

Vandergrift, L. (2003). Orchestrating strategy use: Toward a model of the skilled second language listener. *Language Learning*, 53 (3), 463-496.

Zimmerman, B.J. (1989). A social cognitive view of self-regulated learning. *Journal of Educational Psychology*, 88, 397-410

Zimmerman, B, J. (1990). Self-regulating Academic Learning and Achievement: The Emergence of a Social Cognitive Perspective. *Educational Psychology Review*, 2(2), 1990

Zimmerman, B.J. & Schunk, D.H. (2008). Motivation: An Essential Dimension of Self-Regulated Learning. *Motivation and Self-Regulated Learning*, 1-30

## Appendix A: Self-Checklist

Please answer the following questions on the scale of one to four.

- (4) I strongly agree.  
 (3) I agree.  
 (2) I disagree.  
 (1) I strongly disagree.

Before class,

(Item 1) I did enough preparation for the class (for a quiz or homework).

During class,

(Item 2) I spoke out loud in English during the class (for chorus reading and shadowing).

(Item 3) I actively participated in pair work.

(Item 4) I tried to control myself and fight off my negative feelings. (I tried not to whisper or sleep.)

(Item 5) I highlighted my textbook (underlining and using circles).

(Item 6) I jotted down notes.

(Item 7) I inferred the meaning of unknown words from the context.

(Item 8) I made full use of a dictionary.

(Item 8') I used a dictionary to check my comprehension.

After class,

(Item 9) I tried to be self-reflective, going through all I learned and trying to catch the gist of the class.

# An analysis of basic verbs in Japanese junior and senior high school textbooks

Norifumi Ueda<sup>1</sup> and Kota Wachi<sup>2</sup>

<sup>1</sup>School of Foreign Language, Mejiro University, <sup>2</sup> Shiba Junior/Senior High School

uedan@mejiro.ac.jp, rsc36319@nifty.com

## Abstract

The purpose of this presentation is to examine semantic and syntactic environments that basic verbs used in junior/high school English textbooks contain to develop a vocabulary test.

In the learning process, L2 learners should integrate the conceptual differences and separate concepts between L1 and L2 words, which cause difficulties in acquiring L2 vocabulary. This problem happens especially in the acquisition of L2 synonyms and polysemys. It is beneficial for L2 learners to identify what semantic information hinder learning vocabulary.

In this study, we selected synonyms and polysemys of English verbs, especially basic verbs, used in Japanese junior and senior high school textbooks. We analyzed semantic and syntactic environments which differentiate one word from others with similar meanings. The results show what kinds of semantic and syntactic information affect L2 learners' vocabulary learning and imply what kinds of information should be tested to examine L2 vocabulary knowledge.

## Keywords

SLA, textbook analysis, syntactic pattern

## 0 Introduction

In the learning process, L2 learners should integrate the conceptual differences and separate concepts between L1 and L2 words, which cause difficulties in acquiring L2 vocabulary. (Ijaz, 1986; Dong et al., 2005) This problem happens especially in the acquisition of L2 synonyms and polysemys. Besides the acquisition of the conceptual differences between L1 and L2, L2 learners should also learn related information such as syntax, collocation and usage in context in order to use synonyms and polysemys precisely.

To acquire L2 vocabulary, L2 learners in

Japan make use of lexical input mainly through textbooks.

The purpose of this study is to examine English basic verbs in Japanese textbook, especially to examine (1) what kinds of syntactic and semantic information are contained the basic verbs in Japanese textbooks and (2) what kinds of semantic information are involved in the verbs.

## 1 Vocabulary knowledge and vocabulary acquisition

Many kinds of lexical knowledge are hypothesized such as syntax, collocation and usage in context. (Nation 1990) Richards (1976) described vocabulary knowledge as follows:

1. The native speaker of a language continues to expand his vocabulary in adulthood, whereas there is comparatively little development of syntax in adult life.
2. Knowing a word means knowing the degree of probability of encountering that word in speech or print. For many words we also know the sort of words most likely to be found associated with the word.
3. Knowing a word implies knowing the limitations on the use of the word according to variations of function and situation.
4. Knowing a word means knowing the syntactic behaviour associated with the word.
5. Knowing a word entails knowledge of the underlying form of a word and the derivations that can be made from it.
6. Knowing a word entails knowledge of the network of associations between that word and other words in the language.
7. Knowing a word means knowing a semantic value of a word.
8. Knowing a word means knowing the different meanings associated with a word.

There are many studies on vocabulary acquisition in the field of psycholinguistics and bilingual studies, where lexical relations between concepts and words

in L1 and L2, and L2 acquisition process have been examined. In the concept-word relation study, it is suggested that L2 learners should integrate the conceptual differences between two languages in the process of acquiring the other vocabulary. (Ding et al., 2005; Ijaz, 1986; and Ueda, 2007)

In any case, lexical information such as syntax, collocation and usage in context is necessary and essential in the process of acquiring L2 vocabulary. Hence, it is worth examining what kind of lexical information L2 learners get. In this study, we examine textbooks used in a Japanese junior and senior high school in order to examine lexical information.

## 2 Textbook Analysis

### 2.1 Textbooks as materials

We analyzed the textbooks used a private junior and senior high school in Tokyo. Table 1 shows the names of them.

We chose 12 basic English verbs: hear, listen, see, look, watch, gaze, stare, say, talk, speak, tell and utter. This is because when learners acquire these verbs, they should integrate the conceptual differences and separate concepts between L1 and L2 words.

Table 1: Names of the textbooks used in junior/senior high school

Grade	Name of Textbook (Publishing Company)
7th	<i>Just English Is It 1</i> , Renew (Robertson)
8th	<i>Just English Is It 2</i> , Renew (Robertson)
9th	<i>Just English Is It 3</i> , Renew (Robertson)
10th	<p>&lt;Former Half&gt;</p> <p>Reading class (English I): <i>CROWN English Series I</i>, New Edition (Sanseido)</p> <p>Grammar class (Oral Communication I): <i>Forest Extensive English Grammar in 47 Lessons</i>, 5th Edition (Kiriara Shoten)</p> <p>&lt;Latter Half&gt;</p> <p>Reading class: started to use supplemental workbooks</p> <p>Grammar class: Keep using the text above</p>

#### 2.1.1 Textbooks used in Junior high school

In this private school, the textbooks used for 7th to 9th grade students were extremely high level in the vocabulary size and the speed of teaching the contents, compared with the authorized textbooks widely used in public schools in Japan. (These textbooks were adopted by many high-level private schools in Tokyo area.)

These textbooks were composed of grammar-based contents, and students learned the

basic rules of grammar of English with the basic sentences at first, and worked on a variety of exercises with the drills based on each rule. After the exercises, students were to try the reading materials. Each unit of the textbooks had one reading material which contains the grammatical rules focused in each unit. Students tried to translate the sentences in the material, contain the grammatical rules focused in the unit, into Japanese as a confirmation of their understanding.

#### 2.1.2 Textbooks used in high school

Almost the same pattern of contents was used in the textbooks for the grammar class in 10th grade except reading materials. In each lesson, subjects were to learn some new grammar rules with basic sentences and exercises based on the grammatical rules focused in each lesson. As for the confirmation of their understanding, students tried not to translate English sentences in Japanese but to put some Japanese sentences into English.

As for the textbooks for the reading class in the same grade, it consists of many reading materials, grammar exercises, conversation practices, etc. However, subjects learned only the reading materials in their reading class so that they can use the grammar rules in the reading process. In the classroom, teachers didn't explain the rules of grammar minutely, instead, focused on the techniques to read rapidly and effectively. Translation into Japanese was sometimes assigned as confirmation these techniques for reading.

### 2.2 How much input the learners from the textbooks through the classroom.

We will review how much input the students in junior and high school get through textbooks. In junior high school of subjects, 5 credits are assigned to the 7th and 8th grade, 6 credits to 9th grade. Therefore, students took at least one English class a day when they were in junior high school. Subjects learned 2 more classes per week than students in public junior high schools (Students are to take 3 credits of English in junior high school, following the National curriculum guideline for defining the basic standard for education.).

In senior high school, Subjects took two English classes; English I (mainly for reading) and Oral Communication I (mainly for learning grammar). Each class has 3 credits in the 10th grade, following the National curriculum guideline. In the 11th and 12th, subjects are to take two classes in each grade; English I/Reading (4-5 credits to each class) and Writing (2 credits to each class).

As a whole, subjects take much more classes of English; therefore, they have lots of

opportunities to get input from textbooks.

### 2.3 Method for analysis

We analyzed the textbooks by using a lexical analysis software, *WordSmith Ver. 5*. We examined syntactic and semantic environment. As for the syntactic environment, syntactic frameworks were analyzed, and about semantic environment, arguments and their structures were examined. In this study, we followed the frameworks used in Ueda et al. (1999)<sup>1</sup>

## 3 Results

We found that sentences for gaze, stare and utter were not available. Hence, we got the sentences of *talk*, *say*, *tell*, *watch*, *look*, *hear* and *speak*.

### 3.1 Results of *talk*, *say*, *speak* and *tell*

First, we will see the syntactic and semantic structure of the synonyms: talk, say speak and tell. Table 2 shows the syntactic framework of *talk*. The most frequently used framework is “talk + about”, and its semantic structure is [Agent [+human] \_ (about) + Theme [-human]].

Table 2: Syntactic framework of *talk*.

framework	number
talk + to + Noun	4
talk + about + Noun	10
talk + with + Noun	4
talk#	1
total	19

# means no word come after the word.

Table 3: Semantic structure of *talk*

Semantic structure
Agent [+human]_ (about) Theme [-human]
Agent [+human]_ (to) Recipient [+human]
Agent[+human] _ (with) commitative [+human]

Table 4 shows syntactic framework of *say*. Three types of the syntactic framework have almost the same frequencies: “say + to”, “say + complement” and “say + noun”. And their semantic structures are followings in Table 5: [Agent [+human] \_ (to) Receiver [+human]], [Agent [+human]\_ (complement)], and [Agent [+human] \_Theme [-human]].

In the case of *tell*, there are the most various kinds of syntactic structures in Table 6. The most frequent structures are “tell + about + Noun” and

Tell +Direct Object (DO)”. The semantic structure, on the other hand, is the same in both syntactic structures: [Agent [+human] \_ Theme [-human]].

Table 4: Syntactic framework of *say*

framework	number
say + (that) complement	19
say + DO	13
idioms	2
total	48

# means no word come after the word and DO, Direct Object. Complement contains quotation in this case.

Table 5: Semantic structure of *say*

semantic structure
Agent [+human]_ (to) Receiver [+human]
Agent [+human]_ complement
Agent [+human] _Theme [-human]

Table 6: Syntactic framework of *tell*

form	number
tell + (that) complement	3
tell + DO + about + Noun	9
tell + DO	7
tell + person + to do	1
tell DO + to IO	2
tell + DO + IO	6
total	28

# means no word come after the word; Do, Direct Object; and IO, Indirect Object, respectively

Table 7: Semantic structure of *tell*

semantic structure
Agent [+human]_ (complement)
Agent [+human]_ (about) Theme [-human]
Agent [+human] _ Patient [+human]
Agent [+human] _Theme[-human], Patient [+human]
Agent [+human]_ Patient [+human], (about) Theme [-human]

The most frequent syntactic structure in *speak* is “speak + DO”, and its semantic structure is only one: [Agent [+human] \_ Theme [-human]].

Table 8: Syntactic framework of *speak*

framework	number
speak + DO	31
Speak + to	7
speak#	4

# means no word come after the word

From the results, the most frequent structure for *say*, *tell*, and *speak* is “V + DO”, and their semantic structure is the same: [Agent [+human] \_ Theme

<sup>1</sup> In this study, we analyze the present forms of verbs only.

[-human]]. Careful examination shows that the kinds of nouns for Theme are different. The nouns used as the object in *speak* are related to a language.(for example, English, Spanish, etc.) In the case of tell, the “fact”, a “truth” and a “lie” were used as the object. This information differentiates the usage among these three verbs.

### 3.2 Results of *watch*, *see* and *look*

There were various frameworks found in look in Table 11, whereas the semantic structure is only one:[Agent [+human] \_ ]. In the case of watch, just one syntactic and semantic structure were found in Table 9 and 10. *See* had various syntactic framework like *look*, and had similar semantic structure to *watch*. But in the case of *see*, theme can have the attributes [+human] or [-human]. Hence we can say, in the textbooks, these three verbs have completely different semantic and syntactic environments.

Table 9: Syntactic framework of *watch*

Syntactic framework	number
watch +DO	22

Table 10: Semantic structure of *watch* and *look*

Semantic structure	
watch	Agent [+human] _ Theme [-human]
look	Agent [+human] _
see	Agent [+human] _ Theme [+/-human]

Table 11: Syntactic framework of *look*

framework	number
look + at	36
look + like	1
look + adj	3
look#	8
total	48

# means no word come after the word.

Table 12: Syntactic framework of *see*

from	number
see +DO	63
see +DO + PLACE	2
see + DO + TIME	2
see + DO + Ving	2

### 3.3 Results of *hear* and *listen*

The most frequent syntactic framework of hear is “hear + complement” and “hear + DO”. On the other hand, in the case of listen, the typical syntactic structure is “listen to”. In the point of semantic structure, hear and listen had the same structure.

Table 13: Syntactic framework of *hear*

framework	number
hear + that complement	4
hear +DO	5
hear + person + V	1

Table 14: Syntactic framework of *listen*

framework	number
listen + to	7
listen#	2

# means no word come after the word.

Table 15: semantic structure of *hear* and *listen*

Semantic structure	
hear	Agent [+human] _ Theme [+human]
	Agent [+human] _ Theme [-human]
listen	Agent [+human] _

## 4 Conclusion

Through the analysis, we found what kinds of input the Japanese L2 learners obtain from the textbooks, and what information to differentiate semantic and syntactic information in acquiring English synonyms is included in the textbooks.

The information we found in this study can be useful in developing L2 vocabulary test to examine the breadth of vocabulary knowledge, “How much do you know of a word?” And also, the results of this study can be helpful for teaching vocabulary. Moreover, the results imply what kinds of information should be tested to examine L2 vocabulary knowledge.

## References

- Dong, Y., Gui, S. and MacWhinney, B. (2005). Shared and Separate meanings in the bilingual mental lexicon. *Bilingualism: Language and Cognition* 8, 3. 221-238.
- Ijaz, H. (1986). Linguistic and cognitive determinants of lexical acquisition in a second language. *Language Learning*, 36, 401- 451.
- Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. Heinle and Heinle Publishers: Boston.
- Richards, J. (1976). The role of vocabulary teaching. *TESOL Quarterly*. 10-1, 77-89.
- Ueda, N., Miyasaka, N. and Yamazaki, T. (1999). L2 Input and Output: Do L2 Textbooks Contain Enough Syntactic Information? *Essays on English Language and Literature*. Pp.21-27. The English Language and Literature Society of Waseda University

# Listening Comprehension Strategies of Turkish EFL Students and the Effects of Proficiency Level on Strategy Use

Funda Gercek

The School of Foreign Languages, Anadolu University

fgercek@anadolu.edu.tr

## Abstract

This study attempts to identify the type and frequency of listening comprehension strategy used by students learning English as a foreign language and determine the effects of proficiency level on strategy use. The participants were Turkish. A total 369 students studying English in the Intensive Language Program at the Preparatory School at Anadolu University participated in the study. The students were given a survey of listening comprehension strategies adapted from Oxford (1990). The data was gathered from Elementary, Low-Intermediate, Intermediate and Upper-Intermediate language level students.

Findings indicated that all language proficiency level students used listening comprehension strategies in the medium range in general.

The results indicated that the students at the low-intermediate level used overall strategies significantly more than the students at elementary level.

Memory and Cognitive strategies were used by low-intermediate students more often than elementary level students. Intermediate level students used cognitive strategies significantly more often than the students at elementary level. The students at the upper-intermediate level used memory strategies significantly more often than the ones at elementary level. When the students in low-intermediate level was compared with the ones in intermediate level and the students in low-intermediate group was compared with the ones in up-intermediate level, no significant differences were found between their means for individual SILL categories.

## Keywords

Listening Comprehension, Listening Comprehension Strategies, Learning Strategies.

## Introduction

Learning strategies are defined by Oxford(1990:8) as “the specific actions taken by the learner to make learning easier, faster, more enjoyable, more self

directed, more effective and more transferable to new situation”. Nevertheless, studies learning have mostly focused on reading, writing speaking strategies of EFL and ESL learners. Learning strategies in listening skill have widely ignored by researchers. Vogely (1995) claims that “Although our knowledge about Listening Comprehension and Language Learning Strategies has increased in the past decade a gap still exist between research and theory and the classroom reality. We still need research that document empirically the relationship between what theory says and what learners actually know and more importantly do. More specifically information is needed on the skills and strategies learners actually bring or do not bring to the Listening Comprehension Task”. Since the importance of listening comprehension cannot be underestimated and given the lack of research on how learners acquire listening ability, the main focus of this study will be;

What listening comprehension learning strategies are employed by elementary, low-intermediate, intermediate and upper-intermediate level EFL students at Anadolu University Prep classes. This study attempts to identify the listening comprehension learning strategies used by Turkish students studying English in Intensive Language Program at Anadolu University. It also examines the relationship between language proficiency and listening comprehension strategies.

## 1 Aim and the Scope of the Study

This study attempts to identify listening comprehension learning strategies used by Turkish students of English studying at the Preparatory School Prep classes at Anadolu University.

It also examines the relationship between language proficiency and listening comprehension learning strategies frequency use.

This study emphasizes the following five strategies; memory, cognitive, metacognitive, and social-effective, based on the taxonomy of Oxford’s (1990) Language Learning Strategies.

This study was limited to students attending

elementary, low-intermediate, intermediate and upper-intermediate level classes at Intensive English Program of Preparatory School of Anadolu University, Eskişehir. Beginner and Advanced group classes are not included in this research. Communicative strategies were not included. This study was also limited to the listening skill. Sex differences, age and English background of the subjects were not considered.

## 1.1 Research Questions

The research questions to be answered in this study are as follows;

1. Which listening comprehension learning strategies do Turkish students of English use in general?
2. Is there a relationship between language proficiency and listening comprehension learning strategies frequency use?

## 2 Methodology

### 2.1 Subjects

A total 368 students attending the Intensive English Program at Anadolu University, Eskişehir participated in the study. The students were given a placement test examination and the results of the exam revealed that their language levels ranged between beginner and advanced. The students were given the first placement test at the beginning of the Fall Term and the same students were given the second placement test at the beginning of the Spring Term of the same year. The students had taken 16 weeks English instruction according to the grades that they had in the first exam in the Fall term. After the second placement exam that they had, they were placed in different levels. During the study they were enrolled in Intensive English Program in Prep School where they received between 31 and 20 hours of English classes in all skills weekly. At the time of the study, they had all completed a period of 32 weeks English instruction.

### 2.2 Instruments

The following instruments were used to determine different factors such as language level and strategy use of students.

#### 2.2.1 Placement Test

Students language levels were determined according to the scores on a standard placement test (Michigan Placement). The goal of the test was to measure the language level of the students. The highest score possible in Michigan placement examination is 100. There are four parts of the exam; 20 listening comprehension questions, 30

grammar and structure items, 30 vocabulary items and 20 reading comprehension items. Students were given to two hours to complete the test.

#### 2.2.2 Listening Strategy Inventory

To determine the listening strategies students use SILL (Strategy Inventory for Language Learning) was used. An adopted version of Oxford's (1990) SILL for listening comprehension was administered. Only the items related to listening skill in SILL were given to the subjects in this study. The Inventory consisted of 26 statements on a 5 point Likert scale ranging from 1('Never or almost never') to 5 (Always almost true of me'). These items were translated to Turkish by the researcher herself. To determine the linguistics validity of Turkish version of Oxford's (1990) SILL, 20 randomly chosen students from all levels were asked to comment on the statements in terms of language use, grammatical points, punctuation, content and sentence structure. Again, revisions were made based on these students' comment. The modified version was given to 10 teachers of listening comprehension. The teachers were native Turkish speakers.

### 2.3 Procedure

All data were collected during the last week of the spring semester. Students were informed that SILL was designed to identify their listening comprehension strategies. They were recommended to give responses to the statements as sincerely as possible. There was no time limit for the SILL.

#### 2.3.1 Data Analysis

Students responses (ranging from 1 to 5) to the SILL items were scored on the SILL answer sheets. Responses to each strategy group were totaled and then the results were divided by the number of each item in each strategy group to provide an average for each group. These averages were rounded off the nearest hundredth, i.e. 2.25.

Significant variation in mean strategy use across the entire SILL and in differences in mean strategy use in the memory, cognitive, compensation, metacognitive, affective and social strategies categories as related to language level was determined by using one-tailed t-test, employing Stat View 512+ statistical software. The results were considered significant if they reached the level .05 or below.

## 3 Results

### 3.1 Differences in Overall Strategy Use

Frequencies of listening comprehension strategy

use in general revealed that the students participated in the study used the listening comprehension strategies in the medium range (mean 3.13). Students' overall means for individual SILL categories were in the medium use range as well.

**Table 1: Summary of Variation in Listening Overall Strategy Use by Language Level**

Language Level	n	Mean
Elementary	96	3.08
Low -intermediate	81	3.19
Intermediate	76	3.11
Upper-intermediate	116	3.14

**Table 2: Summary of Variation of Means in Six Strategy Group in Listening Comprehension Use by Language Levels**

Level	Memory	Cognitive	Compensation	Metacognitive	Affective	Social
Elementary	2.86	2.99	3.33	3.31	2.67	3.30
Low-int	3.02	3.24	3.43	3.39	2.81	3.30
Intermediate	2.86	3.20	3.41	3.26	2.76	3.30
Up-int	3.02	3.24	3.32	2.77	3.13	3.14
<b>Mean</b>	<b>2.94</b>	<b>3.16</b>	<b>3.37</b>	<b>3.18</b>	<b>2.84</b>	<b>3.26</b>

**Table 3: Summary of Variation in Listening Overall Strategy Use by Proficiency Levels**

Comparison of Proficiency Level	n	mean	t value and significant level
Elem vs L-int	96/81	3.08 /3.19	1.757 p<.04
Elem vs Int	96/76	3.08 /3.11	ns
Elem vs Up-int	96/116	3.08 /3.13	ns
L-int vs Int	81/76	3.19 /3.08	ns
L-int vs Up-Int	81/116	3.19 /3.31	ns
Int vs Up-Int	76/116	3.11 /3.13	ns

To determine the general listening strategies students use, all 26 items were investigated. These 26 items were classified as memory, cognitive, compensation, metacognitive affective and social strategies.

Overall strategy use, according to the t-test results did not vary significantly by language levels of elementary –intermediate ( $t = .423$ ,  $d.f = 170$   $p = .3363$ ), elementary –upper-intermediate ( $t = .868$ ,  $d.f = 210$ ,  $p = .1931$ ); Low-intermediate – intermediate ( $t = -1.467$ ,  $d.f = 155$ ,  $p = .0722$ ); upper-intermediate – low-intermediate ( $t = .887$ ,  $d.f = 155$ ,  $p = .1882$ ); intermediate – up-intermediate ( $t = .364$ ,  $d.f = 190$ ,  $p = .558$ ). However overall strategy use according to the t-test result varies significantly by low-intermediate and elementary groups ( $t = 1.757$ ,  $d.f = 175$ ,  $p = .0403$ ;  $p < .05$ )

### 3.2 Variation in Use of the Six Categories of Strategies

The t-test results as table 4 presents, indicated that the students at low-intermediate level used memory strategies significantly more often than the students at elementary level ( $t = 1.733$ ,  $p < .04$ ). The means were 2.85 and 3.02 respectively.

The difference between low-intermediate and elementary students' uses of cognitive strategies (means were 3.24 and 2.99 respectively) was close to significance ( $t = 2.33$ ,  $p < .01$ ).

In three of the four strategy groups (memory, metacognitive, social) low-intermediate level

students' means were higher than those of elementary level students yet the difference did not approach statistical significance. Low-intermediate level students' means for the six SILL categories were in the medium use range, except for compensation strategies (mean 3.43) which was barely over the upper limit of the range. Elementary level students' means were again in the medium use range.

As table 5 shows, indicated that the students at Intermediate level used cognitive strategies significantly more often than the students at elementary level ( $t = 1.969$ ,  $p < .02$ ). The means were 3.20 and 2.99 respectively. In two of other five strategy categories (memory and social) Elementary and Intermediate group students' means were equal. The mean was 2.85 respectively the same for both groups. The other two strategy categories (compensation and affective) Intermediate level group students' means were slightly higher than elementary group students' means. For metacognitive strategies the reverse was true. However the differences were far from being significant. Intermediate level students' means for the six SILL categories were in the medium range, except for compensation strategies (mean 3.41) which was over the upper limit of the range. Elementary level students' means were in the medium range again.

**Table 4: Summary of Variation in Elementary and Low-Intermediate Students' Means for Strategy Categories**

SILL Category	Elementary	Low-Int	t Value and Significance Level
Memory	3.02	2.85	1.733 ( $p < .04$ )
Cognitive	2.99	3.24	2.33 ( $p < .01$ )
Compensation	3.32	3.43	ns
Metacognitive	3.31	3.39	ns
Affective	2.66	2.80	ns
Social	3.29	3.30	ns

When the students in elementary group compared

with the ones in Upper-Intermediate group as table 6 presents, the students at upper-intermediate level used memory strategies significantly more often than the students at elementary level ( $t=1.797$ ,  $p<.03$ ). The means were 3.01 and 2.85 respectively. Moreover, the students at Upper-intermediate level used cognitive strategies significantly more often than the students at elementary level ( $t=2.728$ ,  $p<.003$ ). In two of the other five strategy categories (compensation and metacognitive), elementary and upper-intermediate group students' means are equal. The means were 3.32 and 3.31 respectively the same for both groups. The other strategy category (affective) Upper-intermediate level group students mean was slightly higher than elementary group students' mean. For social strategies the reverse was true. Although the differences were far from being significant both level students' means for the six SILL categories were in the medium range.

Table 5: Summary of Variation in Elementary and Intermediate Students' Means for Strategy Categories

SILL Category	Elementary	Low-Int	t Value and Significance Level
Memory	2.85	2.85	ns
Cognitive	2.99	3.20	1.969 ( $p<.02$ )
Compensation	3.32	3.41	ns
Metacognitive	3.31	3.26	ns
Affective	2.66	2.75	ns
Social	3.26	3.30	ns

Table 6: Summary of Variation in Elementary and Upper-Intermediate Students' Means for Strategy Categories

SILL Category	Elementary	Low-Int	t Value and Significance Level
Memory	2.85	3.01	1.797 ( $p<.03$ )
Cognitive	2.99	3.24	2.728 ( $p<.003$ )
Compensation	3.32	3.32	ns
Metacognitive	3.31	3.31	ns
Affective	2.66	2.77	ns
Social	3.29	3.12	ns

Table 7: Summary of Variation in Low-Intermediate and Intermediate Students' Means for Strategy Categories

SILL Category	Elementary	Low-Int	t Value and Significance Level
Memory	2.80	2.75	ns
Cognitive	3.30	3.30	ns
Compensation	3.19	3.11	ns
Metacognitive	3.02	2.85	ns
Affective	3.24	3.20	ns
Social	3.43	3.41	ns

Table 8: Summary of Variation in Intermediate and Upper-Intermediate Students' Means for

## Strategy Categories

SILL Category	Elementary	Low-Int	t Value and Significance Level
Memory	2.85	3.01	ns
Cognitive	3.20	3.24	ns
Compensation	3.41	3.32	ns
Metacognitive	3.26	3.31	ns
Affective	2.75	2.77	ns
Social	3.30	3.12	ns

When the students in Low-intermediate group was compared with the ones in Intermediate no significant differences were found between their means for individual SILL categories (See Table 7). Although Low-intermediate group students' means were higher than those of the students in intermediate group, the differences between their means were higher than those of the students in Intermediate group, the differences between their means were found to be non-significant.

When the students in Intermediate group was compared with the students in Upper-intermediate no significant differences were found between their means for individual SILL categories (See Table 8)

## 4 Conclusion

The results of this study show that the students' overall listening strategy mean across the whole SILL and general means for individual SILL categories were in the medium range, which means that the students used listening strategies "sometimes".

The results indicated that the students at the low-intermediate level used overall strategies significantly more than the students at elementary level.

Memory and Cognitive strategies were used by low-intermediate students more often than elementary level students. Intermediate level students used cognitive strategies significantly more often than the students at elementary level. When the students in low-intermediate level and the students in intermediate level was compared with the ones up-intermediate level, no significant differences were found between their means for individual SILL categories.

## References

- Oxford, J.R. (1990). *Language Learning Strategies: What Every Teacher Should Know*. Heinle and Heinle Publishers.
- Vogely, A. (1995) Perceived Strategy Use During Performance on Three Authentic Listening Comprehension Tasks, *The Modern Language Journal*. 75, 1, 41-56

# Students' perception of intercultural communication through new distance learning model

Bok-Myung Chang<sup>1</sup> and Eunjung Bang<sup>2</sup>

<sup>1</sup>Department of English, Namseoul University, <sup>2</sup>Namseoul University

bmchang@nsu.ac.kr, eunjungbk@yahoo.co.kr

## Abstract

This study tries to examine students' perception of the newly introduced model of NWCCDLP (Namseoul-Waseda Cross-Cultural Distance Learning Project) which focuses more on improving students' communicative competence.

In the newly introduced NWCCDLP, the course is divided into five different discussion sessions. For each discussion topic, there is additional one to two preparation class and students are also expected to do reading or research outside of class time. Then students have Joint Class for each topic. After the Joint Class, students submit reflection papers three times during the whole semester.

The results from students' reflection papers and private interviews reveal that

- 1) the preparation class and homework are very helpful to get the idea of the discussion topic so that students can contribute to discussion meaningfully
- 2) the preparation class and homework also help students be aware of various kinds of intercultural communication styles which result in upgrading the interactions through CCDL
- 3) students' satisfaction for the course is higher than out-of-class text chatting that students have done with Japanese students
- 4) this new type of distance learning model (preparation class + reading or research + in class voice chatting + reflection paper) is very effective in language learning especially in enhancing students' communicative competence.

## Keywords

Communicative Competence, English as an International Language (EIL), Distance Learning, NNS(non-native speakers)-NNS Interaction

## 1 Introduction

As communicative competence gains more attention in language learning, the needs for employing various ways of enhancing communicative competence have also been emphasized by language practitioners.

For example, even though English is not used as an official language in Korea, many people invest their money and energy in English education. One of the main purposes of English education in Korea has become raising students' communicative competence. However, despite increasing emphasis on communicative competence in English education, Korean people lack of communicative competence comparing to their grammar and vocabulary level. So teachers and practitioners in Korea are trying to develop a variety of ways enhancing students' communicative competence.

Meanwhile, as there are more NNS-NNS interactions taking place worldwide, people begin to recognize the importance of English as an International Language (EIL), so language practitioners in the countries of Expanding Circles have devised various models of intercultural distance learning. These kinds of cross-cultural distance learning model provide students with chances to use English as an International language, and by doing so, students can also strengthen their communicative competence.

NWCCDLP (Namseoul-Waseda Cross-Cultural Distance Learning Project) is such a model which is designed to improve students' communicative competence by using English as an international language through intercultural communication between students in Namseoul and Waseda universities.

Especially, NWCCDLP which is newly introduced this semester is focusing more on improving students' communicative competence. This model consists of preparation class, reading or research, in class voice chatting discussion and reflection papers for each discussion topic and this schedule is expected to maximize students' contribution to the discussion. Also, students are taught to be aware of a variety of communication skills so that students can recognize communication differences between individuals and cultures.

In other words, students can practice different kinds of intercultural communication styles in addition to negotiating various cultural issues

through the new model. By doing so, students are expected to widen their communicative competence.

In this light, we try to examine students' perception of intercultural communication through this newly introduced model of NWCCDLP which increases the chance to use English as an International Language for their communication

In this vein, the following research questions are formed:

- 1) Do the preparation class and homework activate students' intercultural communication in the Joint Class?
- 2) Is this course more effective in improving students' communicative competence than other out-of-class chatting?

## 2 Literature Reviews

The notion of communicative competence, first proposed by Hymes (1974), not only includes linguistic competence but also a range of other sociolinguistic and conversational skills that enable the speaker to know how to say what to whom, when. In other words, students with high communicative competence are thought to have ability to apply grammatical, discourse and cultural knowledge to communicate effectively in particular contexts for particular purposes (Nunan, 1999).

In the meantime, Crystal (1997) argues that currently, there are seventy-five territories where English is spoken either as a first language (L1), or as an official (i.e. institutionalized) second language (L2) in fields such as government, law and education.

From the Crystal's figures, it is implied that English is commonly used not only by its native speakers but also by people whose native language is not English.

However, there are still other non-native speakers of English for whom English has no official function within their countries, who probably number around one billion and whose proficiency levels range from reasonable to bilingual competence.

They are speakers referred to as speakers of English as an International Language, reflecting the fact that these English users from, for example, Europe and Japan, speak English more frequently as a contact language among themselves than with native speakers of English (Jenkins, 2003).

According to her, non-native English speakers who speak English as a second or foreign language are thought to be over one billion while speakers of English as a native language are thought to number around 350 million.

Even though the figure is not uncontroversial, it is clear that non-native speakers of English outnumber native speakers of English.

This can be also seen in Kachru's three-circle model of World Englishes. As English is spoken worldwide, Kachru (1992) divides World Englishes into three concentric circles, the Inner Circle, the Outer Circle and the Expanding Circle. According to him, the Expanding Circle occupies the biggest circle representing that there are more cases of communication between non-native speakers of English than with native speakers of English and NNS-NNS interaction is critical in successful L2 learning.

Cross-cultural Distance Learning is one of the possible models which can enhance NNS-NNS interactions. NWCCDLP provides students in the Expanding Circle with a new environment to strengthen their communicative competence by raising their sociolinguistic and conversational skills through interactions with other NNSs.

Warschauer (1997) claims that the special features of online communication such as computer-mediated, many-to many, time- and place-independent, usable across long distances provide an impressive array of new ways to link learners. He also argues that these features make online learning a potentially useful tool for collaborative language learning when viewed in the context of sociocultural learning theory, which emphasizes the educational value of creating cross-cultural communities of practice and critical inquiry.

## 3 Methods

The newly introduced cross-cultural distance learning model designed by Waseda University is adopted for this study. This section begins by describing the course.

### 3.1 Course Description

In the newly introduced NWCCDLP, the whole semester is devoted to the course. The course is divided into five different discussion topics. For each discussion topic, there is additional one to two preparation class and students are also expected to do reading or research outside of class time. Through the preparation class and homework, students are encouraged to participate in the interaction more actively and contribute to discussion meaningfully. Moreover, students can practice different kinds of intercultural communication styles provided by the textbook in addition to be prepared for negotiating various cultural issues.

Then students have Joint Class (in class voice chatting discussion) for each topic. The students from Namseoul and Waseda Universities are matched as partners and meet as a small group through the Live On system 5 times during the whole semester.

During the discussion, students are encouraged only to speak in English (voice chatting) and use text box only when there is a communication breakdown.

After the Joint Class, students submit reflection papers three times during the whole semester.

For this course, material developed by Waseda University is used as a main textbook through the whole semester.

All of the students in this study have experience of out-of-class text chatting, so we investigate students' perception of two different cross-cultural distance learning models.

### 3.2 Data Collection and Analysis

The data are collected from the students' reflection papers and interviews. Qualitative methodology is used to analyze the data (reflection papers and interviews) from the students who participated in the course.

Reflection papers are submitted three times during the sessions and the private interview is carried out at the end of the semester.

## 4 Results

Here are some of students' reflection papers.

NSU 03141119

The Live On was very interesting to me. Making conversation with different country's people is difficult, but I could try to get them to understand. And I was more comfortable when I talked in English with Japanese students than Korean students. Also, we could know about the other countries. This class was helpful to me to have confident about English.

NSU 04141008

Even though there was hardware problem, I could communicate with real foreigners. I do not have chance to talk to the foreigner, but I could. I was also very relaxed talking to foreigners.

Even though most of the students point out that technical problem causes too much communication breakdown, students are satisfied with the course in general.

However, we are not centering on general advantages of cross-cultural distance learning in this study. In addition to the above reflection papers,

there are many other students who report that the NNS-NNS interaction helps them become more interested in Japanese culture, more confident in using English and so on.

Instead, we are more concerned with the student's perception of the new cross-cultural distance learning model which is focusing more on enhancing students' communicative competence by preparing for the discussion before the Joint class.

Here are other examples of students' reflection papers.

NSU 03141513

We share many opinions. When each other come with another opinion, there were problems of communication. But because of the textbook and preparation class, there were more understanding with each other.

NSU 01141547

When I have preparation class, it was difficult to make a decision about our topic. Because I didn't have much information about topic before the preparation class. However, when I talk with other students in Liveon, it was easier than preparation class because I had some information about our topic which I've prepared in before class.

NSU 01141107

Only the hardest part was the hardware systems. It shows why we need over preparation for these kinds of CCDL classes

Analysis of students' reflection papers and interviews regarding two research questions reveal the followings.

- 1) the preparation class and homework are very helpful to get the idea of the discussion topic so that students can contribute to discussion meaningfully
- 2) the preparation class and homework also help students be aware of various kinds of communication styles which results in upgrading the interactions through CCDL.
- 3) students' satisfaction for the course is higher than out-of-class text chatting as students can recognize communication differences between individuals and cultures through the preparation class and homework prior to the real discussion. Moreover, students can have actual chances to speak and concentrate on the course through the whole semester as there is consistency in their learning.
- 4) In general, this new type of distance learning model (preparation class + reading or research + in class voice chatting + reflection paper) is very effective in language learning especially in enhancing students' communicative competence.

## 5 Discussions and Conclusions

As English is increasingly used as a communication tool between NNSs, teachers of English in the Expanding Circle have attempted to investigate various types of models to improve students' communicative competence by using English as an International Language.

NWCCDLP is one of possible models which provide students with the chances to interact with other NNSs of English by using English as an International Language.

The newly introduced NWCCDLP, focusing more on raising students' communicative competence, consists of preparation class, reading or research, in class voice chatting and reflection paper. Through this course, students can participate in the discussion more meaningfully by practicing different kinds of communication styles prior to the real discussion. Also students can prepare for the intercultural communication by doing previous research or reading or even discussion practice with Korean students.

Students announce that the preparation class and homework seem to be more effective in intercultural communication as computer-mediated cross cultural communication is more difficult than the normal face-to-face interaction with the speakers of the same language.

In fact, the preparation class and homework prior to the real discussion enable students to recognize communication differences between individuals and cultures. Moreover, with these preparations, students felt more relaxed and confident during the real discussion so that they participated in the discussion more actively.

The results from the students' reflection papers and interviews imply that this new type of NWCCDP is very helpful in raising communicative competence by not only acquiring English as an International Language through the interaction with Japanese students but also learning various kinds of communication skills.

However, nearly all the students report that there is too much technical problem. Especially, in computer-mediated distance learning, as students have to utilize electronic devices to communicate with each other, technical problem can cause serious communication breakdown.

With these results, this research will help us to improve the environment which can activate the interaction between non-native speakers.

## References

- Alptekin, C. (1993). Target-language culture in EFL materials. *ELT Journal*, 47(2), 136-143.  
 Chang, B. M. (2005). World English acquisition

through CCDL project. *Journal of Pan-Pacific Association of Applied Linguistics*, 9(2), 241-261.

- Chang, B. M. (2008). New environment for acquiring English as an International Language. *Proceedings of the 13th Conference of Pan-Pacific Association of Applied Linguistics*, 129-132.  
 Chapelle, C. A. (2001). *Computer applications in second language acquisition: Foundations for teaching, testing and research*. Calbridge: Cambridge University Press.  
 Crystal, D. (1997). *English as a global language*. Cambridge: Cambridge University Press.  
 Jenkins, J. (2003). *World Englishes*. New York: Routledge.  
 Kachuru, B. (1992). *The other tongue: English across cultures* (2<sup>nd</sup> ed.). Urbana, IL: University of Illinois Press.  
 Nunan, D. (1999). *Second language teaching and learning*. Boston: Heinle & Heinle Publishers.  
 Warschauer, M. (1997). Computer-mediated collaborative learning. *Modern Language Journal*, 81(4), 470-481