

Breaking Theory: New Directions in Applied Linguistics

*Proceedings of the 48th Annual Meeting
of the British Association for Applied Linguistics*

3-5 September 2015

Aston University, Birmingham



Edited by
Tilly Harrison, Ursula Lanvers & Martin Edwardes

Breaking Theory: New Directions in Applied Linguistics

*Proceedings of the 48th Annual Meeting
of the British Association for Applied Linguistics*

3-5 September 2015
Aston University, Birmingham



Edited by
Tilly Harrison,
Ursula Lanvers
& Martin Edwardes



Breaking Theory: New Directions in Applied Linguistics

Proceedings of the 48th Annual Meeting
of the
British Association for Applied Linguistics

*3-5 September 2015
Aston University, Birmingham*

Published by Scitsiugnil Press
1 Maiden Road, London, UK
and produced in the UK

Copyright © 2016

Copyright subsists with the individual contributors severally in their own
contributions.

This publication may be reproduced in part or whole only with
acknowledgement of the source.

First Published 2016

By the British Association for Applied Linguistics
Edited by Tilly Harrison, Ursula Lanvers & Martin Edwardes

*Product Design by Martin Edwardes
Scitsiugnil Press logo design by Jack Bruce*

ISBN: 978-0-9559533-8-5

Breaking Theory: New Directions in Applied Linguistics

Proceedings of the 48th Annual Meeting
of the
British Association for Applied Linguistics

*3-5 September 2015
Aston University, Birmingham*

Contents

Chapter	<i>Authors</i>	Page
1 Exploring EFL Reading as a Microsystem in Social Contexts: From the Ecological Systems Approach	Hiroyuki Matsumoto & Chikako Aoki	1
2 The Smoke and Mirrors of Linguistics: Challenging the hidden metaphors	Martin Edwardes	13
3 Challenging government policy on English language teaching in Japan through collaborative action research: theory and method	Tamiko Kondo	25
4 New directions created by Putonghua and English's "double domination" of Zhuang language in China	Alexandra Grey	37
5 Revisiting the action research cycle: Critically examining three cases of classroom-based language learning inquiry	Gerald Talandis Jr. & Theron Muller	53
6 The Influence of Experience as Non-Native Speakers on Beliefs of Native Speakers	Koichi Shimahara	75
7 Unconscious Attitudes toward L2 Speakers	Ali H. Al-Hoorie	87

Chapter	<i>Authors</i>	Page
8 (Re)thinking the Interface Hypothesis and its Implications for Language Teaching	Joana Teixeira	93
9 L2 Pragmatic Problems: Do Short-Term Study Abroad Experiences Help Learners to Reduce Them?	Yoko Sato	111
10 The implicit role of classroom silence in Japanese EFL contexts	Seiko Harumi	121
11 Some conversational challenges of talking to children on the radio: The case of the Chris Evans Breakfast Show	Ian Nakamura	131
12 A comparative study of evaluative acts in classical concert reviews in British and Hong Kong newspapers	Fong-wa Ha	145
13 Do EFL Learners Discriminate Text-Level Main Ideas, Paragraph-Level Main Ideas, and Details During Reading? A Study With Self-Paced Reading	Yoshinobu Mori	163
14 Language-specificity of forensic linguistics: the example of Hungarian	Eszter Tóth	173
15 Vocabulary learning grounded in an ESP community: Design and effect of a basic medical ESP course	Hiroko Suzuki & Miho Fujieda	185
16 An investigation of brain synchrony between students and their teacher during a secondary school EFL lesson	Reiko Yamamoto, Takayuki Nozawa, Hyeonjeong Jeong, Shigeyuki Ikeda, Kohei Sakaki, Ryuta Kawashima, Craig Smith & Yasushige Ishikawa	201
17 ‘An international journal publishing high quality, original research’: self-evaluative categories in journal descriptions	Davide Simone Giannoni	213

Chapter	<i>Authors</i>	Page
18 A comparison study of teacher supported peer feedback and teacher feedback for EFL writing	Dr Huahui Zhao	231
19 On-line and Off-line EFL Sentence Processing of Reduced and Unreduced Relative Clauses: Evidence from Think-aloud Protocols and Translation Task	Natsumi Tanaka	255

1 Exploring EFL Reading as a Microsystem in Social Contexts: From the Ecological Systems Approach

Hiroyuki Matsumoto & Chikako Aoki
Hokkai Gakuen University

Abstract

The present study investigates EFL reading as an ecological system in social contexts, focusing on strategy use, motivation, and beliefs. Based on Bronfenbrenner (1979), ecological systems for reading development consist of the microsystem, the mesosystem, the exosystem, and the macrosystem. A set of questionnaires was used to gather data from 40 university students enrolled in two EFL reading courses. The relationships of their microsystem (EFL reading factors) with their upper systems (socio-environmental factors) were examined using the Spearman rank-order coefficients and Mann-Whitney tests with exact estimation. The results indicated the possibility that general relationships exist between the participants' EFL reading and socio-environmental factors, and that their socio-environmental factors, particularly the mesosystem, are more influential on their strategy use than their motivation and beliefs. Specifically, it appears that the participants' present and past L1 reading experiences at home and in the library positively influenced their present EFL reading strategy use. These results support the validity of considering EFL reading as an ecological system in social contexts.

Introduction and Background

Previous L2 reading research has extensively focused on cognitive aspects of reading such as comprehension processes, or models of reading, proficiency, and strategy use. For example, a plethora of studies have investigated the relationship of L2 readers' strategy use with their proficiency, revealing that their strategy development can contribute to the enhancement of reading proficiency (e.g., Carrell, 1989; Grabe, 1991; Kern, 1989). Yet, these cognition-centered studies have not included reading motivation, which has been considered one of the most important affective factors in L2 reading (Grabe, 2009). To date, L2 research on reading motivation has been limited.

Although L2 reading motivation has not been extensively explored, Mori (2002) examined the motivational structure of EFL reading using a questionnaire based on Wigfield and Guthrie's (1997) theory of L1 reading motivation for English-speaking children. In Mori's study, four factors similar to those identified in the L1 theory were confirmed for EFL university students: intrinsic motivation, extrinsic motivation, the importance of reading, and reading efficacy. This indicates that the motivational structure of L2 reading for EFL students is close to that of L1 reading for children, and that both intrinsic and extrinsic motivation are key concepts in L2 reading. L1 research on reading motivation (e.g., Guthrie & Wigfield, 2000) also implies that students with high intrinsic motivation tend to be active readers, high achievers, and adept users of reading strategies.

The social contexts of L2 reading have seldom been discussed. Although Grabe (2009) examined social factors which can affect L1 reading development for school children in the United States and other areas (see Table 1), these factors must be re-examined to determine if they are applicable to the reading of university L2 students, who are at a different developmental stage and learn in different school systems and social settings. Nevertheless, these L1 social factors can provide helpful implications for the study of social influence on L2 reading. In several L1 studies, the socioeconomic status (SES) of parents is considered a predictive factor for their children's cognitive abilities, including reading comprehension and academic achievement (Fletcher & Reese, 2005; Noble, Farah, & McCandliss, 2006). The home literacy environment (HLE), and in particular, the involvement of mothers, have also been reported as influential factors in the development of children's L1 reading ability (Dieterich, Assel, Swank, Smith, & Landry, 2006; Mistry, Biesanz, Chien, Howes, & Benner, 2008; Niklas & Schneider, 2013).

Socioeconomic status (SES) of the parents and family
Values and attitudes of the family and wider social group
Educational levels of the parents
Patterns of interactions between parents and children
Uses of language and literacy resources in the home (home literacy environment: HLE)
Literacy resources available in the community and schools
Training of teachers
Policies and practices of educational institutions of all types

Table 1: Major social factors affecting L1 reading
Excerpted from Grabe (2009, p156)

Purpose of the Study

To address the issues above, the present study investigated EFL reading as an ecological system in social contexts, focusing on strategy use, motivation, and beliefs. This study is a preliminary to exploring the relationships between EFL reading and social factors. We adopted the ecological systems approach (Bronfenbrenner, 1979) as a framework to study the influence of social factors on EFL reading. Based on this approach, reading development consists of four ecological systems (see Table 2 and Figure 1).

Microsystem	Individuals' learning to read: strategy use, motivation, and beliefs
Mesosystem	Individuals' learning environments: families and schools
Exosystem	Communities that indirectly affect individuals' learning
Macrosystem	Sociocultural and/or ideological influences on individuals' learning

Table 2: Four ecological systems for reading development
From Bronfenbrenner (1979)

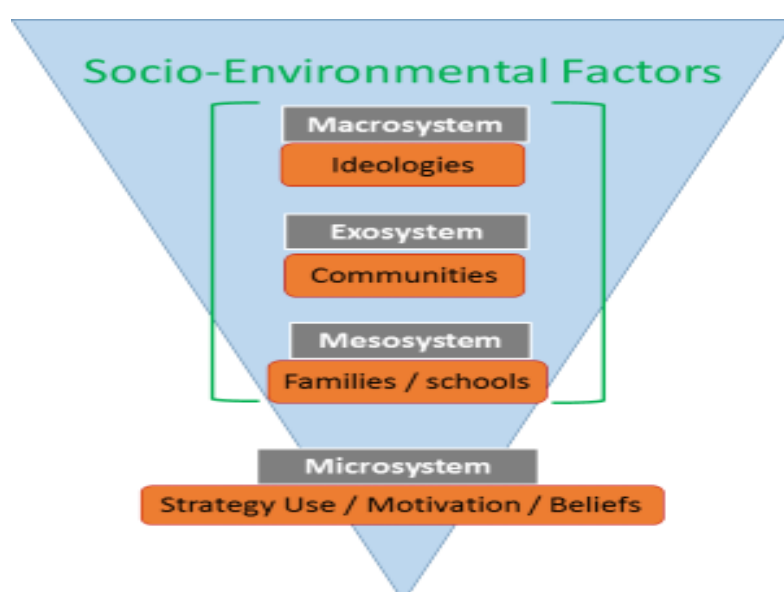


Figure 1: Framework of this study based on the ecological systems approach

Participants

The participants were 40 Japanese first-year university EFL students, enrolled in two classes of a reading-centered English course, in a semester of 15 classes of 90 minutes each. They were selected on the basis of their accessibility and proximity to the researcher (non-probability convenient sampling). Their ages were 18 to 20 years old, and the proportion of the sexes was almost equal. The purpose of the course was for students to acquire basic English reading comprehension.

Method and Instruments

A set of questionnaires was used to gather data from the participants. All the items were rated on a five-point Likert scale ranging from 1, *strongly disagree* to 5, *strongly agree*. All the instructions and statements were written in Japanese, the participants' first language, so that they could easily understand the questionnaires.

To evaluate the microsystem (their strategy use, motivation, and beliefs), three existing scales (refer to Matsumoto, Hiromori, & Nakayama, 2013) were slightly changed. The questionnaire consisted of 38 items in total: 14 for strategy use, 12 for motivation, and 12 for beliefs.

For strategy use, the SORS (Survey of Reading Strategies: Sheory & Mokhtari, 2001) was revised. It was originally composed of three subsets of metacognitive strategy (10 items), cognitive strategy (12 items), and support strategy (six items) to measure ESL reading strategy use. However, some items needed revision to match the EFL context of Japan, and finally the following subscales were created: main idea strategy (five items; $\alpha = .824$; e.g. *I read for the purpose of understanding the main idea*), reasoning strategy (three items; $\alpha = .829$; e.g., *I check to see if my inferences are right or wrong*), adjusting strategy (three items; $\alpha = .816$; e.g. *I read slowly and carefully when the text is difficult*), and monitoring strategy (three items; $\alpha = .820$; e.g. *I check my understanding when I come across new information*). The scale's factor model was validated by confirmatory factor analysis (GFI = .915, AGFI = .864, RMSEA = .056): main idea strategy ($\beta = .964, p < .001, R^2 = .929$), reasoning strategy ($\beta = .690, p < .001, R^2 = .476$), adjusting strategy ($\beta = .550, p < .001, R^2 = .308$), and monitoring strategy ($\beta = .857, p < .001, R^2 = .735$).

For motivation, the MRQ (Motivation for Reading Questionnaire: Wigfield & Guthrie, 1997) was modified. The original MRQ consisted of three subscales: extrinsic motivation (15 items), intrinsic motivation (14 items), and reading efficacy (eight items). For this study, some items were revised to match the Japanese EFL context, holding its original three subscales: extrinsic motivation (five items; $\alpha = .789$; e.g. *I would like to gain a higher score than my classmates on reading tests*), intrinsic motivation (four items; $\alpha = .718$; e.g. *I like challenging books written in English*), and reading efficacy (three items; $\alpha = .784$; e.g. *I can read English proficiently with effort*). The validity of the scale's factor model was confirmed (GFI = .981, AGFI = .957, RMSEA = .022): extrinsic motivation ($\beta = .624, p < .001, R^2$

= .389), intrinsic motivation ($\beta = .573$, $p < .001$, $R^2 = .329$), and reading efficacy ($\beta = .793$, $p < .001$, $R^2 = .629$).

For beliefs, a belief scale originally developed by Ueki (2002) was revised for use. The scale consisting of three subcategories (orientations of strategy, environment, and effort: six items for each) to estimate learner beliefs among Japanese high school students was adjusted slightly to fit to Japanese university students: strategy orientation (four items; $\alpha = .778$; e.g. *It is effective to think of the way of learning*), environment orientation (four items; $\alpha = .715$; e.g. *It is important to be in a good environment for learning*), and effort orientation (four items; $\alpha = .836$; e.g. *It is effective to keep working hard with effort*). Confirmatory factor analysis validated the scale's factor model (GFI = .994, AGFI = .981, RMSEA = .000): strategy orientation ($\beta = .793$, $p < .001$, $R^2 = .629$), environment orientation ($\beta = .636$, $p < .001$, $R^2 = .404$), and effort orientation ($\beta = .708$, $p < .001$, $R^2 = .501$).

Mesosystem ($\alpha = .696$)	1. I read Japanese books at home now.
	2. I read Japanese books at home in my childhood.
	3. My family read me Japanese picture books in my childhood.
	4. My family likes reading Japanese books.
	5. I attended a cram school(s) to pass university entrance exams.
	6. I decided to attend a cram school(s) by myself.
	7. My family's ideas about education influenced me.
	8. My family often enjoys aspects of foreign culture.
	9. High school English classes were useful.
	10. English teachers at school influenced my ideas about learning English.
Exsosystem ($\alpha = .689$)	11. Someone in my community influenced my ideas about learning English.
	12. I read Japanese books in the library now.
	13. I read Japanese books in the library in my childhood.
Macrosystem ($\alpha = .682$)	14. English education in Japan is effective.
	15. University entrance exams in Japan provide good English questions.
	16. English textbooks used in Japanese secondary schools are useful.
	17. English is essential in the global community.

Table 3: Questionnaire items for socio-environmental factors

To evaluate the three upper systems (socio-environmental factors), a questionnaire of 17 items was used (Table 3). In the development process (in a preceding pilot survey), 17 items were chosen from a pool of 40 statements prepared, on the basis of the existence of correlations with any of the microsystem items: mesosystem (10 items; $\alpha = .696$), exosystem (three items; $\alpha = .689$), and macrosystem (four items; $\alpha = .682$).

Data Analysis

The relationships of the microsystem with the three upper systems were examined by the Spearman rank-order coefficients for correlations and Mann-Whitney tests with exact estimation for differences.

Results and Conclusion

Results are shown below in order of strategy use, motivation, and beliefs, each with the correlations and differences included.

Table 4 shows the correlations between the participants' reading strategy use and socio-environmental factors, and also the differences in their reading strategy use by socio-environmental factors. Several items in the three components of strategy use correlated significantly with their mesosystem and exosystem: four of these items concerned family reading experiences, one item school English teachers, and three items reading experiences in the library. The general trend of these correlations is that the participants' reading experiences in L1 Japanese both at home and in the library were related to their present strategy use in reading L2 English. This is confirmed by the significant differences in the participants' strategy use according to their mesosystem, although their exosystem made only one difference in their strategy use. Given the priority of their L1 reading experiences to their L2 strategy use, this can be interpreted as evidence that the participants' mesosystem (and possibly exosystem) positively influenced their strategy use in reading English.

	Mesosystem					Exosystem	
	1	2	3	4	10	12	13
	I read Japanese books at home now.	I read Japanese books at home in my childhood.	My family read me Japanese picture books in my childhood.	My family likes reading Japanese books.	English teachers at school influenced my ideas about learning English.	I read Japanese books in the library now.	I read Japanese books in the library in my childhood.
Main Idea Strategy							
I identify the topic sentence of each paragraph.					.325*		
					.049*		
The primary purpose of my reading is main idea comprehension.							.344*
							.788
I pay attention to the organization of a text.						.382*	.385*
						.057	.214
I differentiate the main idea of a text from its detailed information.	.365*	.430**	.313*				
	.014*	.018*	.025*				
Reasoning Strategy							
I infer the meaning of a text backward while reading.	.376*	.372*	.360*				
	.047*	.092	.038*				
I guess the meaning of unknown words from the context while reading.		.429**	.371*	.366*			
		.030*	.045*	.027*			
I infer the meaning of a text based on my background knowledge.						.389*	.344*
						.011*	.277
Adjusting Strategy							
I adjust my reading speed according to the difficulty of a text.	.404**	.400*	.542**				
	.037*	.061	.010*				

Note. $N = 40$, Spearman's rank correlation coefficients were calculated, and significant ones are listed (upper values); then accordingly, Mann-Whitney tests were performed to evaluate the differences of strategy use by the high mean rank (4 and 5 on a Likert scale) and low mean rank (1 to 3 on the same scale) participants, and two-tailed exact probabilities are listed (lower values).

* $p < .05$. ** $p < .01$.

Table 4: Correlations and differences about the participants' strategy use and socio-environmental factors

Table 5 shows the correlations and differences concerning the participants' reading motivation and socio-environmental factors. Both their extrinsic and intrinsic motivations have a few items that correlated significantly with their mesosystem, exosystem, and macrosystem. This result indicates that the participants' motivation to read L2 English had some relationship with their socio-environmental background. However, this tendency is not necessarily confirmed by the differences. For example, the macrosystem item mentioning that using English is indispensable in the globalized society made only one difference in the participants' intrinsic motivation, although the mesosystem appears to have made some differences to their extrinsic and intrinsic motivation. Despite the complexity of the results, this shows the possibility that the participants' mesosystem had more influence on their L2 reading motivation than their macrosystem.

	Mesosystem			Exosystem	Macrosystem
	6	8	9	11	17
	I decided to attend a cram school(s) by myself.	My family often enjoys aspects of foreign culture.	High school English classes were useful.	Someone in my community influenced my ideas about learning English.	English is essential in the global community.
Extrinsic Motivation					
Reading English well is one goal of my learning English.			.356*		.344*
			.019*	✓	.103
I like to score high on the reading section of English proficiency tests.				✓	.343*
					1.000
I want myself recognized as a good English reader.			.430**		
			.020*		
I like to learn something by reading English. ✓	.321*				.380*
	.461				.079
Intrinsic Motivation					
Reading English is fun.		.349*	.382*		
		.020*	.016*		
I am happy if I find something new while reading English.		.319*	.319*	.318*	.347*
		.075	.045*	.396	.141
Increasing knowledge by reading English is fun.					.394*
					.043*

Note. $N = 40$. Spearman's rank correlation coefficients were calculated, and significant ones are listed (upper values); then accordingly, Mann-Whitney tests were performed to evaluate the differences of motivation by the high mean rank (4 and 5 on a Likert scale) and low mean rank (1 to 3 on the same scale) participants, and two-tailed exact probabilities are listed (lower values).

* $p < .05$. ** $p < .01$.

Table 5: Correlations and differences about the participants' motivation and socio-environmental factors

Table 6 consolidates the results of the correlations and differences in the participants' beliefs about strategy, environment, and effort. Their mesosystem and macrosystem correlated significantly with their beliefs. More specifically, the mesosystem items resulted in positive relations, and the macrosystem items in negative relations. A negative relation in their macrosystem, for example, shows that students, who think of English textbooks used in their secondary schools as useful, do not agree that it is effective to think of how to read English. The positive relations in their mesosystem were confirmed by the differences; however, the negative relations in their macrosystem were not strengthened by the differences. This implies that their mesosystem was more influential on their beliefs than their macrosystem.

	Mesosystem				Macrosystem	
	3	5	6	10	15	16
	My family read me Japanese picture books in my childhood.	I attended a cram school(s) to pass university entrance exams.	I decided to attend a cram school(s) by myself.	English teachers at school influenced my ideas about learning English.	University entrance exams in Japan provide good English.	English textbooks used in Japanese secondary schools are
Strategy Orientation						
It is effective to think of how to read.						-.354* .685
It is effective to change how to read voluntarily.					-.328* .263	-.324* .582
It is necessary to think of how reading goes well.	.399* .012*			.316* .037*		
Environment Orientation						
It is important to be in a good reading environment.						-.339* .317
It is important to be in an environment where I can ask questions.				.343* .022*		
Effort Orientation						
It is effective to keep working hard with effort.		.370* .083	.446* .052			

Note. $N = 40$. Spearman's rank correlation coefficients were calculated, and significant ones are listed (upper values); then accordingly, Mann-Whitney tests were performed to evaluate the differences of beliefs by the high mean rank (4 and 5 on a Likert scale) and low mean rank (1 to 3 on the same scale) participants, and two-tailed exact probabilities are listed (lower values).

* $p < .05$. ** $p < .01$.

Table 6: Correlations and differences about the participants' beliefs and socio-environmental factors

In conclusion, the following possibilities are indicated.

- 1) General relationships existed between the participants' microsystem (EFL reading factors: strategy use, motivation, and beliefs) and their upper systems (socio-environmental factors).
- 2) Their mesosystem (families and schools) was more influential on their microsystem than their exosystem (communities) and macrosystem (ideologies).
- 3) Their socio-environmental factors (in particular, their mesosystem) were more influential on their strategy use than their motivation and beliefs.
- 4) The participants' present and past L1 reading experiences at home (mesosystem) and possibly in the library (exosystem) positively influenced their present EFL reading strategy use.

As with studies on the influence of HLE (home literacy environment) on L1 reading development (Dieterich, Assel, Swank, Smith, & Landry, 2006; Mistry, Biesanz, Chien, Howes, & Benner, 2008; Niklas & Schneider, 2013), we found the possibility of HLE as an influential factor for the participants' EFL reading, in particular, their strategy use. By contrast, we could not find

anything suggesting the influence of SES (socioeconomic status) of their parents, because of the deficits of our questionnaire items for socio-environmental factors.

We believe that the findings in this study all support the validity of considering EFL reading as an ecological system in social contexts. However, we need more in-depth research to explore the influences of socio-environmental factors on EFL reading through multiple data sources including qualitative information.

Acknowledgements

This research was supported by JSPS (Japan Society for the Promotion of Science) *KAKENHI* (Grant-in-Aid for Scientific Research) Number 25370693.

References

- Urie Bronfenbrenner (1979).** *The ecology of human development: Experiments by nature and design*. Harvard University Press: Cambridge, MA, USA.
- Patricia L. Carrell (1989).** Metacognitive awareness and second language reading. In *The Modern Language Journal* 73, pp121-134.
- Susan E. Dieterich, Mike A. Assel, Paul Swank, Karen E. Smith & Susan H. Landry (2006).** The impact of early maternal verbal scaffolding and child language abilities on later decoding and reading comprehension skills. In *Journal of School Psychology* 43, pp481-494.
- Kathryn Fletcher & Elaine Reese (2005).** Picture book reading with young children: A conceptual framework. In *Developmental Review* 25, pp64-103.
- William Grabe (1991).** Current developments in second language reading research. In *TESOL Quarterly* 25, pp357-406.
- William Grabe (2009).** *Reading in a second language: Moving from theory to practice*. Cambridge University Press: Cambridge, UK.
- John T. Guthrie & Allan Wigfield (2000).** Engagement and motivation in reading. In Michael. L. Kamil, Peter B. Mosenthal, P. David Pearson & Rebecca Barr (Eds.), *Handbook of reading research* 3. Lawrence Erlbaum Associates: Mahwah, NJ, USA, pp403-422.
- Richard G. Kern (1989).** Second language reading strategy instruction: Its effects on comprehension and word inference ability. In *The Modern Language Journal* 73, pp135-149.
- Hiroyuki Matsumoto, Tomohito Hiromori, & Akira Nakayama (2013).** Toward a tripartite model of L2 reading strategy use, motivations, and learner beliefs. In *System* 41, pp38-49.
- Rashmita S. Mistry, Jeremy C. Biesanz, Nina Chien, Carollee Howes & Aprile D. Benner (2008).** Socioeconomic status, parental investments, and the cognitive and behavioral outcomes of low-income children from immigrant and native households. In *Early Childhood Research Quarterly* 23, pp193-212.

- Setsuko Mori (2002).** Redefining motivation to read in a foreign language. In *Reading in a Foreign Language* 14, pp91-110.
- Frank Niklas & Wolfgang Schneider (2013).** Home literacy environment and the beginning of reading and spelling. In *Contemporary Educational Psychology* 38, pp40-50.
- Kimberly G. Noble, Martha J. Farah & Bruce D. McCandliss (2006).** Socioeconomic background modulates cognition-achievement relationships in reading. In *Cognitive Development* 21, pp349-368.
- R. Sheorey & K. Mokhtari (2001).** Differences in the metacognitive awareness of reading strategies among native and non-native readers. In *System* 29, pp431-449.
- Rie Ueki (2002).** Structure of high-school students' beliefs about learning. In *Japanese Journal of Educational Psychology* 50, pp301-310.
- Allan Wigfield & John T. Guthrie (1997).** Relations of children's motivation for reading to the amount and breadth of their reading. In *Journal of Educational Psychology* 89, pp420-432.

2 The Smoke and Mirrors of Linguistics: Challenging the hidden metaphors

Martin Edwardes

King's College London

Martin.edwardes@btopenworld.com

Abstract

In 1957, Noam Chomsky introduced a new concept to linguistics. Not the idea that a single engine lies behind all the different human languages – von Humboldt had discussed this in 1836. Chomsky's idea was that we needed to understand only the mechanism of that single engine to study all languages.

Chomsky's Generative Linguistics programme directed linguistic study away from individual languages, and toward a universal system of language which generates all languages. Instead of descriptive accounts of grammar, phonology, lexis and semantics, linguistics became an essentialist study of grammar and syntax; and instead of studying language users, Generative linguists concentrated on inward analysis of their own systems of language. Linguistics descended into a fog of grammatical forms, investigated largely by self-regard: smoke and mirrors.

This paper looks at three conventionalised metaphors in academic linguistics, representing unproven assumptions which we seldom question: **LANGUAGING IS BEING HUMAN**; **LANGUAGING IS BEING CLEVER**; and **LANGUAGING IS A GOOD THING**. These metaphors have emerged directly from the Generativist enterprise, and they are often assumed in the work of linguists who otherwise abhor the Generativist tradition. They are key metaphors because they have allowed us to pretend that all languages are representations of a still-unspecified acme we call language, that this unspecified substance is what sets us apart from other animals, and that the substance is a significant feature in all our species specialisms. The paper discusses why we now need to abandon these metaphors, and the assumptions behind them, and reconfigure our subject area in terms of the wider topics of communication, socialisation, and human evolution.

Introduction

Linguistics, as the study of languages, is an ancient topic. It can be traced back at least to the Greek philosophers, and probably even earlier to the Vedic scholars of Sanskrit. However, before the European Renaissance it

was not assumed that this was an exclusively human activity: languaging seemed to be what every animal did when they were communicating. All we needed to do to understand the inner worlds, or minds, of other animals was to interpret the language they were signalling in. Mythology, fable and children's tales are replete with nonhuman animals communicating with humans, or with each other, in ways that can only be described as linguistic.

Indeed, in the mythologies of early civilisations (and in modern pre-literate hunter-gatherer groups) the difference between humans and other animals is a matter of degree rather than an unbridgeable distance. Animals advise (e.g. Athena's Owl; Odin's ravens, Huginn and Muninn; and the Celtic salmon of Llyn Llyw); they deceive (e.g. the Serpent in the Garden of Eden; Red Riding Hood's Big Bad Wolf; and Anansi the trickster spider); and they are key figures in our cultures (e.g. the Pelasgian snake-god Ophion; the Egyptian cat-goddess Bastet; and Hanuman, the Hindu monkey god). There was a belief in a constant exchange between human and animal form, with the assumption that the individual's physical form was a surface effect rather than an intrinsic quality.

Then came René Descartes (1649 [1998]), and his "humans are special" viewpoint. According to Descartes, only humans have souls, so only humans have minds, so only humans are capable of making choices; other animals are just automata, meat machines for our use. Suddenly, human languages were cut free from any relationship with the rest of natural signalling; and with no comparators to constrain our definitions, we began to think of language as a discrete thing with its own qualities, and not just another means to a communicative end.

This reification of language culminated in the rise of Noam Chomsky and Generativism. Generativism is a set of theories which emerged from the Cartesian enterprise (Chomsky, 2009) via three assumptions: first, that all languages are products of a single, innate and exclusively human capacity to produce languages; second, that languages represent something special and complex about human cognition; and third, that language is the primary mechanism in the rise to dominance of *Homo sapiens*. The concept of all languages being based on a single set of rules was not strictly original when Chomsky published the first agenda for Generativist linguistics in 1957; what was new was the idea that this base set of rules was the only explanation needed for languaging. Wilhelm von Humboldt had proposed in 1836 that all languages shared a common source, form and rule base, an argument that accorded with the Tower of Babel story in the Bible; but he did not envisage

any cognitive function exclusive to humans beyond the capacity to communicate in a complex way. The driving engine of language was the desire to communicate, not an innate structure of thought which existed even without communication.

What counts as language?

The Generativist assumptions have one big advantage over other views of language: they turn the study of language into a measurable and natural phenomenon which can be explained mathematically (Chomsky, 2013, p35). However, this advantage is bought at a high price. The list of features which differentiate language from other communication systems has become increasingly restricted over the nearly 60 years of Generativist Linguistics, and is now composed of a single recursive cognitive mechanism known as MERGE. Yet MERGE (which is often characterised, using von Humboldt's definition, as "the infinite use of finite means") is not how most people would define their everyday use of language. For the majority, languaging is a background communicative process they use for specific, delimited messaging; it is infinite neither in terms of use or need. And there is good reason to accept this mundane definition of language: non-linguists produce the vast majority of language in the world, and they do it both competently and consciously. It seems likely, therefore, that they have some useful personal definitional knowledge about it. If Chomsky's (1965, pp18-27) view is true, that the route to linguistics is through the intuitions of the "native speaker" (whatever one of those might be), then the unfussy introspection of the vast majority of language is users is as valid as the convoluted introspection of the linguistics professor.

Indeed, many linguists nowadays question the value of the narrow definition of language that MERGE provides (e.g. Crystal, 2005, pp232-233). However, they also often attempt to retain the structural precision brought by seeing language as a separately definable, natural thing (e.g. *ibid*, chs1-3). This approach leads to all kinds of definitional issues (Crystal calls them "edges"), because language, if it is actually a separate thing, is deeply embedded in the human communicative system. This system encompasses limbic signals, autonomic signals, automatic signals, conventional signals, conventionalised signals and socio-cultural signals, which between them use a range of communicative modes – vocal, other audial, gestural, other visual, tactile, and possibly even pheromonal. The human communicative system is, like other communicative systems in nature, multimodal and highly redundant: we signal the same thing simultaneously in different ways. Is it really

sensible to try to tease out one part of that system and elevate its role above the other parts?

This last question is something that both integrationist linguists (e.g. Toolan, 1996, ch3) and distributionist linguists (e.g. Cowley, 2011) would agree needs to be asked. It is the counter to the Generativist view that language is a measurable, natural, and above all, discrete entity. Instead, it views language as a fuzzily-defined function which may, or may not, represent a differentiable cognitive system; but which, either way, cannot be properly studied except as part of a whole-body, integrated communication system.

Being Human?

The first assumption of Generative Linguistics, that all languages are products of a single, innate and exclusively human capacity to produce languages, generates the first hidden metaphor of linguistics, that **LANGUAGING IS BEING HUMAN**. Humans have language because language is encoded in our genes; if other animals had the same genetic coding then they would have language; but no other animal has language; so the language genes are specific to humans; and only humans can use language. The natural conclusion of this long and somewhat circular chain of reasoning is the unvoiced assumption that **LANGUAGING IS BEING HUMAN**.

This assumption has become a truism of linguistics – and, indeed, of the wider scientific world. The British Science Museum website says:

“Human beings are the great communicators of the animal world. They are the only living creatures that use language – words or symbols that represent objects, actions, qualities, feelings and ideas. Other animals communicate in much less complex ways.”¹

Yet when we test the claims in this short extract, we find that they are all questionable. The great communicators? True, in the modern world we are surrounded by instances of language; but is that the natural human state, or is it a recent product of our modern culture? Do ants, for instance, which are constantly surrounded by the pheromonal communications of their nest, regularly seek relief from the hubbub in meditation or solitude? If not, then the award for being the greatest communicators probably belongs to one of the many eusocial species, and not us.

¹ <http://www.sciencemuseum.org.uk/WhoAmI/FindOutMore/Yourbrain/Whatisspecialabouthumanlanguage.aspx>

So are we the only living creatures that use language, when it is defined as “words or symbols that represent objects, actions, qualities, feelings and ideas”? Clearly not, as nonhumans who have been introduced to human language keep on demonstrating. Whether we define the communicative capacities of these nonhumans as language or not, we have to accept that Kanzi the bonobo understands the symbol-object-meaning relationship that composes a nounlike word, and the symbol-action-meaning relationship that composes a verblike word (Segerdahl *et al*, 2005); Alex the grey parrot understood the symbol-attribute-meaning triad that composes an adjectivelike word (Pepperberg, 1999); and Rico the collie dog understood the symbol-object-identifier nature of naming – even when he was faced with novel symbols and novel objects which had to be paired (Kaminski *et al*, 2004). The capacities of animals as guests in the human communicative universe is impressively languagelike in many respects.

The final claim, that the communication of other animals is less complex than human communication, could well be true; but, equally, it could well be false. We do not currently have good evidence of what, or even how, most animals communicate. It was only recently discovered that elephant signalling includes a large component of subsonic calls transmitted through the earth and not through the atmosphere (Günther *et al*, 2004). Even more recently, Clay *et al* (2015) have shown that bonobos have a multifunctional peep call, which seems to be a general attention-getter. It is used in a range of circumstances, and has a functional flexibility that was formerly thought possible only in human communication. To make claims about the complexity of nonhuman communication would seem to be premature at this stage, but it is a predictable consequence of the unspoken assumption that **LANGUAGING IS BEING HUMAN**.

So what about the last bastion of difference espoused by Generativism, the recursive nature of MERGE? Is this truly the Rubicon between language and nonhuman communication? Pinker and Jackendoff (2005) are among many who see problems with this approach. While they are definitely of the view that language is peculiarly human, they see it as the product of a series of evolutionary developments which are driven by communicative need, not cognitive serendipity. They take the view that there are many factors contributing to human communicative systems; and, after showing the genetic improbability of MERGE as an evolutionary outcome, they say:

“The alternative in which language is an adaptation for the communication of knowledge and intentions ... is consistent with behavioural and genetic evidence that language shows multiple signs of partial specialization for this

task rather than grafting one component (recursion) onto a completely unchanged primate base.” (p231).

For Pinker and Jackendoff, language cannot be the product of a single, discrete evolutionary event; and MERGE has no plausible evolutionary story to tell. The current consensus in language evolution studies is that Pinker and Jackendoff are likely to be right (e.g. Tallerman, 2014).

Being Clever?

The second assumption of Generative Linguistics, that languages represent something special and elaborate about human cognition, generates the second hidden metaphor of linguistics, that **LANGUAGING IS BEING CLEVER**. This is detectable in the Science Museum’s final statement, that nonhuman communication is “much less complex”: there seems to be the implication (the use of “much less complex” rather than “simpler”) that the complexity of language requires a more advanced form of cognition than other communication systems. Language is complex because we have the cognitive capacity to allow it to be so; the cleverness we need to produce language is intimately tied to our general human cleverness, and this is attested by our many human-specific accomplishments.

So what do we base our understanding of human cleverness on? Usually we point at our nervous system, and particularly our brain; but it is not the basic neurophysiological attributes of *Homo sapiens* that seem to be behind the “smarts”. The human brain does not have the greatest cortical complexity in the animal kingdom (some of the cetaceans are more complex); it is not the largest (elephants are larger); it does not even have the greatest ratio against body size (Neanderthals narrowly beat *Homo sapiens* in that contest, and modern *Homo sapiens* has a brain about 5% smaller than our ancestors of only 30,000 years ago with no diminution of body size). Only the complex formulation of the Encephalisation Quotient (EQ), developed to explain rather than measure the relative cognitive complexity of humans (Jerison, 1973), shows us as consistently better than all other animals – although the suspiciously large difference between us and our nearest competitor (66%) emphasises that the EQ is an explanation after the facts, not of them. Suzana Herculano-Houzel (2012) has recently shown that, on many of the measures we use to differentiate human brains, there is no real difference from other animals; and in some cases there is no real justification to consider them as valid measures of cognitive capacity in the first place. Humans may use language because of the type of brain we have, but that does not mean that our type of brain has to be cognitively exceptional in the natural world.

Yet there is something about humans that does seem to be clever. We have become the dominant species on the planet, established complex cultures, conquered space, invented computers ... The list is long. All of these achievements can be defined as clever because they indicate a species-specific capacity for complex organisation. Indeed, we humans do seem to be exceptional in our organisational cleverness, as shown by our willingness to cooperate, our trust, and our capacity to rely on others. Does this mean, however, that language drove the organisational cleverness, or even that language was a necessary component of it? Could language have been an outcome – or even a side-effect – of the communicative methods we used to support our organisational cleverness?

We can say, in terms of intraspecies organisation, that humans are certainly clever (although probably not exceptionally so); but we cannot say with any certainty that language, out of all of our cooperative strategies, is a key marker of that cleverness; and we certainly cannot designate language as the cause, or even an index, of that cleverness.

A Good Thing?

The third assumption of Generative Linguistics, that language is the primary mechanism in the rise to dominance of *Homo sapiens*, generates the third hidden metaphor of linguistics, that **LANGUAGE IS A GOOD THING**. Like the cleverness metaphor, this relies heavily on the assumption that being human involves some kind of distancing from the rest of nature: there are things (usually complicated things) that only humans can do. Typing the phrase “only humans can” into Google produces a fascinating list of some of those things, including: “have morality”; “think about thinking”; “have insight”; “dance”; “make ice-cream soda”; “make a tree into art”; and “understand the creation of the new”.

However, this list of what only humans can do is more controversial than it sounds. In a series of experiments, Carel van Schaik’s team in Zurich has demonstrated that chimpanzees show behaviours we would class as moral (e.g. von Rohr *et al*, 2015); Beran *et al* (2013) have shown that language-trained chimpanzees do seem to understand what they know; and, in a review of several experiments in the area, Sara Shettleworth (2012) shows that chimpanzees do seem to have their “eureka!” moments. As for dancing, one can only assume the writer has never seen a bird of paradise. This leaves making ice cream sodas, making trees into art, and understanding the creation of the new (whatever that may be) – none of which this particular human can do. Any list of exclusively human characteristics is likely to

produce similar results: either capacities which turn out not to be exclusively human, or capacities which are exclusive to only a subset of humanity.

The third hidden metaphor is perhaps the most pernicious, but least evidenced, of the three: if **LANGUAGING IS A GOOD THING** in and of itself, then having a greater capacity to language would seem to make you better at being human. Chatty innovators and elaborators of language should be more popular than the taciturn essentialist, and therefore more successful at getting their genes into the future. Yet it is simplicity and adherence to convention which are the prized language skills in everyday discourse; and even they take second place to clarity of meaning. As Plato expressed it, “Wise men speak because they have something to say; fools because they have to say something”.

Being linguistically conventional is particularly important. Individuals who adhere to the majority intuitions about language are seen as more central to a social group, while less conformist individuals are marginalised. This is an anthropological phenomenon we see at work throughout our modern societies, and it may well indicate that one of the evolutionary fitness functions of language is to differentiate “us” from “them” – which is not really a Good Thing if language is there to help us cooperate with each other. Socially, we use language differences as a Hogwarts-like “sorting hat”: you speak this way, so your choices in our group are limited to these, and your role is this. However, as individuals such as Baron Sugar of Clapton show, a human can have as much – or as little – grammar, lexis and phonology as they damn well please and still be successful.

Smoke and mirrors?

The three hidden metaphors of language affect the study of linguistics in several ways, most notably in how we define language itself. They selectively point our attention at particular landmarks in the landscape of human communication, allowing us to ignore other equally significant features. For instance, conventionally signed communication is as effective as conventionally spoken communication, and they should be seen as working at the same level. To recognise this, we label them both as language; but is this a valid synthesis? If you are interested in the destination, then travel by air and sea are comparable; but if you are studying the journey then they are quite different. Charles Hockett, when preparing his list of the design features of language (1960), identified the vocal/auditory channel as the first feature because it was “perhaps the most obvious”, having “the advantage – at least for primates – that it leaves much of the body free for

other activities” (p 6). So what changed the validity of this definition (and, in later years, Hockett’s mind)? It was not a diminution of the vocal/auditory channel advantage, but our definition of language: from communication while doing to communication as doing.

Another strange feature of language is that it lumps spoken and written language together, despite the fact that the way the receiver engages with the signal is very different in the two communicative modes. Yet we still exclude pictures from language because ... well, why? Both pictures and writing are composed of segmented, differentiated and hierarchical components, they are both communicative and they both involve a deferred receiver. In fact, the communicative similarities between a picture and a piece of writing would seem to be greater than between a piece of writing and a speech act. So why do we partition the field of human communication in the way we do?

If we separate out a part of our communication system and declare it to be specifically human, reliant on complexity, and a Good Thing, then we do not need to address what it actually is. We can define our object fluidly, without reference to the object itself. If something new appears that is clearly part of what we want to call language, but which has features we have excluded from our definition, we can expand our definition. Equally, if we insist that real language is exclusive to humans, as Hauser, Chomsky and Fitch have done (2002), then we can reduce our definition until it consists of only a single feature, the recursive power of MERGE. This definitional fluidity has generated 50 years of sometimes quite acrimonious debate over what constitutes language – a debate that Lewis Carroll actually settled long ago:

“When I use a word,” Humpty Dumpty said, in rather a scornful tone, “it means just what I choose it to mean – neither more nor less.”(ch6).

Another way?

In 1957, Chomsky attempted to reduce the scope of linguistics to just grammar, when he stated that “linguists must be concerned with the problem of determining the fundamental underlying properties of successful grammars” (p11). Willard Quine (1960, p11) took a very different view, that language is nothing but “a fabric of sentences variously associated to one another and to nonverbal stimuli by the mechanism of conditioned response”. Quine’s view may be unpopularly behaviourist, but it nonetheless presents a stark contrast to Chomsky’s approach. Yet both of these extremes rely on the unspoken assumption that there is a real thing called language; and they both, in different ways, incorporate the three hidden metaphors. What happens if

we take away the thingness of language, and talk about communicative mechanisms instead?

The first thing that disappears is the language-shaped hole in our head. Instead, we have an instinct to communicate which relies on any appropriate cognitive resources to get the job done. Over our lifetimes, we learn new mechanisms which enhance our communicative experience. In this model, language is an arbitrary set of those mechanisms, a contextually useful definition that can change as context changes – but that means we have to define its meaning for each usage. Talking about language as a thing requires careful definition of what that thing is in each case. Effectively, the word “language” is demoted from a concrete noun to an adjectival descriptor.

Another thing that disappears are the concepts of first and second languages, and the contrast between them. Every brain has its own idiolect, which can be composed of mechanisms for communicating with a single group or for communicating with several groups at different levels of proficiency. An individual does not “speak a language”, they communicate in their own way, which can be comprehended more or less effectively by their audience. Second-language learning becomes just another way of incrementing the individual’s communicative mechanisms, it does not need to be seen as an esoteric process that requires a completely new way of thinking and learning.

Where does this leave the three metaphors, **LANGUAGING IS BEING HUMAN**; **LANGUAGING IS BEING CLEVER**; and **LANGUAGING IS A GOOD THING**? First we should not be surprised if other animals can use communicative mechanisms that we have co-opted into a particular definition of language; and we should not be smug if they cannot. Human communication is for humans, by definition; but it uses mechanisms that are generally communicative, and some of those mechanisms will be available to some nonhumans – no big deal. Second, human communication is important in the job of being human, so we should not be surprised if things that humans think are clever turn up in our communication systems; but that does not mean that the communication systems are clever, it means that they are adaptive – they support the job of being human. So, once again, no big deal. Finally, if human communication is important in the job of being human, then it is definitely a Good Thing; but this does not mean that it is a better thing than other communication systems, which are important in the jobs of being other animals. As a natural phenomenon, human communication is just a thing, neither good nor bad – once again, no big deal.

What happens to the smoke and mirrors? Once we realise that grammar is an outcome of languagelike communication, not a cause, we can concentrate on how grammar assists communication, instead of trying to link it to general cognition. It may well be that grammar turns out to be a significant component in the way humans think; but we should start with its role in communication (where brains do need to negotiate to similar mechanisms) rather than in cognition (where they don't). When we can see languagelike grammar as just a part of human communication, then the smoke will begin to clear.

And the mirrors of intuitive self-study will be changed to plain glass simply by changing the focus of attention. If human communication is really about communication then it is not the sender's mind that the sender has to model, it is the receiver's mind. Studying the area of human communication we call linguistics cannot be effectively achieved just by looking at how signals are produced, we also have to look at how they are received. There is still reflection, in all senses of the word, but it is reflection of – and off – an external surface: the receiver.

Currently our understanding of language is a variant of the *Blind Men and the Elephant* story: we all cling to our own definitions, and act as if our understanding of one aspect explains the whole elephant; but, even if we add all our definitions together, we still will not describe the animal. Language is a list of symptoms of a subset of human communication which we have, somewhat arbitrarily, reified as a Thing; but “thingness” is not a property of the list of symptoms, or even of the aspects that generated the symptoms, it is a property of the object that has those aspects – if there is actually such a thing. We need to address the whole of the elephant in the room.

References

- Michael J. Beran, J. David Smith & Bonnie M. Perdue (2013).** Language-Trained Chimpanzees (Pan troglodytes) Name What They Have Seen but Look First at What They Have Not Seen. In *Psychological Science* 24:5, pp660-666.
- Lewis Carroll (1871).** *Through the Looking-Glass, and What Alice Found There*. Macmillan: London, UK.
- Noam Chomsky (1957).** *Syntactic Structures*. Mouton de Gruyter: Berlin, Germany.
- Noam Chomsky (1965).** *Aspects of the Theory of Syntax*. MIT Press: Cambridge, Mass, USA.
- Noam Chomsky (2009).** *Cartesian Linguistics: a chapter in the history of rationalist thought* (third edition). Cambridge University Press: Cambridge, UK.
- Noam Chomsky (2013).** Problems of Projection. In *Lingua* 130, pp33-49.
- Zanna Clay, Jahmaira Archbold & Klaus Zuberbühler (2015).** Functional flexibility in wild bonobo vocal behaviour. In *PeerJ*, 3:e1124.

- Stephen J. Cowley (2011).** Distributed Language. In Stephen J. Cowley (ed.), *Distributed Language*. John Benjamins B.V.: Amsterdam, Netherlands.
- David Crystal (2005).** *How Language Works: How babies babble, words change meaning, and languages live or die*. Penguin Books: London, UK.
- René Descartes (1649 [1998]).** Descartes to Henry More, 5 February 1649. In *Meditations and Other Metaphysical Writings*. Penguin: London, UK, pp173-175.
- Roland H. Günther, Caitlin E. O’Connell Rodwell, Simon L. Klemperer (2004).** Seismic waves from elephant vocalizations: a possible communication mode? In *Geophysical Research Letters* 31, L11602, pp1-4.
- Marc D. Hauser, Noam Chomsky & W. Tecumseh Fitch (2002).** The Faculty of Language: what is it, who has it, and how did it evolve? In *Science* 298, pp1569-1579.
- Suzana Herculano-Houzel (2012).** The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. In *PNAS* 109 suppl1, pp10661-10668.
- Charles F. Hockett (1960).** The origin of speech. In *Scientific American* 203, pp4-12.
- Wilhelm von Humboldt (1836 [1971]).** *Linguistic Variability and Intellectual Development*. University of Miami Press: Miami, FL, USA.
- Harry J. Jerison (1973).** *Evolution of the Brain and Intelligence*. Academic Press: New York, NY, USA.
- Juliane Kaminski, Josep Call & Julia Fischer (2004).** Word Learning in a Domestic Dog: Evidence for “Fast Mapping”. In *Science* 304, pp1682-1683.
- Irene Maxine Pepperberg (1999).** *The Alex Studies: Cognitive and communicative abilities of grey parrots*. Harvard University Press: Cambridge, Mass, USA.
- Steven Pinker & Ray Jackendoff (2005).** The faculty of language: what’s special about it? In *Cognition* 95, pp201-236.
- Willard V. O. Quine (1960).** *Word and Object*. MIT Press: Cambridge, MA, USA.
- Claudia Rudolf von Rohr, Carel P. van Schaik, Alexandra Kissling & Judith M. Burkart (2015).** Chimpanzees’ Bystander Reactions to Infanticide: An Evolutionary Precursor of Social Norms? In *Human Nature* 26, pp143-160.
- Pär Segerdahl, William Fields & Sue Savage-Rumbaugh (2005).** *Kanzi’s Primal Language: The cultural initiation of primates into language*. Palgrave Macmillan: Basingstoke, Hants, UK.
- Sara J. Shettleworth (2012).** Do Animals Have Insight, and What Is Insight Anyway? In *Canadian Journal of Experimental Psychology* 66:4, pp217-226.
- Maggie Tallerman (2014).** No syntax saltation in language evolution. In *Language Sciences* 46, pp207-219.
- Michael Toolan (1996).** *Total Speech: An integrational linguistic approach to language*. Duke University Press: Durham, NC, USA.

3 Challenging government policy on English language teaching in Japan through collaborative action research: theory and method

Tamiko Kondo
York St John University, UK

Introduction

The Japanese Ministry of Education, Culture, Sports, Science and Technology (known as MEXT) has repeated its announcement of plans to reform English language education in recent years. These movements show the government's desire for an urgent improvement of Japanese students' English ability which puts considerable pressure on many Japanese English teachers. In preparation for the 2020 Tokyo Olympic Games, in 2013, MEXT announced the latest plan (2013), identifying specific levels or scores in external English proficiency tests both as the goal of Japanese secondary level students and as the minimum requirement of English proficiency for Japanese English teachers. MEXT is also considering the idea of introducing the TOEFL test into university entrance examination.

Amongst Japanese English teachers themselves, there has been some debate about the credibility of this decision. Some of them have questioned whether scores in standardised tests can assess communicative competence. Moreover, in 2014, MEXT convened a number of Leaders of English Education Projects (known as LEEPs), one from each of the Japanese municipalities, and began a series of training courses for the LEEPs with the support of the British Council. Japanese English teachers are watching anxiously how MEXT and the British Council may require them to change their practice, and some tension has arisen as the result.

From the perspective of critical applied linguistics, MEXT's policy apparently supports 'native-speakerism' (Holliday, 2005, p.6) and assumptions about 'dominant varieties' (British/American) of English (Canagarajah, 2006, p.229). This instrumental approach to education, its

focus on outcomes and its ‘monolithic’ conceptualization of English goes against my own values. These include a commitment towards authentic communication through the development of communicative competence and a ‘plurilithic’ approach to the conceptualisation of English (Hall, 2013, p.211). In order to challenge the national notion of native-speakerism, I have convened an action research group of Japanese Junior High School English language teachers, who are the research participants for this study. We have tried to find ways of facilitating our students’ communicative competence, and developing ourselves and our practice. Ultimately, we aim to make a positive contribution to the development of English language education and English teacher education in Japan. Using a workshop for Japanese English teachers (Wicaksono and Kondo, 2014), I have also tried to raise consciousness of, and develop sensitivity to, changing Englishes and their professional development according to their local needs and contexts.

In the spirit of ‘bottom-up applied linguistics’ which challenges the existing ‘top-down transmission model’ of teaching and learning in applied linguistic studies (Hall et al., 2011, pp. 19-20), this project takes a collaborative action research approach to generating new knowledge. Action research is described by McNiff and Whitehead (2011, pp.7-10) as a cyclical process of learning that includes: ‘observe - reflect - act - evaluate - modify - and move in new directions’. The purpose of action research is not only to ‘improve our practice’ but also to ‘generate new knowledge’ which ‘feeds into new theory’ (ibid., p.14). Importantly, action researchers are ‘insider researchers’ who ‘see themselves as part of the context they are investigating’ (ibid., p.8). This inevitably requires them to work with others throughout the research process, leading the process of ‘knowledge creation’ to ‘a collaborative process’ (ibid., p.32). On the basis of this approach to practice-based action research, I have worked with the research participants through facilitating their action-reflection cycles, in order to influence the political context which condones native-speakerism, by putting new knowledge about English language education and English teacher education into the public domain.

In this context, this paper aims to answer these questions: What is the impact/ are the gaps in English teacher education programs conducted by an ‘external professional expert’? What does communicative competence mean? How could a plurilithic understanding of English have impact on the English language education context of Japan? The findings from the following qualitative data analysed by means of content analysis are provided and discussed: the questionnaire responses from: 10 LEEPs who participated in the second British Council-led teacher education program in October 2014;

15 teachers who participated in a LEEP-led teacher education program in August 2015 and the feedback from 15 workshop (Wicaksono and Kondo, 2014) participants.

What is the impact/ are the gaps in English teacher education programs conducted by an 'external professional expert'?

In 2014, MEXT commenced a new 'cascade' design project for English teachers' professional development 'in cooperation with the external professional expert', the British Council (MEXT, 2014a). The aim of this new project is 'promoting the training of English teachers with communicative competence in English appropriate to globalization' (2014a). Actually, this project started in response to MEXT (2013) which states that English should be taught 'basically in English' at junior high school level. It is MEXT's view that monolingual instructional strategy would 'increase students' English experience', leading to the development of their communicative competence (2014b). Under this project, LEEPs have attended two kinds of education programs conducted by the British Council, the first one focusing on their learning of practical teaching strategies, the second one focusing on how to instruct other teachers (2014a). Having been educated there, LEEPs are required to educate other English teachers in each municipality. The LEEP-led teacher education programs are mandatory programs which all English teachers have to attend 'within around the next five years' (2014a). In order to attain this, in 2015, new LEEPs were selected in each municipality to be sent to the governmental (the British Council-led) education programs which are going to conduct the LEEP-led teacher education programs, in the same way as the first LEEPs.

The LEEPs who participated in the British Council-led second education program in October 2014, and the teachers who participated in a LEEP-led education program in August 2015, share discourse in their description of the impact and the gaps in those programs as part of the MEXT's cascade project. While most of them are positive about the content of the program, they identify the gaps between the content of the program and their expectations/needs in their real-life teaching context. The analyses of the LEEPs' and the teachers' written reflections are presented in Table 1 and Table 2.

Analyses (Frequency)	The LEEPs' written responses
<u>Positive impact of the program</u> / All are very positive about the program	"It was a very good stimulus to learn the totally different way of teaching from mine." "If I had only to attend the program, that was great, helpful to my future practice"
<u>Benefit of the program</u> / Networking and interacting with other LEEPs which took place 'outside the program', was mentioned most (5)	"Meeting with the highly respectable LEEPs as individuals and professionals was the greatest benefit from the program." "Interacting with the LEEPs from other municipalities was specifically new, sharing the same experience and discussing with them was invaluable."
<u>Gaps identified 1</u> / Relevance between the program and their future practice (2)	"Why do we attend these education programs? How will these lead to the future English language education? The programs would have become more effective if we had made them clear before the start of this project."
<u>Gaps identified 2</u> / Relevance between teaching strategies they learnt at the program and their real-life teaching context (3), ^(a) in terms of students' level, ^(b) in terms of the idea of following the same way with the British Council, ^(c) in terms of LEEPs' time commitment to get ready as a teacher educator	^(a) "The content should be more conscious of how we can apply it to the first year students in junior high school. Both the content and the materials should be more conscious of students in the real-life situations." ^(b) "Realistically, it is doubtful whether I will be able to practice a hundred percent what I learnt at the program at school." ^(c) "Considering what I am required to do as a teacher educator, I need much more time."

Table 1: Observations on the questionnaire responses from 10 LEEPs who participated in the British Council-led second education program in October 2014

Analyses (Frequency)	The teachers' written responses
<u>Positive impact of the program</u> / Their general opinion about the program is mostly positive: ^(a) good (11), ^(b) maybe good (2), ^(c) needs reconsideration (1), no answer (1)	^(a) "This training course is really good for me to improve my teaching skills." ^(b) "It is good, if we do not have to think about English for entrance exams." ^(b) "It is good to develop students' practical English skills, but we should be trained more if we use the method." ^(c) "We should discuss monolingual instructional strategy."
<u>Benefit of the program</u> / ^(a) Certain sessions or features of the program were mentioned most (8), ^(b) got new ideas (4), discussion with other participants (1), knowledge update (1), no answer (1)	^(a) "It was beneficial that we learnt how to increase authentic English use in class." ^(a) "Writing session and vocabulary session were so useful." ^(b) "I could get many activities that I can use in lessons!"

Analyses (Frequency)	The teachers' written responses
Gaps identified 1 / Relevance between teaching strategies they learnt at the program and their real-life teaching context (10), in terms of ^(a) level, ^(b) time, ^(c) relevance with the Course of Study, ^(d) language use	^(a) “It is hard for me to adopt the method we learnt with my students.” ^(b) “It takes us more time to teach with the method. Progress with classes would be behind.” ^(c) “The idea of the program isn’t relevant to the Course of Study or the MEXT-authorized textbook.” ^(d) “Japanese explanation should be necessary depending on the content.”
Gaps identified 2 / Gaps between the program and their expectation/needs (2)	“We should have discussed how we can give MEXT-authorized textbook-based lessons only in English.” “I wanted to learn theory of second language acquisition.”

Table 2. Observations on the questionnaire responses from 15 teachers who participated in a LEEP-led education program in August 2015

The following points should be noted:

1. None of the LEEPs and the teachers mentioned learning and improving monolingual instructional strategy, which is the focus of MEXT (2013), as the benefit of the program. This shows another gap between MEXT and the participants. This also implies that it is not realistic for the participants to consider the development of students’ communicative competence in connection with monolingual instructional strategy. Actually, monolingual instructional strategy is questioned by eight LEEPs in the same questionnaire.
2. The gaps between the content of the programs and the participants’ real-life teaching context imply the challenge of teacher education programs conducted by an ‘external professional expert’. Specifically, the relevance between the program and their future practice, which was identified by two LEEPs, suggests the significance of the cycles of (re)evaluation of teacher education programs in relation to the changes in the real-life teaching context.

What has given rise to the current situation described above are as follows: the gap between MEXT’s and teachers’ thinking on communicative competence, and the lack of the consistency of MEXT’s viewpoint for English teachers’ professional development. In other words, discussing what communicative competence means has been left out, as was discussing what the requirements for Japanese English teachers should be, in this postmodern, globalised era in which a variety of ‘Englishes’ are spoken all

over the world. It is the primary purpose of this study to spotlight these issues and stir up discussions among those involved.

What does communicative competence mean?

The nature of communicative competence and its implications for language teaching has been much debated, over several decades, for example by Hymes (1972), Canale and Swain (1980), Canale (1983) and Savignon (1976, 1997 and 2002). On theorising my theoretical framework of communicative competence, I mainly look at Savignon's (2002) approach to communicative competence, in that Savignon's theory suits the present-day context where communication in English occurs a lot in 'nonnative-nonnative interactions' (Canagarajah, 2006, p.233). Savignon's (2002) theoretical framework includes four domains of communicative competence: grammatical competence, discourse competence, strategic competence and sociocultural competence. Savignon (2002, p.9) broadens the perspective of sociolinguistic competence of Canale and Swain (1980), and presents instead a new concept of sociocultural competence, emphasising 'an understanding of the social context in which language is used'. My preliminary research, in collaboration with the research participants, has suggested that teachers conceptualise communicative competence as 'a positive attitude towards communication', consisting of the following six factors: willingness, empathy, openness, creativity, originality and confidence. Having got this initial conceptual framework, I then looked at the relevant literature in an attempt to link the framework with various theories of communication, and finally produced the model of my conceptual framework of communicative competence (Figure 1). From our perspective, these six factors should be emphasised in English language education context in order to help students become communicatively competent in real-life social-discursive context.

In the responses to the above-mentioned questionnaire for LEEPs, eight out of ten LEEPs hold similar views to us, mainly emphasising the following two points: a positive attitude towards communication (including oneself, the other, the relationship with the other), and a capacity to have, give and share our own opinions. These factors seem likely to be relevant to the notion of 'dispositions that favor translingual communication and literacy' (Canagarajah, 2013, p.5). Canagarajah (2014, p.91) also explains that those dispositions are 'developed in social contexts through everyday experience, as in *habitus*'. In his study on 'skilled migrants in English-dominant countries', Canagarajah highlights one migrant's disposition of being 'comfortable with using English in combination with local languages and even having influences of one's own values and identities' (ibid., p.95).

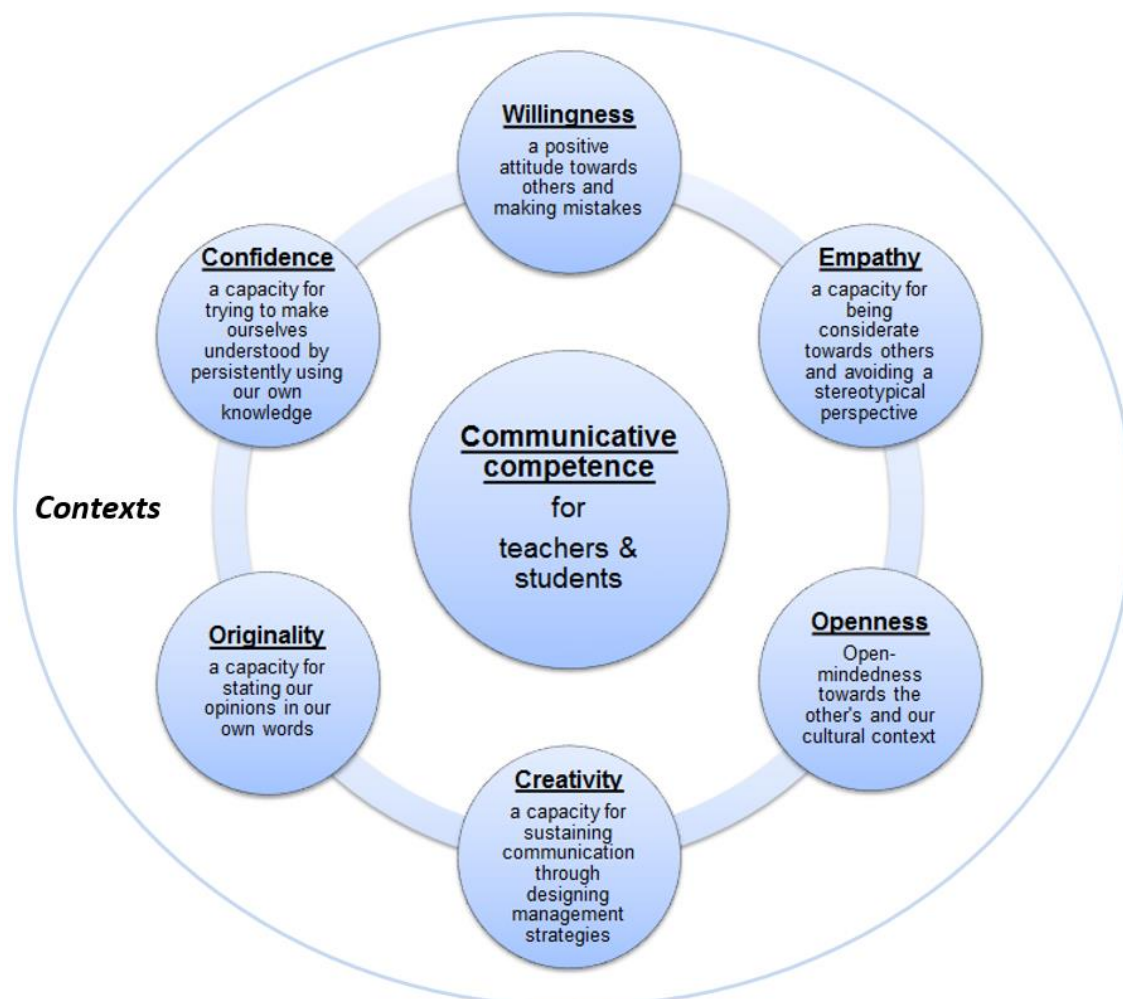


Figure 1: What helps us become communicatively competent? (Kondo, 2015)

Although this statement assumes a multilingual communication setting, I think that this idea can be emphasised in the same manner in the additional language education context in a monolingual society such as Japan. That is because this open-mindedness towards speaking English even with influences of our own language would help our students feel comfortable and become confident with their English. With this hope, our interpretations of willingness (a positive attitude towards others and making mistakes) and openness (open-mindedness towards the other's and our cultural context) include a positive attitude towards 'ourselves' and 'our culture'. This model also reveals a significant gap in the interpretations of communicative competence of teachers and MEXT, who have simply tried to equate communicative competence with scores in external English proficiency tests. By presenting this, I hope that this conceptual model will be of some help for other language teachers to theorise their own perspective towards communicative competence.

How could a plurilithic understanding of English have impact on the English language education context of Japan?

Monolithic conceptualization of English is consistent with a belief in dominant varieties of English (British/American English), tending to rely on well-known ‘standardised’ English tests (Canagarajah, 2006, pp.229-230). This monolithic conceptualisation of English has greatly influenced the discourse in English language education policy in Japan, showing apparent contradiction between its discourse and the principle of the Course of Study, where authentic communicative language skill is valued. A plurilithic conceptualisation of English, influenced by the thinking of critical applied linguists on World Englishes and English as a lingua franca, has challenged traditional monolithic ideas about language learning and teaching. In a plurilithic concept of English, the language consists of ‘multiple, coalescing objects’, has ‘fuzzy boundaries’ and ‘an ambiguous shape and form’, and is ‘variable, hybrid, and dynamic’ (Hall and Wicaksono, 2015). From this perspective, the adoption of ‘standardised’ English tests targeting English used in specific areas of the world (such as the UK/the USA) would narrow English diversity in the educational context, and does not necessarily help students communicate effectively and satisfactorily in a range of different contexts. On the basis of this plurilithic approach to English, we need to help both Japanese English teachers and students to identify themselves as ‘a speaker of their own English’, not ‘a non-native speaker’. What should be highlighted is Japanese English teachers’ ‘expertise’, not their non-nativeness (Rampton, 1990). Accordingly, we need to stop being misled by traditional notions and re-evaluate the goal of English language learning, by emphasising that we do not have to be ‘trying to become exactly like the other person’ (‘a native speaker’) (Gumperz, 1979, p.273).

On the other hand, it is a challenge to spotlight plurilithic conceptualisation of English in a monolingual context such as Japan, where people have very little opportunity to interact with people from other countries in English. As a first step, a workshop (Wicaksono and Kondo, 2014) was held with the aim of raising Japanese English teachers’ ‘awareness of the variable and dynamic nature of global English’ and reflecting on ‘implications for *their* professional practice, according to *their* local needs and contexts’. Part of the workshop was focused on the ontologies of English, comparing monolithic and plurilithic conceptualizations of English. The feedback from 15 participants show five kinds of impact which the workshop had on them, as shown in Table 3.

Impacts	The participants' written reflections
⁽¹⁾ Raised awareness of varieties of English and the contradiction in MEXT's policy	"I am getting to have a big question to the Japanese Government ideas for English language education they are planning now. Because its goal is depending on "Standard English" being tested by TOEFL."
⁽²⁾ Identified what is really the issue and spotlighted dilemma	"I agree that the concept of "Standard English" is useless because actually nobody can define what "Standard English" is. However, when teachers teach English to students, particularly beginners or low-level learners, we need to teach them what is correct so that they won't get confused."
⁽³⁾ Encouraged them to challenge the existing reality positively	"But we teachers are required to open mind toward the world and keep studying what English is." "We can be flexible with our ideas and teaching methods, and for this we need to have many workshops."
⁽⁴⁾ Encouraged them to think how they can make a difference in their context	"The people outside the English education context strongly believe there is one "Standard English". How we can change their belief is one of the big challenges."
⁽⁵⁾ Possibly laid the foundation of teacher community culture	"To make today's session more useful experience, we will reunite again in the near future and share our experience."

Table 3: Observations on the feedback from 15 participants in the workshop (Wicaksono and Kondo, 2014)

These reflections communicate that the workshop gave them the opportunity to think for themselves how 'they themselves are influenced by' the context, beyond thinking about the ontologies of English (Sheppard 1998, cited in Edy 2000, p.57). At the same time, it is implied that there should be more opportunities for Japanese English teachers to think about how English language should be understood in this postmodern, globalised era. Those opportunities could help teachers become aware of the equality of all varieties of English, which would be communicated to their students. With this hope, plurilithic conceptualisation of English should be focused and discussed in the English teacher education context. At the outset, we should shift our perspective on Japanese English teachers, from 'teachers of English as a foreign language' to 'teachers of English as a lingua franca', because 'the idea of English as a *foreign* language' belongs to 'native speakers only' (Hall and Wicaksono, 2015).

Implications

This account of my action research could suggest the way of incorporating a critical applied linguistic approach into action research, and possible classroom-based responses to ‘plurilithic conceptualisation of English’ in the monolingual context. While MEXT have repeatedly referred to ‘globalization’ (‘gurōbaruka’ in Japanese) in policy (MEXT 2003, 2011 and 2013), they still value the dominant varieties of Englishes, as I discussed above. Nowadays, ‘much of the communication in English happens among multilingual speakers in nonnative-nonnative interactions’ (Canagarajah, 2006, p.233). My point is that true globalisation means being able to value all languages and speakers (all Englishes and English speakers) in this globalised era, not being able to speak the English used in the UK or the USA. It is this ability that gives Japanese students ‘global citizenship’ which MEXT desires them to have. The present situation shows the limitations of MEXT’s (2013) rushed decision to push forward a monolingual instructional strategy in junior high schools for the development of students’ communicative competence. A monolingual instructional strategy could be a means for developing students’ communicative competence, however, it should not be the goal of English language education or teacher education programs. Likewise, ‘standardised’ English tests should not be the goal of our students or the requirement for English teachers. What could be emphasised is to help Japanese English teachers and Japanese students recognize and appreciate the many varieties of Englishes, and ‘the ability to shuttle between different varieties of English and different speech communities’ (p.233). This would contribute to Japanese English teachers’ professional development and Japanese students’ life-long English language learning. I hope that this paper could trigger discussions among those concerned.

References

- Suresh Canagarajah (2006).** Changing communicative needs, revised assessment objectives: testing English as an international language. In *Language Assessment Quarterly* vol 3 3, pp229-242.
- Suresh A. Canagarajah (2013).** Introduction. In Canagarajah A. Suresh (ed), *Literacy as translingual practice: between communities and classrooms*. Routledge: Oxon, UK, pp1-10.
- Suresh A. Canagarajah (2014).** Theorizing a competence for translingual practice at the contact zone. In May Stephen (ed), *The multilingual turn: implications for SLA, TESOL, and bilingual education*. Routledge: New York, USA, pp78-102.
- Michael Canale (1983).** From communicative competence to communicative language pedagogy. In Richards Jack C. & Schmidt Richard W. (eds), *Language and communication*. Longman: London, UK, pp2-27.

- Michael Canale & Merrill Swain (1980).** Theoretical bases of communicative approaches to second language teaching and testing. In *Applied Linguistics*, vol 1 1, pp1-47.
- Maureen Edy (2000).** Understanding professional development. In Brechin Ann, Brown Hilary & Edy Maureen A. (eds), *Critical practice in health and social care*. The Open University: London, UK, pp48-69.
- John Gumperz (1979).** Cross-cultural communication. In Harris Roxy and Rampton Ben (eds), *The language, ethnicity and race reader*. 2003. Routledge: London, UK, pp267-275.
- Christopher J. Hall (2013).** Cognitive contributions to plurilithic views of English and other languages. In *Applied Linguistics*, vol 34 2, pp211-231.
- Christopher J. Hall & Rachel Wicaksono (2015).** *Changing Englishes: an interactive course for teachers*. [Online]. [Accessed 25 March 2015]. Available from: <http://www.yorks.ac.uk/changing-englishes/changing-englishes.aspx>
- Adrian Holliday (2005).** *The struggle to teach English as an international language*. Oxford University Press: Oxford, UK.
- Dell Hathaway Hymes (1972).** On communicative competence. In: Pride J. B. and Holmes Janet (eds), *Sociolinguistics*. Harmondsworth, Penguin Books: Middlesex, UK, pp269-293.
- Ben Rampton (1990).** Displacing the 'native speaker': expertise, affiliation and inheritance. In: Harris Roxy and Rampton Ben eds. *The language, ethnicity and race reader*. 2003. London: Routledge, pp107-111.
- Sandra J. Savignon (1976).** *Communicative competence: theory and classroom practice*. Central States Conference on the Teaching of Foreign Languages, 23 April, Detroit, pp2-23.
- Sandra J. Savignon (1997).** *Communicative competence: theory and classroom practice texts and contexts in second language learning*. 2nd ed. McGraw-Hill: New York, USA.
- Sandra J. Savignon (2002).** Communicative language teaching: linguistic theory and classroom practice. In: Savignon Sandra J. (ed), *Interpreting communicative language teaching: contexts and concerns in teacher education*. Yale University Press: London, pp1-27.
- The Japanese Ministry of Education, Culture, Sports, Science and Technology (2003).** *Eigo ga tsukaeru nihonjin no ikusei notameno kōdōkeikaku*. [Online]. Japan: The Japanese Ministry of Education, Culture, Sports, Science and Technology. [Accessed 13 August 2014]. Available from: <http://www.mext.go.jp/>
- The Japanese Ministry of Education, Culture, Sports, Science and Technology (2011).** *Five proposals and specific measures for developing proficiency in English for international communication*. [Online]. Japan: The Japanese Ministry of Education, Culture, Sports, Science and Technology. [Accessed 17 June 2015]. Available from: <http://www.mext.go.jp/>
- The Japanese Ministry of Education, Culture, Sports, Science and Technology (2013).** *Gurōbaruka ni taiōsihita eigokyōikukaikaku jisshikeikaku*. [Online]. Japan: The Japanese Ministry of Education, Culture, Sports, Science and Technology. [Accessed 13 August 2014]. Available from: <http://www.mext.go.jp/>
- The Japanese Ministry of Education, Culture, Sports, Science and Technology (2014a).** *Heisei 26 nendo eigokyōikusuishinrīdā chūōkensyū jisshiyōkō*. [Online].

- Japan: The Japanese Ministry of Education, Culture, Sports, Science and Technology. [Accessed 10 August 2015]. Available from: <http://www.mext.go.jp/>
- The Japanese Ministry of Education, Culture, Sports, Science and Technology (2014b).** *Kongono eigokyōiku no kaizen-jūjitsuhōsaku nitsuite hōkoku: gurōbaruka ni taiōshita eigokyōikukaikakuno itsutsuno teigen*. [Online]. Japan: The Japanese Ministry of Education, Culture, Sports, Science and Technology. [Accessed 17 June 2015]. Available from: <http://www.mext.go.jp/>
- Rachel Wicaksono and Tamiko Kondo (2014).** *Are we teachers of English or English as a lingua franca? Breaking away from 'non-native English teachers' myth and heading for Japanese students' and our communicative competence*. [Letter of invitation, 20 December].

4 New directions created by Putonghua and English's "double domination" of Zhuang language in China

Alexandra Grey
Macquarie University, Australia
alexandra.grey@students.mq.edu.au

Abstract

This paper examines new directions in the mobilization of linguistic resources for identity and inclusion in circumstances where Standard Mandarin (Putonghua) and English dominate a Chinese minority language. While existing literature examines tensions between Putonghua and English, it rarely considers China's official minority languages: another tension exists between a "peripheral language" and the combined practices of "supercentral" Putonghua and "hypercentral" English (de Swaan's categories (2001)). Drawing on a 2014-2015 ethnographic study of Zhuang, the Kam-Tai language of the largest official minority group in China, this paper argues that Putonghua and English dominate in a relatively stable coalition, and that, contrary to some other minority language cases in China, the imbalance of power vis-à-vis Zhuang results in its destabilisation even as a resource of home-making and ethnic identity, the traditional power bases of minority language.

I identify complementarities in the relationship between Putonghua and English enabled by capital in different fields and value constructed on national and international "scales" (Blommaert, 2007). I explain two language ideologies that stabilise this double domination. Putonghua and English, but not Zhuang, are constructed as features of a sought-after "*habitus* of mobility" (Grey, 2015), a *habitus* of the kind Blommaert (2007, p. 11) predicts will "prevail over others because [it] connect[s] to higher and more powerful scale-levels". Consequently, not only are people shifting away from learning or using Zhuang but, through discourse, reconstructing Zhuang identity as authentic even without Zhuang language practices.

Background and Literature

This paper holds that languages exist as discrete entities in China in the sense of "a language" being a common, meaningful social construction. One such is "minority language" (*shaoshu minzu yuyan*), referring to 54 language

varieties recognised as belonging to official minority ethnolinguistic groupings (called *minzu*). All citizens are classified into a minority *minzu* (together 8.49% of the population (Zhou, 2013, p. 1)), or the majority Han *minzu*. This results in widespread official and popular discourses within China about the nation as a *minzu tuanjie/tongyi duominzu guojia*, conventionally expressed in English as “unitary multi-ethnic country” (China Encyclopedia Compilation Group, 2008, p. 38) or “unified country made up of diverse nationalities” (Harrell, 1993, p. 97). Officially, minority languages are distinguished from the many Mandarin dialects (*fangyan*) in China.

There is substantial literature on English in China and its relationship with Putonghua. English is generally observed to index international mobility and prosperity (Blommaert, 2007, p. 13), and these are valued in China. However, globalization, and therefore English language features which index it, also causes tensions in China. As Gao Shuang (2015, p. 7) observes,

[g]iven the gatekeeping role of English in both educational and professional advancement, English has become an important linguistic capital that people compete for. Over emphasis on English, however, has always been a hot topic of debate.

Zhang (2011, p. vi) explains a common conceptualization in China’s English education sector of English and Putonghua harmoniously occupying different functions: “English for *yong* – modern uses and international communication; Chinese for *ti* – national cohesion and harmony”, but notes this is “a simplistic view of complementary language use”. Zhang (2011, p. vii), and the prominent academic Gao Yihong (2009, p. 58), suggest the complex, overlapping uses and values of the two languages create an “identity dilemma” for Chinese learners of English. Such cases are “up-scaled” (Blommaert, 2007, p. 6) to a national “identity crisis” in the literature (e.g. Zhang, 2011, p. 10; Zheng, 2009), and, likewise in government discourses.

While Putonghua has been a tool of national identity construction since the PRC’s early days, it has become a government priority to mobilise Putonghua as a “rallying point” (Heller, Pujolar, & Duchêne, 2014, p. 542); (Coupland, 2001, p. 18); (Fishman, 1999, p. 364) for national cohesion and Chinese identity in reaction to globalization and English. For example, in 2001, alongside its WTO entry, China enacted a law obliging Putonghua in many domains (National People’s Congress, 2000), and plans to remove the obligatory English university entrance examination from 2017, while soft

power initiatives to teach Putonghua are now global and well-funded. Nevertheless, there are also government-propelled discourses constructing English and Putonghua as harmonious not competitive. For example, Zhang, (2011, p. 18), in her work on English and the 2008 Beijing Olympics, notes "Beijing citizens were encouraged to express their patriotism and glorify the whole nation by learning English."

However, the literature examining Puonghua-English tensions seldom considers how the spread of English affects power between China's majority and minority languages or how the "identity dilemma" differs for Chinese who speak minority languages: the "dilemma" is presented in terms of people with a mainstream (Han) identity (e.g. Y. Gao, 2009). Similarly, Gao Shuang's (2015) otherwise excellent analysis of the changing indexicality of multilingualism within Chinese "neo-nationalism" (p.3) overlooks the indexicality of minority multilingualism in China today and historically, when today's official minority languages were, like English, considered languages of foreign "barbarians" (p.4) (on the foreignness of China's minorities, see also Tapp & Cohn, 2003). Gao Shuang's data – texts from Chinese media celebrating the foreign language multilingualism of "ordinary but great citizen[s]" (p.12) – illustrate how minority multilingualism does not count as multilingualism at all in influential discourses. Thus, questions as yet unanswered in the literature include whether English in China is in tension with minority languages or destabilises minority ethnolinguistic identities.

These questions cannot be answered generically: China is a vast linguistic environment and the histories of contact between minority languages and state-sanctioned standard Mandarin varieties (of which Putonghua is the latest) vary widely. Moreover, while the circulation of resources for learning and legitimating English is unequal in China in ways that mirror general minority inequality in socioeconomic resource distribution, there are significant socioeconomic differences among China's various minority *minzu*. Therefore, this paper studies one specific minority language, Zhuang. The Zhuang *minzu* numbers around 16 million people (Luo, 2008, p. 317). Zhuang language has a standardised dialect and local dialects. The majority of Zhuang speakers today are from South-Central China, in particular Guangxi Zhuang Autonomous Region, a territory with provincial-level government status.

While the spread of English in Zhuang-speaking areas is not widely discussed in the literature, research exists on the impact of English on other

minority languages practices in China. There is some suggestion that English provides identity resources for constructing non-Han identities: regarding young Uyghurs, Finley and Zeng (2015, p. 22) suggest English learning is driven by a “wish to transcend PRC state discourses and national boundaries and become citizens of the world”. However, such studies are particular to groups whose language use is semiotised in China within broader inter-ethnic political and physical conflicts. Zhuang-Han relations are typically unlike the unrest of Tibetan and Uyghur regions, and so power relations between English, Putonghua and Zhuang differ from relations with those minority languages. Closer to Zhuang’s situation are minority languages in Yunnan Province, bordering Guangxi. The impact of English on Yunnan’s minority languages has been studied, particularly within schooling (e.g. Ru Blanchford & Jones, 2011; Yuan et al., 2015). Yang’s (2012, pp. i-ii) study of English, Putonghua and Naxi minority language in Lijiang, a Yunnan town, finds:

The prevalent belief in Lijiang is that Chinese and English constitute a form of linguistic capital which enables individual, socio-economic mobility whereas Naxi is mostly seen as of symbolic value, tied to Naxi ethnic heritage and identity, and thus of limited socio-economic value ... Naxi is still the main language of the home and community and a marker of their ethnic identity.

That is, from the perspective of Naxi, English and Putonghua are stable between themselves and dominating: I hypothesise the same for Zhuang. Yang’s finding of an enduring, though symbolic, value for Naxi suggests a certain stability in the patterning of Lijiang’s societal multilingualism, despite a power imbalance. I hypothesise that Zhuang is likewise relatively stable as a resource for ethnic identity construction and binding to home communities. High status did not mean high usage of English in Yang’s (2012, p. ii) study; suggesting internal stability between Putonghua and English can arise through having relatively few competing domains. I therefore have examined practices and language ideologies in my data that construct Putonghua and English as mutually supportive rather than competing.

Method

To investigate these hypotheses, I applied a Bourdieuan (1991) lens, as extended by Heller’s (2010) theory of language as linguistic resources circulating within a political economy. I interviewed 43 Zhuang *minzu*/Zhuang-speaking university students, and a secondary cohort of 18 minority language professionals and activists, in four cities across China. I selected locations to compare experiences within and outside of traditionally

Zhuang-speaking areas. Zhuang urbanisation is rising, hitting 34% at the last count (in 2005: Guo, 2013, p. 5), although just 10% in 1999: (Stites, 1999, p. 112). China's overall urbanisation surpassed 50% in 2012 (Bloomberg News, 2012) (recent Zhuang-specific rates are not available). Younger generations are socialised into this urbanisation and the marketizing, industrializing and globalizing of China. China's societies now place a high value on mobility across internal regional, rural-urban and class borders, often facilitated through the education infrastructure which structures access to urban residency rights, employment, and urbane lifestyle. Minority and rural students receive university entrance concessions under a national affirmative action policy (see Kaup (2000, p. 137)), but as individual universities partly control entry, minority enrolments at well-regarded universities remain disproportionately low. My student participants are comparatively highly educated. University students embody a *habitus* of mobility, hence their selection.

Through education, the students started personal journeys of migration, living in on-campus dormitories and visiting home only once or twice annually. Given this in-between lifestyle, my analysis incorporates Levitt and Glick Schiller's (2004, p. 1003) notion of simultaneous identity construction by migrants with "multi-sited" lives: they theorise "ways of belonging" (p. 1010) which are self-aware actions undertaken to maintain both Home and Host identities. I considered language practices as likely resources for belonging in the multi-sited lives of young minority people. This shaped the semi-structured interviews around language in adaptation, study, social networks and recreation at university, contacting and visiting home, and plans after graduation. Bourdieu's notion of "*habitus*", a disposition formed in particular through youth experiences which develop a subject's expectations about the "distinction" attainable through certain language practices (1977, p. 660), informed my treatment of participants' childhood home and school experiences as formative language market interactions, so home and school languages were also interview topics.

We discussed these topics in semi-structured "walk and talk" (Qian, Qian, & Zhu, 2012) interviews through the students' everyday linguistic landscapes, literally applying Kvale's (2009, p. 5) "wandering together" metaphor in order to better situate language practices amongst environmental texts and norms. (If students declined to walk because of outdoor temperatures, we met indoors on campus and they talked me through a virtual tour.) I also spent informal time in active observation of participants and their peers.

Findings and Discussion

Modifying Multilingual Repertoires

The data show these young Zhuang speakers and Zhuang group members live within busy linguistic ecologies of Zhuang, Standard and Dialectal Mandarins, other minority languages, and Englishes, but that a strong hierarchy exists with Zhuang subordinated. The students' repertoires are summarised below.

All 43 were fluent and literate in Putonghua, as was evident in our interviews and correspondence, and had obtained university entry in Putonghua-medium examinations, and were undertaking Putonghua-medium degrees. Three students reported speaking only Putonghua but as two of those were English Majors I counted them as partial English speakers. All others reported being bi- or multi-lingual. Thirty-eight students self-reported as Zhuang speakers, though one self-reported as a partial speaker. The self-reported monolingual Putonghua speaker additionally self-reported receptive competency in Zhuang. Zhuang speakers were mainly, but not exclusively, Zhuang *minzu* members. None reported speaking Standard Zhuang, although they knew of it. All Zhuang speakers used Zhuang rarely, for calling and visiting their families. For some, in-family communications had transitioned from Zhuang to Putonghua. Most did not use Zhuang with Zhuang-speaking classmates. The students did not know of Zhuang pop music or watch Zhuang TV, though some recalled TV in Zhuang previously existed and some knew of Zhuang folk songs from older relations. (Few could sing them.) They reported forgetting words in Zhuang after childhood and the *hanhua* ("Mandarization") of nouns, grammar and prosody when speaking Zhuang. Some non-Zhuang-speaking students were raised by Zhuang-speaking parents who intentionally limited their children's learning of Zhuang. Many students commented that this is more common amongst younger generations; a trend corroborated in interviews with secondary participants.

Although all 43 grew up in autonomous Zhuang areas, including Guangxi, where bilingual Zhuang-Putonghua primary education is permitted, few had Zhuang-Putonghua primary schooling. None had bilingual middle/high schooling. None knew anybody who learnt Zhuang other than informally in Zhuang-speaking home communities. Of the 38 Zhuang-speakers, only three reported having personal Zhuang literacy practices. (Rudimentary Zhuang literacy was being acquired at university by 14 in minority language and culture majors but they did not have Zhuang literacy practices out of class). For the literate students, having parents literate in Zhuang was a significant

resource in their own acquisition of Zhuang literacy. Very few participants wanted jobs for which Zhuang skills would be crucial or prioritised moving to Zhuang-speaking areas after graduation.

Regional Mandarin dialects were common in the repertoires, acquired from parents, playmates and communities in childhood. None had been the medium of instruction at school but had been playground varieties. By university, Putonghua was used more in everyday life than Mandarin dialects, and exclusively for studies. Some students used dialects in preference to Zhuang with siblings and old friends.

Four undergraduates were taking English university majors. While all other undergraduates (35 students) had compulsory English minors, and everyone had taken English throughout school, only two undergraduates and one PhD student reported English as a language they spoke. Others were nevertheless interested in English: students who did not report speaking English were keen to arrange English conversation practice sessions with me during fieldwork. Many enjoyed English-medium entertainment media. English was frequently visible in the linguistic landscape, though Putonghua was visually dominant.

Most of the 18 Minority Language and Culture Majors chose that major in order to attend university; they considered it easier to get into this major because of an oral Zhuang entry examination. These students were not taught in Zhuang or expected to become literate. Students attending university in Inner Mongolia (where there are no Zhuang courses or entry examinations) chose their university because in the entry process was less competitive and because they wanted to move to new, urban, places. The students at university in Beijing had attended the specialist Putonghua-medium high school in Beijing whose function is to facilitate university entry for out-of-town minority students.

In summary, the repertoire data reveal Zhuang was subordinated to Putonghua, English and Mandarin dialects on a value hierarchy, and often used less than Putonghua and Mandarin dialects. Zhuang was mainly practiced orally and in the home domain, but fading even in that domain as they aged. Whether students could speak Zhuang, and the purposes they spoke it for, reflected the normative home and school language practices they had reported. Education, meeting people, city life, media, using digital technology, and work are associated mainly with Putonghua. English was an

everyday practise only for a small group but an aspiration for others. Written Zhuang was not expected to be personally practised or seen.

A Putonghua-English Alliance

As well as showing that Zhuang was subordinated to Putonghua, English and Mandarin dialects, the data show that English and Putonghua mutually reinforce one another in significant ways, with a more salient tension constructed between minority and majority languages. The opposition of Putonghua and English is just one construction; mobilizing multiple languages as identity, communicative and other resources, rather than constructing a Putonghua-English contestation also emerged. This may be understood as “cultural strategizing” (Baoud, 2015) (Silverberg, 2007, p. 33). Moreover, Putonghua and English often derive their power from “capital” (Bourdieu, 1991) in different fields. Even within one field, both can be highly valued, with the contrastingly low-value comparator being a minority language rather than each other. In this way, the majority languages’ relationship can be complementary, stable and powerful, rather than tense.

For example, while English has high social capital – being linked to class and spatial mobility – Putonghua remains very highly valued for its cultural capital. Both have high educational capital but they are very clearly kept from competing: students work towards one end-of-schooling exam in English and another in Putonghua. By contrast, Zhuang was positioned as inappropriate for education from an early age in my participants’ experiences. (NB This paper includes only my English translations of interviews, for brevity.)

Researcher: So, if you had the chance, would you study written or spoken Zhuang?

'Una': I wouldn't, no, speaking realistically, it's not necessary; you know, Zhuang-speaking people, basically they are people who are outside of educated circles, basically there aren't people who speak it or write it. If you study it you still, in my opinion, don't have much use for it. [9.7.2014]

Fay: the [Zhuang] language at home is limited to what we speak, [I] did not think much about it. I did not think: hey, that language can become a professional [way of] writing. [27.6.2014]

Fay reported coming to see Zhuang in a different light after commencing her Zhuang Major, but the data reveal even this degree is not taught in Zhuang or necessarily aimed at professional Zhuang literacy, so the expectation of a

low value for Zhuang language practices in the educational field, developed in school, is not necessarily challenged at university.

What Shohamy (2006) calls the “hidden agenda” in language policy – linguistic norms implicitly promoted by the state – and pervasive language ideologies are important in constructing the language hierarchy and values of linguistic resources. In addition to the Patriotic English ideology identified in the literature, my data revealed two ideologies constructing Putonghua and English’s mutual superiority over Zhuang. First, “standard language ideology”, which Piller (2015) defines as:

the belief that a particular variety – usually the variety that has its roots in the speech of the most powerful group in society, that is often based on the written language, that is highly homogeneous and that is acquired through long years of formal education – is aesthetically, morally, and intellectually superior to other ways of speaking the language.

Putonghua is government-endorsed standard Mandarin, the language associated with the Han majority, and was often referred to as *hanyu* (“Han language”) in the interviews. Putonghua has an alphabetic pinyin only for infants’ schooling, but is otherwise written in *hanzi* (“Han characters”). By contrast, Zhuang is now written only in an alphabetic script and its historical character script (explicated by (Holm, 2013)) is almost forgotten. This is powerful, given the historic and continuing symbolism of character script as an index of civility, beauty and “Chineseness”. In short, Putonghua derives enormous capital from standard language ideology.

Similarly, English, as taught in China is also positioned as a standard language, with an American accent and grammar constructed as normatively better than other varieties of English. Standard language ideology shapes a hegemonic expectation that every language has a more correct and higher value variant. English, in China, does not challenge the Standard Language paradigm. The consequence is the naturalization of language hierarchies: it is “natural” that all languages come in better, central and worse, peripheral forms. This bolsters the very ideology that disadvantages Zhuang and other minority languages. Because Putonghua and (standard) English are constructed as each orienting to a different centre, the competition between them is reduced; rather, to be mobile across a greater scale of boundaries, knowing both standards is constructed as valuable.

Moreover, Standard Language Ideology is transplanted onto Zhuang. There is a centre and a periphery amongst Zhuang variants. Standard Zhuang was

created on the dialect of Wuming, which, though just a town today, was a historic centre of education in Guangxi. The importance of Standard Zhuang over other Zhuang variants is made very clear in language policies and legislation which I have reviewed that support and fund teaching, publishing, broadcasting and literacy *only* of Standard Zhuang, with a bias towards script development rather than oral language maintenance. As the data reveals many Zhuang speakers claim not to be able to understand, and read or write Standard Zhuang, their Zhuang variants are relegated to the periphery within the formal structures and discourses of Zhuang language protection.

Patriotic English and Standard Language ideologies bind Putonghua and English together, but the ideologically-structured placement of this bloc at the top of the hierarchy also relies on a third ideology, one that rationalises learning/speaking/promoting “high” languages at the expense of Zhuang. I call this the Zero-Sum Language Ideology. Zero-Sum Language Ideology is the idea that a gain in Zhuang means a loss in Putonghua and/or in English. This zero-sum logic is built on an awareness not only of the comparative ease of naturalistic language learning in childhood – so time spent on Zhuang is a waste while time on Putonghua or English is getting ahead of the curve – but also of a social prejudice against those who have negative transfer from first-language Zhuang in spoken Putonghua. Participants call this the *jia zhuang wenti* (“pinched Zhuang problem”).

My data shows some students’ repertoires were shaped by deliberate choices by parents and grandparents not to use Zhuang at home to avoid costs to their Putonghua, and they saw this as normal amongst their urban peers. Kaup (2000, p. 177) mentions urban second-generation Zhuang not growing up learning Zhuang in Guangxi in the 1980-90s. My participants suggested it now also happens in rural Guangxi, as Una and Lucy describe:

Una: the generation under me, they really speak Putonghua from very young... the head of the household may have the idea of teaching them to speak Putonghua. [9.7.2014].

Lucy: There’s a particular situation now back home [Guangxi], 5-6 year olds don’t want to speak Zhuang and their families don’t want them to either, even if the household are all Zhuang speakers. [30.6.2014]

The data do not suggest a large-scale direct replacement of Zhuang with English, however, this does not mean English places no pressure on Zhuang. If we consider Bourdieu’s notion of distinction, and the constant inflationary pressure which a marketised system and ideologies induces, then a systematic but indirect relationship between English and Zhuang becomes

clearer. As Piller and Cho (Piller & Cho, 2013) argue of South Korea, market ideologies naturalise discourses of competition, and English – an index of globality – becomes a “key terrain where competition is played out” (p. 28). This is applicable also in China, with its own emergent “English Fever” and enormously profitable English learning industry. In this context, the standard of linguistic capital required to attain distinction and profit is constantly driven upwards. Those closer to the top must acquire ever “better” English, perhaps even a distinguishing second foreign language, and the pressure spreads down. Everyone needs “better” Putonghua to meet the rising bar. For those, like many of my participants, whose relatively low socioeconomic position has resulted in limited access to the resources for acquiring proficient, standard English, they must nevertheless attain ever-rising standards of Putonghua to compete for university places and jobs. This puts greater pressure on the home and the school to develop Zhuang children’s Putonghua earlier and therefore better. Fay, a student talking about children in her traditionally Zhuang-speaking home community, expresses this perception of English being the next level of distinction when the value of Putonghua becomes lowered by mass up-skilling, through *shenzhi* (“even”):

Fay: nowadays many little children cannot speak [it]. It’s because from pre-school onwards teachers all speak Putonghua, there are even some already starting to teach some English. [27.6.2014]

Similarly, Mae, whose parents deliberately spoke Putonghua with her in childhood despite one being a Zhuang speaker, intends to speak English to any future children because “English is a tool” [27.6.2014]. Whether or not she one day does so, this reflects the perception amongst some that English replaces Putonghua, which has already replaced Zhuang, in the competition for linguistic distinction.

This market pressure amplifies the Zero-Sum language ideology: to attain distinction with rising standards of Putonghua, the possibility of a bilingual child’s Zhuang transferring into their Putonghua and reducing its linguistic capital is too costly. More profitable, mobile resources are acquired if the child avoids any “pinched Zhuang problem” and invests maximal resources in Putonghua at the optimal time, childhood.

Rationalizing the New Directions

Above, I explained that adapting to linguistic double domination increasingly means raising non-Zhuang speaking children, and that Zhuang-speaking youths are losing willingness and proficiency in Zhuang as they

develop a *habitus* of mobility. "The failure to transmit vernaculars inter-generationally may be rationalised in various ways" argue Woolard and Scheffelin (1994, p. 62), "depending on how speakers conceptualise the links of language, cognition and social life": it is to these rationalities I now turn. Two discourses in my data construct it as increasingly socially "tolerable" (Grin, 1995) for a Zhuang ethnic/cultural/political identity not to be linguistically enacted. However, my data reveals the authenticity of a non-Zhuang speaking Zhuang person not yet acceptable to all.

The first of these discourses naturalises not speaking Zhuang within a perceived openness of the Zhuang towards majority culture. For example, Lucy, explaining her family's shift from using Zhuang together in her early childhood to using Putonghua and a Mandarin dialect, repeats this discourse from authority figures:

Lucy: The Zhuang...are an ethnic group who is relatively happy studying Han ways of speaking. Don't know why. Our teachers also say the same, '[we] don't know why the Zhuang are relatively keen and easily receive Han culture'...They're relatively happy receiving Putonghua. [30.6.2014]

The second rationalization is reflexive, as it involves explicitly weighing up preserving Zhuang language or preserving Zhuang as a meaningful grouping. A clear example comes from Mr B, a trilingual Zhuang-Putonghua-English teacher participating in my secondary interviews. He said of non-Zhuang speaking youngsters, "[I'm] open minded to this group, I don't want to kick them out, [they have] other ties to Zhuang." [7.5.2015, English original.] The students also engaged in this discursive construction of culture as more important than language:

Fay: But I think [you] could not study Zhuang language; culture must absolutely but passed onwards.

Danielle: Yes, I think culture absolutely must be studied. [27.6.2014, *joint interview*]

In another interview, Mae likewise said:

So that is to say [Zhuang] language is also important, but the essence is still more about culture...I think culture is nevertheless the quintessence. [27.6.2014]

I asked Fay and Danielle which cultural characteristics they considered significant:

Fay: [You] cannot let them think 'oh, we Zhuang people and those *minzu* don't have any differences' ...I'd say there are some things you can, for instance through festival activities, [culture] can be reflected.

The oppositional positioning of language and group is not reflexively considered in this discourse, however, so it naturalises the separability of language and identity and the application of a cost/benefit evaluation to language practices.

In addition, official ethnic policy is an influential discourse: the state no longer classifies people into minorities by language, because "imaginable" (Mullaney, 2011, p. 16) ethnic communities are now administratively reproduced based on parents' ethnic classification. These bloodline-based administrative categories, through which everyone's ID card has an ethnic classification, increasingly serve as legitimate bases upon which to self-identify or identify others as Zhuang, in contrast to the historic centrality of shared Zhuang language.

In contrast, a discourse of Zhuang language as crucial for authenticity was expressed by a small number of participants who themselves had high Zhuang proficiencies, including literacy. Hoz, a Masters student in Kunming, told me he had already considered this issue:

But even if you do not have that Zhuang bloodline, but [you] are able to use [Zhuang language], then in the long-term [you're] together with the Zhuang, accept Zhuang beliefs, actually we recognise that person is Zhuang. But if you do not use our language, I would not identify with you much. [11.6.2014]

While this position may be common amongst some Zhuang, it is important that this rejection of non-Zhuang speakers as authentically Zhuang was rare amongst mobile young people. Being able to identify as Zhuang without the "costs" of Zhuang language is in the interests of mobile young people. That Zhuang language is unnecessary as a "way of belonging" (Levitt & Glick Schiller, 2004, p. 1010) to the home community may be accepted by those with whom the mobile students share other ways of belonging, but this will not necessarily endure if an "imagined community" (Anderson, 1991) of Zhuang speakers becomes unimaginable. The students' discourses revealed a reliance on knowing others speak Zhuang to make Zhuang a meaningful category. For example, Liz emphasised addressing my hypothetical of few or no more Zhuang speakers with two discrete outcomes:

Liz: I think if very few people speak [Zhuang]...maybe the *minzu* as a thing can still exist at least in some people's hearts. But if there were no Zhuang

speakers, I think this *minzu* really would have a bit of an “exists in name only” feel. [27.6.2014]

Conclusion

This paper argued that, as hypothesised, from the perspective of Zhuang, the mutual reinforcement of English and Putonghua, rather than instability between them, stands out. I identified key language ideologies constructing this dominant coalition: Patriotic English, Standard Language and Zero-Sum Language Ideology. I explained the complementarity of Putonghua and English in different fields and on different national and international scales through which the value of their capital is constructed. However, contrary to the hypothesis derived from Yang’s study of Naxi, Zhuang is not enduring as a symbolic resource of affinity with home and ethnic identity, at least for those with, or aspiring to create, a *habitus* of mobility. Rather, identifying as a “non-Zhuang-speaking Zhuang” is increasingly common and discursively rationalised, though the authenticity of non-Zhuang speakers remains a concern for some Zhuang. Zhuang language is no longer mobilised without reflexivity as a “way of being” (Levitt & Glick Schiller, 2004, p. 1010); on geographic and socio-economic development scales, Zhuang is re-semiotised not as an index of ethnicity, and therefore a resource of identity and inclusion, but as an index of ‘staticity’ in the peripheries of place and class.

References

- Benedict Anderson (1991).** *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Verso: London, UK..
- Gegentuul Baioud (2015).** ‘Detours’ taken by Mongols on WeChat. Retrieved from <http://www.languageonthemove.com/language-globalization/detours-taken-by-mongols-on-wechat>
- Jan Blommaert (2007).** Sociolinguistic Scales. In *Intercultural Pragmatics*, 4:1, pp1-19.
- Bloomberg News. (2012).** China’s Urban Population Exceeds Countryside for First Time. In *Bloomberg News*, 17.1.2012. Retrieved from <http://www.bloomberg.com/news/articles/2012-01-17/china-urban-population-exceeds-rural>
- Pierre Bourdieu (1977).** The economics of linguistic exchanges. In *Social Science Information*, 16:1, pp645-688.
- Pierre Bourdieu (1991).** *Language and Symbolic Power* (J. B. Thompson Ed.). Polity Press: Cambridge, UK.
- China Encyclopedia Compilation Group. (2008).** *China Encyclopedia (English Version)*. China Intercontinental Press: Beijing, China.
- Nikolas Coupland (2001).** Introduction: Sociolinguistic theory and social theory. In Nikolas Coupland, Srikant Sarangi and Christopher Candlin, *Sociolinguistics and Social Theory* Longman: London, UK, pp1-26.

- Abram de Swaan (2001).** *Words of the World: The Global Language System*. Polity Press: Cambridge, UK.
- Joanne Smith Finley & Xiaowei Zang (2015).** Language education and Uyghur identity: an introductory essay. In J. S. Finley & X. Zeng (Eds.), *Language, Education and Uyghur Identity in Urban Xinjiang*. Routledge: Oxford, UK, pp1-33.
- Joshua A. Fishman (1999).** *Handbook of Language and Ethnic Identity*. Oxford University Press: Oxford, UK.
- Shuang Gao (2015).** Multilingualism and good citizenship: The making of language celebrities in Chinese media *Working Papers in Urban Language and Literacies*. King's College London: London, UK, Paper 166.
- Yihong Gao (2009).** Sociocultural Contexts and English in China: Retaining and Reforming Cultural Habitus. In Joseph Lo Bianco, Jane Orton and Gao Yihong (Eds.), *China and English: Globalisation and the Dilemmas of Identity*. Multilingual Matters: Bristol, UK, pp57-78.
- Alexandra Grey (2015).** *Urbanisation and minority language in China: a case study of urban, upwardly mobile and language-less Zhuang people*. Paper presented at the Australian Linguistic Society conference. Western Sydney University: Sydney, NSW, Australia.
- François Grin (1995).** Combining immigrant and autochthonous language rights: A territorial approach to multilingualism. In Tove Skutnabb-Kangas, Robert Phillipson and Mart Rannut (Eds.), *Linguistic Human Rights: Overcoming Linguistic Discrimination*. Mouton de Gruyter: Berlin, Germany, pp31-48.
- Rongxing Guo (2013).** *China's ethnic minorities: social and economic indicators*. Routledge: Abingdon, Oxford, UK.
- Stevan Harrell (1993).** Linguistics and Hegemony in China. In *International Journal of the Sociology of Language*, 103:1, pp97-114.
- Monica Heller (2010).** Language as a Resource in the Globalized New Economy. In N. Coupland (Ed.), *Handbook of Language and Globalization*. Wiley-Blackwell: Oxford, UK, pp349-365.
- Monica Heller, Joan Pujolar & Alexandre Duchêne (2014).** Linguistic commodification in tourism. In *Journal of Sociolinguistics*, 18:4, pp539-566.
- David Holm (2013).** *Mapping the Old Zhuang Character Script: A Vernacular Writing System from Southern China*. Brill: Leiden, Netherlands.
- Katherine Palmer Kaup (2000).** *Creating the Zhuang: Ethnic Politics in Yunnan*. Lynne Rienner Publishers Inc.: Boulder, CO, USA.
- Steinar Kvale & Svend Brinkmann (2009).** *Interviews: learning the craft of qualitative research interviewing*. Sage: Thousand Oaks, CA, USA.
- Law of the People's Republic of China on the Standard Spoken and Written Chinese Language (2000).** Chinese Law & Government.
- Peggy Levitt & Nina Glick Schiller (2004).** Conceptualising Simultaneity: A Transnational Social Field Perspective on Society. In *International Migration review*, 38:3, pp1002-1039.
- Yongxian Luo (2008).** Zhuang. In Anthony Van Nostrand Diller, Jerold A Edmondson and Yongxian Luo (Eds.), *The Tai-Kadai Languages*. Routledge: London, UK..
- Thomas Shawn Mullaney (2011).** *Coming to Terms with the Nation: Ethnic Classification in Modern China*. University of California Press: Berkeley, CA, USA.

- Ingrid Piller (2015).** Language Ideologies. In K. Tracy (Ed.), *The International Encyclopedia of Language and Social Interaction*. John Wiley & Sons Inc.: Boston, MA USA, pp1-10.
- Ingrid Piller & Jinhyun Cho (2013).** Neoliberalism as language policy. In *Language in Society*, 42:1, pp23-44.
- Junxi Qian, Liyun Qian & Hong Zhu (2012).** Representing the imagined city: Place and the politics of difference during Guangzhou's 2010 language conflict. In *Geoforum*, 43:5, pp905-915.
- Dongran Ru Blachford & Marion Jones (2011).** Trilingual Education Policy Ideals and Realities for the Naxi in Rural Yunnan. In Anwei Feng (Ed.), *English Language Education Across Greater China*. Multilingual Matters: Bristol, UK, pp228-260.
- Elana Goldberg Shohamy (2006).** *Language policy : hidden agendas and new approaches*. Routledge: London, UK.
- Miriam Rom Silverberg (2006).** *Erotic Grotesque Nonsense: The Mass Culture of Japanese Modern Times*. University of California Press: Berkeley, CA, USA.
- Regie Stites (1999).** Writing Cultural Boundaries: National Minority Language Policy, Literacy Planning, and Bilingual Education. In Gerard. A. Postiglione (Ed.), *China's National Minority Education: Culture, Schooling and Development*. Falmer Press: New York, NY, USA, pp95-130.
- Nicholas Tapp & Don Cohn (2003).** *The Tribal Peoples of South West China*. White Lotus: Bangkok, Thailand.
- Kathryn Ann Woolard & Bambi B. Schieffelin (1994).** Language Ideology. *Annual Review of Anthropology*, 23, pp55-82.
- Hongyan Yang (2012).** *Naxi, Chinese and English: Multilingualism in Lijiang*. (PhD), Macquarie University: Sydney, Australia. Retrieved from http://www.languageonthemove.com/wp-content/uploads/2013/03/PhD_Thesis_Hongyan_Yang.pdf
- Yichuan Yuan, Deying Hu, Peng Li, Honghua Zhu, Jinjun Wang, Yun Shang & Hongbin Ba (2015).** A Survey Report on Trilingualism and Trilingual Education in Yunnan. In Anwei Feng & Bob Adamson (Eds.), *Trilingualism in Education in China: Models and Challenges*. Springer: Dordrecht, Netherlands, pp175-198.
- Jie Zhang (2011).** *Language Policy and Planning For the 2008 Beijing Olympics: An Investigation Of The Discursive Construction Of An Olympic City And A Global Population*. . (PhD), Macquarie University: Sydney, Australia.
- Yongnian Zheng (2009).** Identity Crisis, the "New Left" and Anti-Westernization Search for New National Identities: Modernization vs "Westernization". In Yongnian Zheng (Ed.), *Discovering Chinese Nationalism in China: Modernization, Identity, and International Relations* Cambridge University Press: ambridge, UK, pp46-66.
- Yun Zhou (2013).** *Question of Ethnic Group Formulation in the Chinese Census*. Paper presented at the XXVII IUSSP International Population Conference, Busan.

5 Revisiting the action research cycle: Critically examining three cases of classroom-based language learning inquiry

Gerald Talandis Jr. & Theron Muller

University of Toyama

talandis@hmt.u-toyama.ac.jp

theron@las.u-toyama.ac.jp

Abstract

Action research (AR) as a means for language teachers to critically explore and investigate their practice has been advocated as part of teacher training and masters-level development courses for some time (Nunan, 1990). While models of the AR process vary, central to all is a reflective research cycle of planning, acting, observing, and reflecting (Burns, 2010). However, models promoted by AR advocates tend to represent an idealised implementation and may not reflect the actual experience of teachers pursuing classroom-based research. This investigation explores three cases of classroom research from the perspectives of teachers. Data consists of participant research by the two authors as well as the experiences of a third language teacher, collected through semi-structured interviewing (Myers & Newman, 2015) and examination of different versions of her written work submitted and revised for publication. Our findings suggest that disseminating the results of AR projects was often considered very early in the research process and was frequently geared, at least initially, toward oral presentations followed later by written publication. Literature reviews tended to occur after data had been collected, often in concert with preparing for presentations and submission for publication. In addition, investigations frequently ended after one cycle of research. Based on these findings, an alternative model for representing action research that more fully reflects the experience of these three practitioner-investigators is proposed. How the model may help inform the way teacher research is conceptualised and taught is also discussed.

Introduction

This paper describes an investigation into three Japan-based language teachers' experiences of conducting classroom-based research. While there have been a number of models of action research (AR) proposed in the literature (Altrichter, Posch & Somekh, 1993; Hopkins, 1993; Nunan, 1992), these are generally based on researchers observing teachers doing classroom

research (CR), reifying their experiences into models of teacher research practice, then using those models to guide teacher-researchers in in-service teacher development courses. While the literature includes a number of guides for language teacher action research (Burns, 2010; Freeman, 1998; Wallace, 1998), there has been less attention paid to critical analysis of the different steps advocated as part of the AR cycle from the perspective of the teachers doing research. To fill this gap, this investigation addresses the following research questions:

- To what extent do three teachers' CR experiences follow a hybrid model of the traditional AR process?
- How can this hybrid AR model be updated to more accurately reflect their CR experiences?

This paper begins with a review of some of the key AR models from the literature, followed by an overview of the research methodology and the two kinds of data used. The data section explains how the teachers' CR mapped onto a hybrid AR model based on the literature reviewed and explores the extent to which the model reflects the teachers' AR experiences. Finally, an alternative model that more fully represents the three teachers' CR experiences is proposed, with discussion focusing on implications for understanding teacher research.

Examining action research models

This section summarises four models of AR: Kemmis and McTaggart (1988), Elliott (1991), Nunan (1992), and Altrichter, Posch, and Somekh (1993) then concludes by describing the hybrid model of AR we use to analyse our teacher CR data. Key themes that inform the hybrid AR model are examined. The AR models reviewed here were developed from researchers observing teachers as they did teacher research (Hopkins, 1993). Based on these observations, the models were made to represent what teachers do in their classrooms. These representations in turn have been applied to in-service and pre-service teacher training courses as instruction on how to go about doing CR. As teacher-researchers ourselves, one of our objectives has been to examine CR from a teacher's perspective rather than from the viewpoint of researchers studying teachers.

Figure 1 represents one of the earliest AR models available, developed by Kemmis and McTaggart (1988, p. 14). It follows a cycle of "plan, action & observation, reflection" (pp. 11-13), leading into a "revised plan" (p. 14)

which then cycles through another round of investigation. One of this model's major innovations is that it accommodates multiple cycles of teacher research, acknowledging that from the teacher's perspective, AR is an ongoing, reiterative and exploratory process. However, it has been criticised as being too fixed and rigid. This can lead to an overly prescriptive interpretation, one that requires teachers to precisely follow its steps and sequences, with the implicit assumption that AR is unidirectional and centred around a single, unchanging investigative theme across multiple iterations (Burns, 2010).

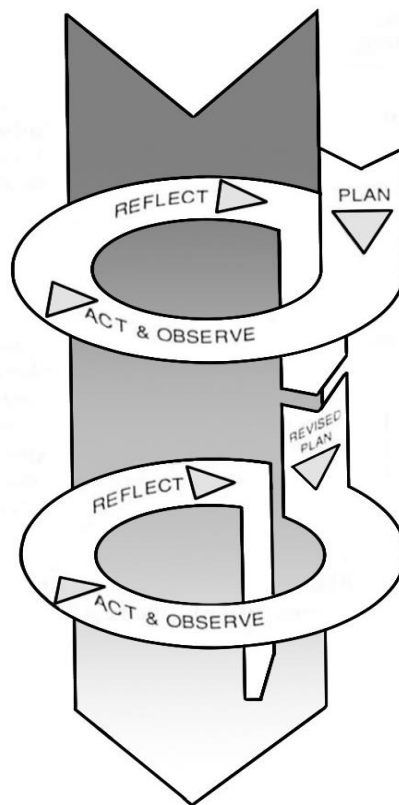


Figure 1: Kemmis & McTaggart's AR model (1988, p. 14)

Elliott's (1991) three-cycle model (Figure 2) provides a "revised" (p. 70) version of Kemmis and McTaggart (1988). One innovation is that the research focus is not static and can evolve within the cycle. This means the original research 'cycle' may not be completed, as steps taken during it may lead to a change in investigative focus. This can, in turn, spawn a new cycle of AR, which can also recursively initiate additional new cycles (Elliott, 1991). In addition, Elliott's (1991) "reconnaissance" (p. 70) step, similar to Kemmis and McTaggart's (1988) plan and reflect, is not depicted as only coming at the beginning or end of the cycle, but can also be embedded between different "action steps" (p. 70). This suggests activities such as fact-

finding or data-analysis are also not restricted to the beginning and end of the AR cycle but can recur throughout the research process. However, in the Kemmis and McTaggart (1988) and Elliott (1991) models, dissemination, where action researchers share their findings, is implicit rather than explicit.

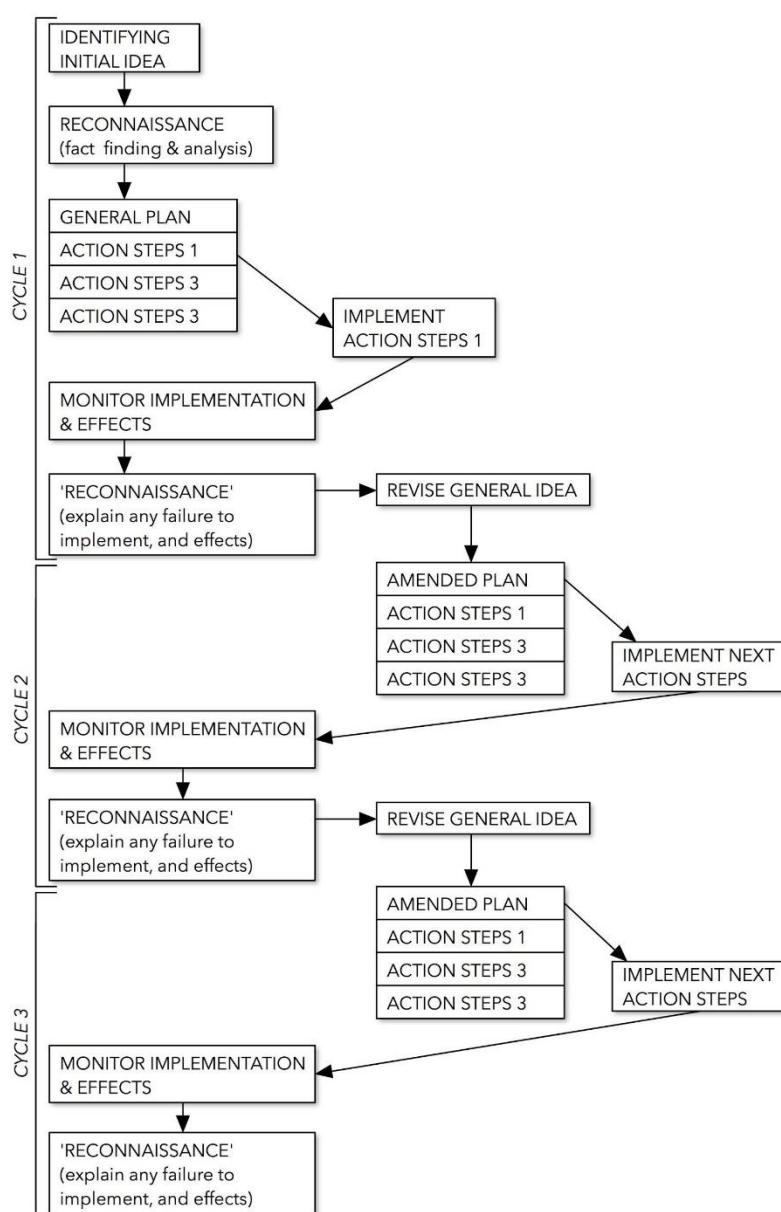


Figure 2: Elliott's AR model (1991, p. 71)

The final two models discussed here (Figure 3), Nunan (1992) and Altrichter, et al. (1993), incorporate explicit mention of dissemination, an important aspect of the teacher data we collected. In Altrichter et al. (1993), the recursive nature of the Kemmis and McTaggart (1988) and Elliott (1991) models is captured through the arrow linking stage C to stage B, with stage D, "Making teachers' knowledge public" (p. 7) representing when teachers

disseminate their research. With the Nunan (1992) model, “results are disseminated” (p. 18) for example, in “a workshop for colleagues and ... a paper at a language conference” (p. 19) and then followed up in “the form of an ongoing cycle ... in which the teacher reflects on, returns to, and extends the initial query” (p. 18). Thus in Altrichter, et al. (1993), dissemination is an end product of the AR cycle, coming after multiple cycles of investigation, while in Nunan (1992), dissemination is built into the recursive AR cycle, repeating as a “step” (p. 19) in the research, with potential for multiple iterations of dissemination in any given investigation.

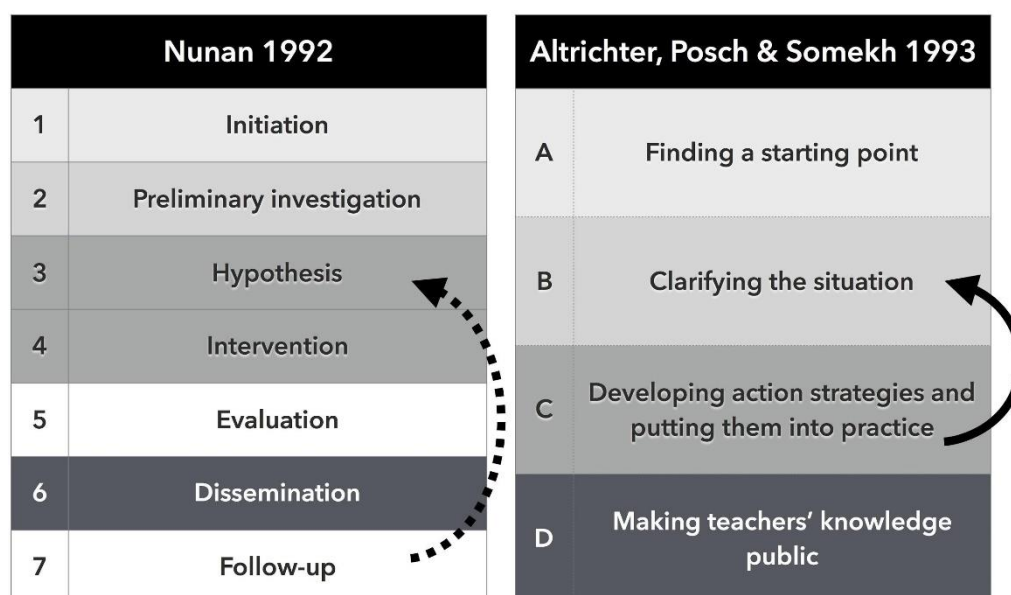


Figure 3: AR models from Nunan (1992) and Altrichter, et al. (1993)

Finally, not all researchers agree that teachers conduct AR in the same order as the graphical representations above suggest (Burns, 2010). As Burns (2010) notes, “AR processes involve many interwoven aspects—exploring, identifying, planning, collecting information, analysing and reflecting, hypothesising and speculating, intervening, observing, reporting, writing, presenting—that don’t necessarily occur in any fixed sequence” (p. 8). Nevertheless, she uses the Kemmis and McTaggart (1988) model as a means of rhetorically organising her discussion of AR. This illustrates that the models themselves can influence the structure of the AR discussion in the literature, and how “they are potentially prescriptive and are disconnected from real-life practices” (McNiff, 2013, p. 63). It is the influence of the models discussed above on the conversation surrounding teacher research that we are seeking to probe and question here.

In terms of the hybrid AR model we map our data onto, we consider the cyclicity of AR an important feature as well as the potential for the focus of research to change. As will become apparent, issues of reconnaissance and dissemination and how they fit into the models are problematic, although the most compatible of the models above is Nunan's (1992), where dissemination can happen as part of the AR cycle rather than outside of it. Having reviewed the literature on AR models and summarised the main features of our hybrid AR model, next we turn to the methods of this investigation.

Methods: Creating classroom research profiles for three Japan-based language teachers

The CR practice of three teachers is examined for this investigation. The data for the first two teachers, the authors of this article, was collected through what Erickson terms (1985) "participant-observational research" (p. 47). This methodology is "concerned with the relation between meaning-perspectives of actors and the ecological circumstances of action in which they find themselves" (Erickson, 1985, p. 29) and is intended to "enable researchers and practitioners to become much more specific in their understanding of the inherent variation from classroom to classroom" (p. 47) in order to "give a detailed view of the actual structure and process of program implementation" (p. 41). In our case, this involved listing the different parts of Gerald and Theron's respective CR projects and working together to organize the information using the hybrid AR model.

The data for 'Kathy' (a pseudonym), the third teacher, was gathered as part of an ongoing investigation into Japan-based authors' writing for publication practice, which incorporates semi-structured interviewing (Gillham, 2005) and collection of multiple versions of manuscripts intended for academic publication (Muller, 2016). Kathy's data was examined to reconstruct the process her CR went through. This information was then similarly applied to the hybrid AR model. In addition, Kathy was consulted via email regarding the interpretations made of her data, and where appropriate we adjusted our analysis based on her feedback. It was felt this recursivity of analysis, consultation with Kathy, and reanalysis would lead to a more robust representation of what the experience of CR was for her (Guba & Lincoln, 1989).

As Gerald and Theron have a number of CR projects with relatively complex narratives, they each selected one for detailed analysis. Since Kathy has completed a total of three projects with relatively straightforward narratives,

all three were analysed and mapped on to the hybrid AR model. In the second part of our analysis, we examined problems in fit between the hybrid model and the three teachers' research experiences. This led to the final stage of our analysis where we considered strategies for modifying the hybrid model to address the problems identified earlier. This revised model was then used to remap multiple CR projects for Gerald and Theron and to remap all three of Kathy's projects in order to explore its robustness and fit for modelling teacher CR practice.

The three teachers' data: Mapping experience onto a hybrid model of action research

This section maps the three teachers' CR data onto the hybrid AR model. Gerald's is the only example of deliberate AR, with Theron and Kathy conducting CR that they did not explicitly position as AR. Since the investigation presented here is participant research, discussion of the data is intentionally presented from a 3rd-person perspective.

Gerald's data: An intentional classroom action research project

Gerald supplied CR data for an intentional AR project conducted with a colleague between 2011-12 at a Japanese university near Tokyo, with dissemination ongoing through 2014. The project involved developing a classroom intervention to make a more interactive English language syllabus, with the goal of improving students' speaking skills. As this project was framed as AR, it intentionally incorporated 'cycles', with three ultimately represented in the data. The first two easily map onto the hybrid AR model. However, the third, which was completed as part of the manuscript revision process after submission to a journal for review, represents a data-focused analysis cycle rather than a cycle of additional research conducted in the language classroom (Appendix A).

The first cycle began in spring 2011 and followed a typical AR process, including planning, action, observation, and reflection. The project deviated from the hybrid AR model in a few ways. First, there was no 'preliminary investigation' or 'fact-finding' stage. In addition, the dissemination process began early on via the submission of a conference presentation proposal and selecting a target journal in which to publish findings, activities that are typically restricted to the end of the research cycle in AR models. The second cycle, in the fall semester, concentrated specifically on a particular subset of

students, a shift reminiscent of Elliott's (1991) representation of the AR process as potentially changing investigative focus.

At the end of the year-long intervention, work on dissemination began in earnest. This involved three separate but related projects: a presentation at an international conference held in Japan (Talandis Jr. & Stout, 2012), a research report in an international journal (Talandis Jr. & Stout, 2015a), and a book chapter reflecting on the AR process and the challenges of balancing teacher and researcher roles (Talandis Jr. & Stout, 2015b).

Theron's data: Investigating writing across classroom contexts

Theron's data consisted of his investigations into writing over a five year period from 2010-15. The thematic focus on writing meant that his CR activities over multiple institutions and classrooms were included (Appendix B). While characterizing these activities as CR, Theron did not explicitly set out to place his investigations within the AR paradigm. Nevertheless, there are several ways in which his research resembles AR: it is classroom based, addresses classroom issues, monitors pedagogical interventions, and is cyclical in nature.

Differences between Theron's research experience and the hybrid AR model arise when examining continuity in his research across different institutions, classes, and teaching contexts. Specifically, Theron's first writing research project was based on CR conducted in Nagano from 2010 to 2011 and was ultimately developed into one presentation (Muller, 2011) and a book chapter (Muller, 2014c). At the end of 2011, Theron took a new position at the University of Toyama and brought his interest in investigating the teaching of writing with him into his new role.

At the University of Toyama, Theron had an initially difficult experience teaching third year medical students (Muller, 2012). This led to a decision to revise his syllabus and change the focus of the course to examining doctor-patient medical discourse. The assessed component of the course changed from oral presentations to student written reflections, which were used to judge teaching effectiveness and make more explicit what students were getting from the class. As students submitted those reflections, their potential for yielding interesting data which could form the basis for further examination became evident. Therefore, Theron retained copies of the students' reports for later examination. This interest in doing further research into his students' writing led Theron to submit presentation and publication

proposals, eventually leading to two conference presentations (Muller, 2012b, 2014c), the journal piece previously cited (Muller, 2012), and one book chapter examining the evidence for student autonomy in their reflective writing (Muller, 2014a).

Kathy's deliberate classroom research: Not 'real action research'

Kathy's CR data comes from 2009 to 2011, when she was working part-time in Japan as an adjunct lecturer. It consists of three independent research projects conducted in Japanese university classrooms (Appendix C). The first project, conducted independently, involved examining the use of films and film clips in class. The second, completed with a full-time colleague, implemented 'paperless teaching'. The third and final project involved students attempting an 'English-only project' that Kathy worked on with two other colleagues.

Each project followed a similar pattern, with an initial planning of the specific intervention and submitting a conference presentation proposal before the implementation stage had even begun. This was followed by the investigation itself. Upon completion of each project, the research findings were disseminated through a conference presentation for each of the three projects followed by a proceedings paper in two cases (*Film Clips in Class* and *English-only*) and a journal publication (*Paperless Teaching*). Kathy has been assured anonymity as a participant in this research, so citations of her presentations and work produced have been deliberately omitted.

Discussion

Problems with fit between the hybrid AR model and the three teachers' data

The data presented above reveals a number of differences between the hybrid AR model and how the three teachers actually carried out their CR. Regarding Gerald's deliberate AR investigation, the project began without a fact-finding or reconnaissance stage, as he and his colleague were clear on what they wanted to do from the start. In addition, a third research-oriented stage followed the two stages of classroom pedagogy-focused AR. This stage does not easily fit into the hybrid AR model, which tends to consider cycles of research revolving around classroom investigations rather than cycles of data analysis. Furthermore, their literature review was completed as they went about disseminating their findings, rather than at the beginning of their

investigation as the models would predict. Finally, dissemination was instrumental in helping Gerald and his colleague develop the narrative of their research. In other words, they learned what to say as they were saying it. As they described their experience:

Looking back, it is clear that the process of publicly sharing our research findings was instrumental in building our current understanding of the project. Reporting what we experienced elicited more questions which resulted in more answers which brought up further questions that led to deeper and deeper insights. Articulation was the key that unlocked deeper clarity and comprehension. In the end, it was not a process of figuring out everything first and then sharing; we had to discover our story through its telling. It was a simultaneous process. (Talandis Jr. & Stout, 2015b, p. 23)

Examining Theron's Nagano writing research from the hybrid AR model perspective, while it was conceived as CR, it was not designed to be cyclical; the intervention instead involved a single academic year. The data was ultimately presented on (Muller, 2011) and published in a book chapter (Muller, 2014c). After moving to the University of Toyama, Theron changed his classroom practice to respond to pedagogic dilemmas and challenges emerging from his new context and students. However, these changes were pedagogical rather than research oriented. As he implemented these changes, Theron felt his students' written reflections were potentially interesting, but he did not formulate a clear plan regarding how to collect and analyse data. Rather, when his interest in researching his students' writing manifested, he began seeking outlets and opportunities for dissemination. In contrast with AR models, Theron's data was first collected, and it was only later when assembling presentation abstracts and publication proposals that the issue of what to examine and how to go about analysing it was addressed. Thus while AR is generally depicted as seeking to improve classroom pedagogy, in this case, the intervention and the research generally were not intrinsically linked. In other words, the results of Theron's research were not necessarily intended to help inform his pedagogy. Rather, the investigative stance taken was one where doing the research into what students were writing (as opposed to the pedagogic assessment of their writing) would yield insights regarding what was going on in the classroom, insights which were not intended to influence Theron's classroom pedagogy to the extent AR models would imply. In addition, the literature he reviewed before teaching was intended for classroom pedagogy purposes, such as examining the medical discourse analysis literature with the intention of selecting specific topics for use in class. Later, when seeking avenues for dissemination, the ESP literature was consulted to situate Theron's classroom practice within the ESP field.

Furthermore, there was thematic carryover from earlier to later projects, specifically the theme of investigating student writing, a theme which crossed institutions and classroom contexts. This topic of professional mobility in English language teaching is raised in the literature generally (Hans-Vaughn & Scherff, 2008), but is not explicitly accounted for in depictions of teacher AR.

Kathy's data suggests she followed an AR process, but one that also varied from the hybrid AR model. First, she began each project with a clear intention to disseminate her findings, initiating dissemination before data collection began. Another difference was that cyclicity was missing in her investigations; each one was timed to be conducted between submission of a conference presentation proposal and the actual presentation, generally encompassing research of about a semester in length. Also, in-depth literature reviews were generally not conducted until the time came for dissemination. Additionally, in each project, there was no preliminary investigation or reconnaissance stage, similar to Gerald's project. Finally, despite some overlap in time where she was writing up one project while beginning another, Kathy considered each project as separate, with no thematic consistency across them. This is in contrast to the consistency of CR themes found in Theron's data. However, this lack of thematicity was not necessarily by design, as Kathy was particularly subject to changing employment circumstances during this period, moving between classes and institutions. As she put it in an email, "... I never stayed in one place long enough to actually complete anything substantial" (personal correspondence, 2015).

In reflecting on her research, Kathy did not consider what she had done AR, writing in an email, "well, you know, it is classroom research, but I didn't plan this as "real" action research. I never did any pilot testing or vetted anything or did anything by the book" (personal correspondence, 2015). This comment suggests she perceives AR as prescriptive. That she did not follow what she thought was the 'correct' way to conduct AR contributed toward a feeling that her investigations were in some sense unsuccessful; despite three conference presentations and three published papers, she concluded "I'm not able to do this" (personal correspondence, 2015).

The topic of Kathy's research objectives is worth exploring in a bit more depth. Reflecting on her CR practice, she wrote in an email:

The research for AR really seems to be a separate entity from teaching and putting them together sometimes works but not always. ... I always feel like I

struggle to get the teaching done well. Throwing research into that mix has always been just stressful, and for someone like me—non-tenure, no hope of tenure, it seems like a waste of time. Describing what I was doing in the classroom seemed to be the most useful in terms of giving other teachers ideas to use in the classroom. That's ultimately all I ever wanted from research—ideas to make teaching easier and more fun for everyone involved (personal correspondence, 2015).

It seems here that rather than seeking to speak to an AR-oriented audience, Kathy was instead envisioning her research as ideally addressed to fellow classroom language teachers. Put another way, she was more concerned with the craft of teaching than with theorising her practice for a wider audience of researchers. It seems, however, that she has been unable to balance the demands and expectations that come with doing CR against the teaching responsibilities that she feels occupy the majority of her professional time and attention.

To summarise, from the viewpoint of the hybrid AR model, each investigation is a separate project, lacking a larger context or potentially influential factors, such as change of workplaces or classes. In Theron's case, a single investigation begun at one school continued on in a different form at another which required representing them as separate projects, resulting in a loss of thematic continuity from his point of view. This is also true with Gerald's data, where prior and subsequent research interests were not taken into account by the hybrid model's focus on single projects. Finally, the dissemination stage was not limited to the last step of the investigation cycle, where it is often placed in AR models.

Moving toward a new means for modelling classroom research

While various aspects of the three teachers' CR practice could be mapped onto the hybrid AR model, there were several incompatibilities, such as lack of pre-planning activities and distribution of the dissemination stage throughout the research process. In addition, research cyclicity was only observed in Gerald's data. All told, these differences suggest that AR models could be updated to more explicitly account for these teachers' experiences.

One way to better accommodate our data would be to move from a model of AR that places specific classroom interventions at its centre to one with the classroom teacher as the central focus (Figure 4). This alternative viewpoint offers certain advantages when representing our teacher research data. With the teacher at the centre, it is possible to accommodate multiple potential

teaching or research themes. As Figure 4 shows, each theme can have one or more projects associated with it. The inner arrow encircling each project indicates a typical AR research cycle, with the outer circle representing dissemination work. This updated model allows for multiple projects to be associated with particular themes, with the dashed arrows representing relationships between different themes and projects. It also accommodates instances where there are cycles of investigation not immediately oriented toward classroom pedagogy, such as in Gerald's and Theron's cases.

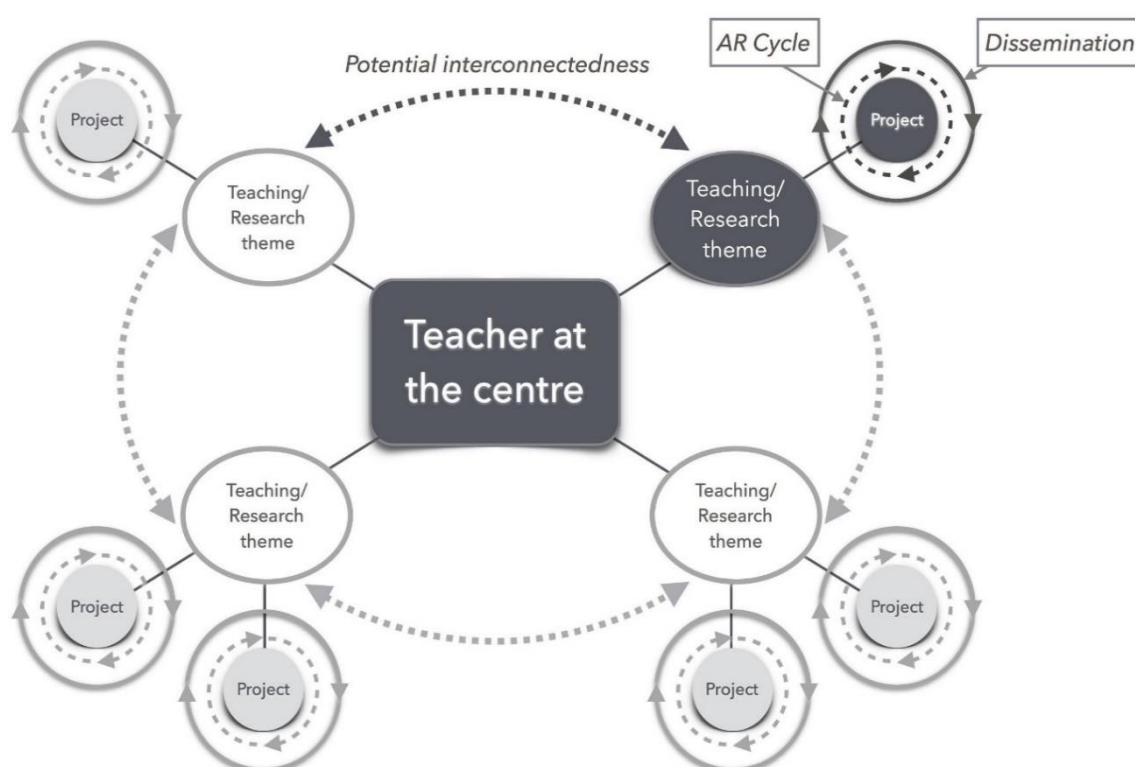


Figure 4: Teacher-centred/thematic view of AR process

Applying the new model to the teachers' data

This section examines the three teachers' data in light of our proposed AR model that puts the teacher rather than the intervention at its centre. We feel there are four primary improvements this model brings: It foregrounds thematic links between different research projects, represents continuity across projects and teaching contexts, accommodates multiple dimensions of teacher characteristics, and potentially allows for revaluation of the role the technical craft of teaching can play in CR. Each of these advantages are explored below through reconfigured representations of the teachers' CR projects.

With respect to Gerald's data, his single isolated AR project (Appendix A) becomes contextualised within related research interests (Figure 5). In Appendix A, there is only one project represented while in Figure 5, that project is subsumed under his general interest in *teaching speaking skills*. This holistic perspective is instructive, as arrows pointing from previous and toward subsequent research themes ('*applied pragmatics*' & *reflecting on classroom recordings*) indicate what came before his AR project and where he has gone with his research since.

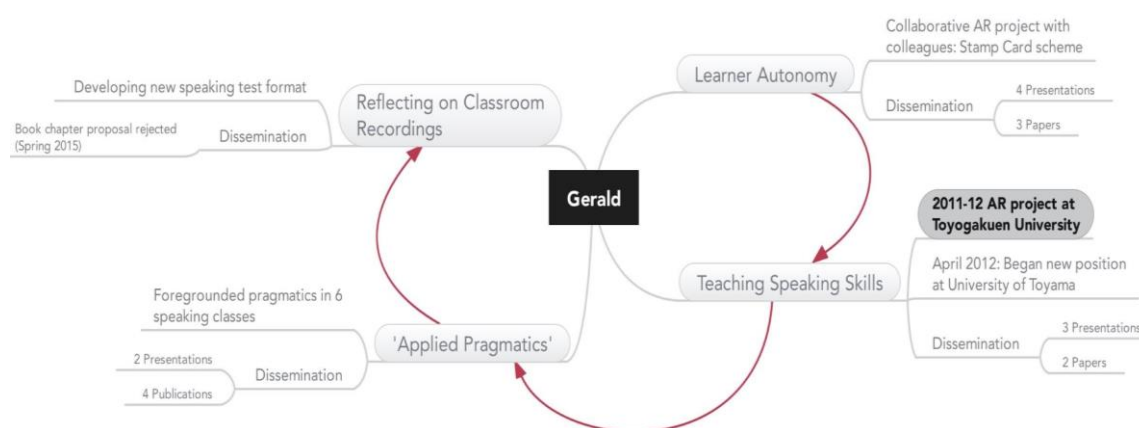


Figure 5: Gerald's data: A teacher-centred view

Theron's data (Figure 6) shows similar progressions and thematic linkages, with *task-based language teaching* represented as an older strand of CR interest which gave way to his more recent investigations into teaching writing. This interest in writing has several projects associated with it, including *medical English* and an exploration of *learner agency & motivation*. Moreover, Theron's interest in *academic writing for publication* represents a more research-oriented than classroom teaching-oriented strand of inquiry. In summary, our updated model allows for representing connections between interests and projects while illustrating how we as teachers have developed professionally in a much more integrated manner than more traditional AR models are designed to accommodate.

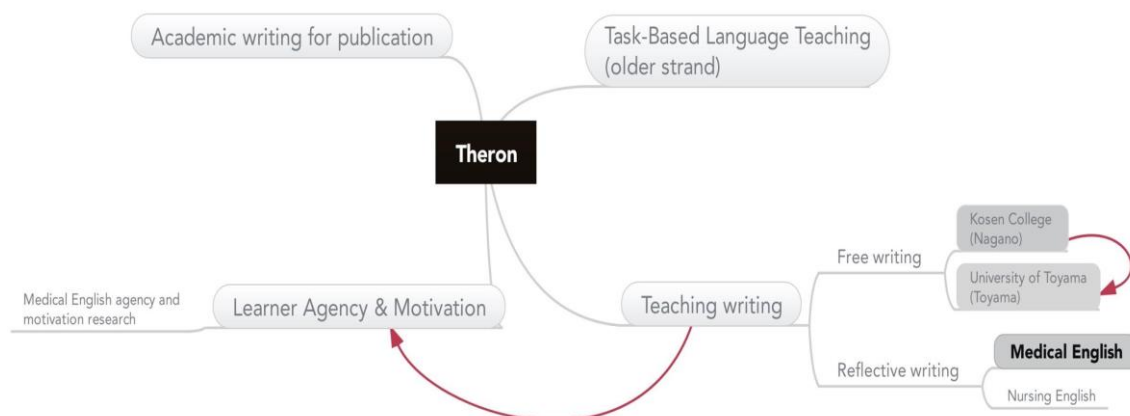


Figure 6: Theron's data: A teacher-centred view

Regarding continuity across projects and contexts, Theron's data is informative, as Figure 6 shows the similarities between Nagano and Toyama regarding his interest in teaching *free writing*. This is also shown in Gerald's data (Figure 5), where he notes his move to the University of Toyama under his *teaching speaking skills* interest. While Kathy's data does not include information about micro-level moves from teaching at one institution to another, there is the potential to track such changes were that information available. Furthermore, had this alternative model been available to her, it may have helped her to better understand how her CR practice had evolved, rather than leading her to conclude that her projects were not 'real action research'.

Considering how the alternative model accommodates multiple teacher characteristics, it is worth contrasting the representation of Kathy's three research projects in Figure 7 with their previous representation in Appendix C. Putting Kathy at the centre of the model allows for a fuller representation of her teacher experience; her teaching English in Japan is one theme, and her later teaching experience in two additional countries can also be represented. Kathy revealed in semi-structured interviews that part of why she was doing CR in Japan was through an interest in becoming a researcher. However, as she was not comfortable balancing teaching with research, part of what led to her move to another country was to find work where research was not as much of an issue. Not included here, but easily accommodated, are further developments in Kathy's career, such as her efforts to gain further teaching credentials through studying for the CELTA.



Figure 7: Kathy's data: A teacher-centred view

Finally, the teacher-centred model we have proposed helps reconfigure the conversation surrounding CR to accommodate more of the technical craft aspects of teaching. As noted by Garton and Richards (2015), the call for teachers to theorise their practice has been largely unsuccessful, partially because it takes a research agenda, generating theory from data, and pushes it onto teachers who may be more interested in their profession's technical craft. This is perhaps most evident in Kathy's data (Figure 7), as she felt it more important to present teaching practices of potential use to other teachers than to represent her CR as a theoretically robust examination of her practice.

Conclusion

In conclusion, we have reviewed currently available models for representing AR and have examined their fit with data from the CR experiences of three English teachers. We found there were some discrepancies between the teachers' experiences and the AR models. This led to a proposal for an alternative model of representing CR which places the teacher at its centre. We feel this alternative model allows for a more explicit representation of dissemination as encompassing CR projects, facilitates the foregrounding of thematic links between different projects, accommodates continuity across projects and contexts, makes space for including multiple dimensions of teacher characteristics and finally encourages revaluation of the role the technical craft aspects of teaching play in CR.

We should acknowledge that our proposed AR model has been developed through an examination of the CR experiences of only three teachers, so the degree to which it is compatible with additional teachers' experiences remains for future research to examine. It is also unclear whether the model we have proposed is compatible with in-service teacher training; the three classroom teachers represented here had all completed MAs in teaching

English as a second or foreign language at the time of research, and it could be that their practice as it is represented here emerged through their experience of teacher research models covered on those courses, which developed in complexity through CR experience. If this is the case, then the model discussed here may end up being incompatible with early career teacher researchers' practice. Finally, the issue of how teachers can go about legitimately disseminating their CR remains an issue in general. In Gerald's case, he adapted the representation of his practice to fit the AR model and couch his research practice in the AR tradition. Theron, on the other hand, has drawn on teaching writing traditions and ESP traditions in framing his research publications. However, in Kathy's case, she appeared unable to rectify the apparent disconnect between her CR practice and the requirements of writing that research for academic publication. She ultimately stopped doing CR to concentrate instead on her teaching craft. The dilemma of how to assist teachers in the dissemination of their CR is a theme that Garton and Richards (2015) raise, and which remains unresolved.

References

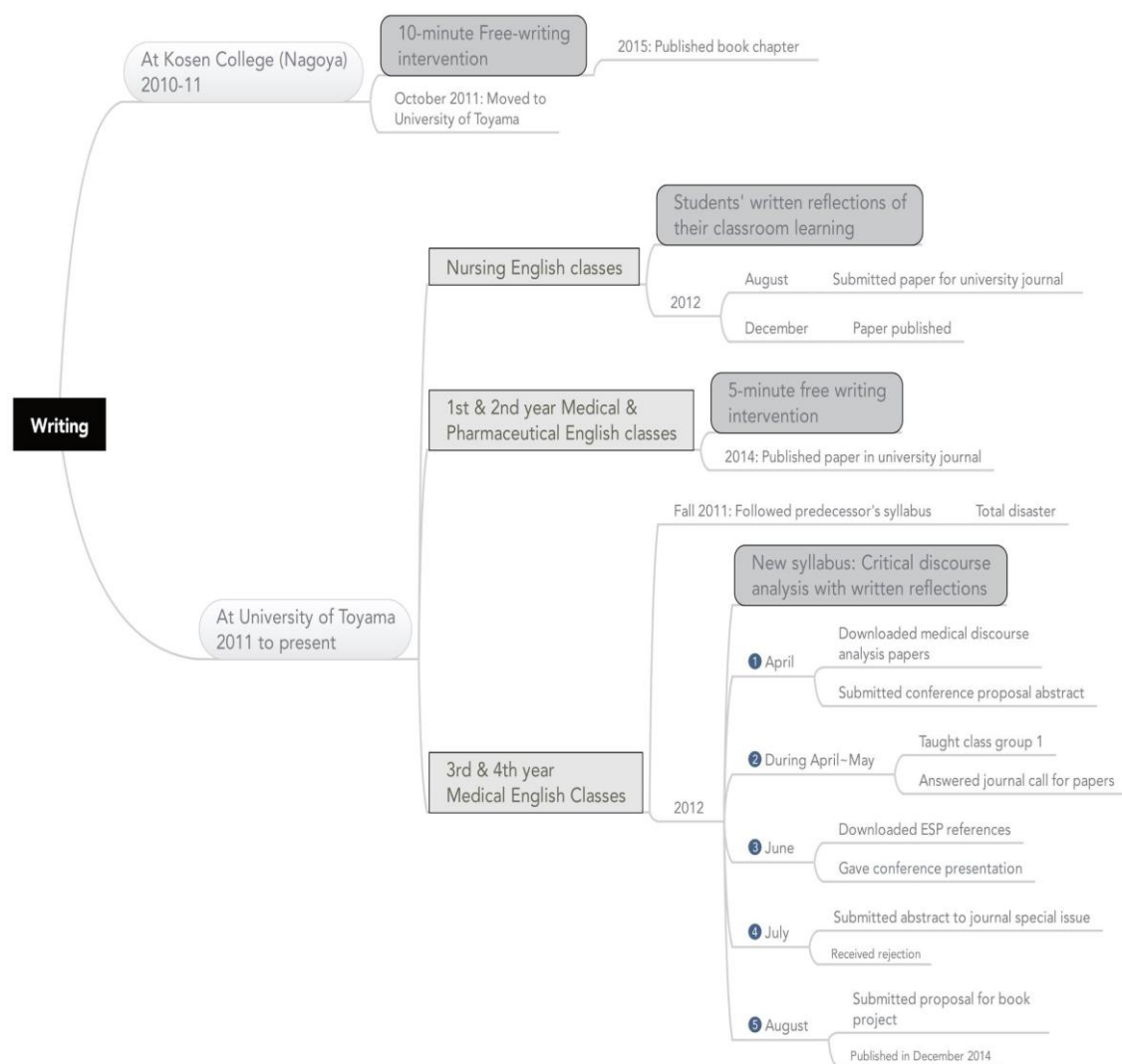
- Herbert Altrichter, Peter Posch & Bridget Somekh (1993).** *Teachers Investigate their Work: An introduction to the methods of action research*. Routledge: London, UK.
- Anne Burns (2010).** *Doing Action Research in English Language Teaching. A Guide for Practitioners*. Routledge: New York, USA.
- John Elliott (1991).** *Action Research for Educational Change*. Open University Press: Milton Keynes, UK.
- Frederick Erickson (1985).** Qualitative methods in research on teaching. *Occasional Paper No. 81*. Michigan State University: East Lansing Institute for Research on Teaching. Available: <http://files.eric.ed.gov/fulltext/ED263203.pdf>
- Donald Freeman (1998).** *Doing Teacher-Research: From Inquiry to Understanding*. Heinle & Heinle: Boston, USA.
- Sue Garton & Keith Richards (2015).** Breaking theory, (re-)valuing practice. Presented at the *48th Annual Meeting of the British Association of Applied Linguistics (BAAL 2015)*, September 3-5. Aston University: Birmingham, UK.
- Bill Gillham (2005).** *Research Interviewing: The Range Of Techniques*. Open University Press: Berkshire, UK.
- Egon G. Guba & Yvonna S. Lincoln (1989).** *Fourth Generation Evaluation*. Sage: Newbury Park, CA.
- Debbie L. Hahs-Vaughn & Lisa Scherff (2008).** Beginning English teacher attrition, mobility, and retention. In *The Journal of Experimental Education* vol 1, pp21-54.
- David Hopkins (1993).** *A Teacher's Guide to Classroom Research, 2nd Edition*. Open University Press: Buckingham, UK.
- Stephen Kemmis & Robin McTaggart (1988).** *The Action Research Planner, 3rd Edition*. Deakin University Press: Victoria, AUS.
- Jean McNiff (2013).** *Action Research: Principles and Practice, 3rd Edition*. Routledge: New York, USA.

- Theron Muller (2011).** Evaluating free writing in Japanese High Schools. Presented at the *PAN SIG 2011 Conference*, May 22. Shinshu University: Matsumoto, JP.
- Theron Muller (2012a).** Critical discourse analysis in ESP course design: The case of medical English. In *Professional and Academic English* 40, pp25-28.
- Theron Muller (2012b).** Using discourse analysis to inspire task design in ESP. Presented at the *Chubu English Language Education Society 42nd Annual Conference*, June 30-July 1. Gifu, JP.
- Theron Muller (2014a).** Critical discourse analysis in a Medical English course: Examining learner agency through student written reflections. In Ping Deters, Xuesong (Andy) Gao, Elizabeth R. Miller & Gergana Vitanova (eds), *Theorizing and Analyzing Agency in Second Language Learning: Interdisciplinary Approaches*. Multilingual Matters: Bristol, UK, pp232-251.
- Theron Muller (2014b).** Free Writing in Medical ESP: Exploring issues of research methods for intact classroom research. In 富山大学杉谷キャンパス一般教養研究紀要 (*The Journal of Liberal Arts and Sciences, University of Toyama, Sugitani Campus*) vol 42, pp29-37.
- Theron Muller (2014c).** Implementing and evaluating free writing in a Japanese EFL classroom. In Theron Muller, John Adamson, Philip Shigeo Brown & Steven Herder (eds), *Exploring EFL Fluency in Asia*. Palgrave Macmillan: Basingstoke, UK, pp163-177.
- Theron Muller (2014d).** Medical English: Introducing doctor-patient discourse analysis in the classroom. Presented at the *JALT PAN SIG 2014 Learner Development Forum*, May 10-11. Miyazaki Municipal University: Miyazaki, JP.
- Theron Muller (2016).** Exploring the practices and experiences of Japan-based language educators writing for academic publication: Examining authors' interactions with editors. *ESBB English Scholarship Beyond Borders* 2. vol 2. Available: <http://www.englishscholarsbeyondborders.org/46-2/publications>
- David Nunan (1992).** *Research Methods in Language Teaching*. Cambridge University Press: New York, USA.
- Gerald Talandis Jr. & Michael Stout (2012).** Teaching conversation: An action research project. Presented at the *37th annual meeting of the Japan Association of Language Teaching (JALT)*, October 12-15. Hamamatsu, JP.
- Gerald Talandis Jr. & Michael Stout (2015a).** Getting EFL students to speak: an action research approach. In *English Language Teaching Journal* 69:1, pp11-25.
- Gerald Talandis Jr. & Michael Stout (2015b).** Towards new understandings: Reflections on an action research project with Japanese university students. In Simon Borg & Hugo S. Sanchez (eds), *International Perspectives on Teacher Research*. Palgrave Macmillan: Basingstoke, UK, pp14-28.
- Michael J. Wallace (1998).** *Action Research for Language Teachers*. Cambridge University Press: Cambridge, UK.

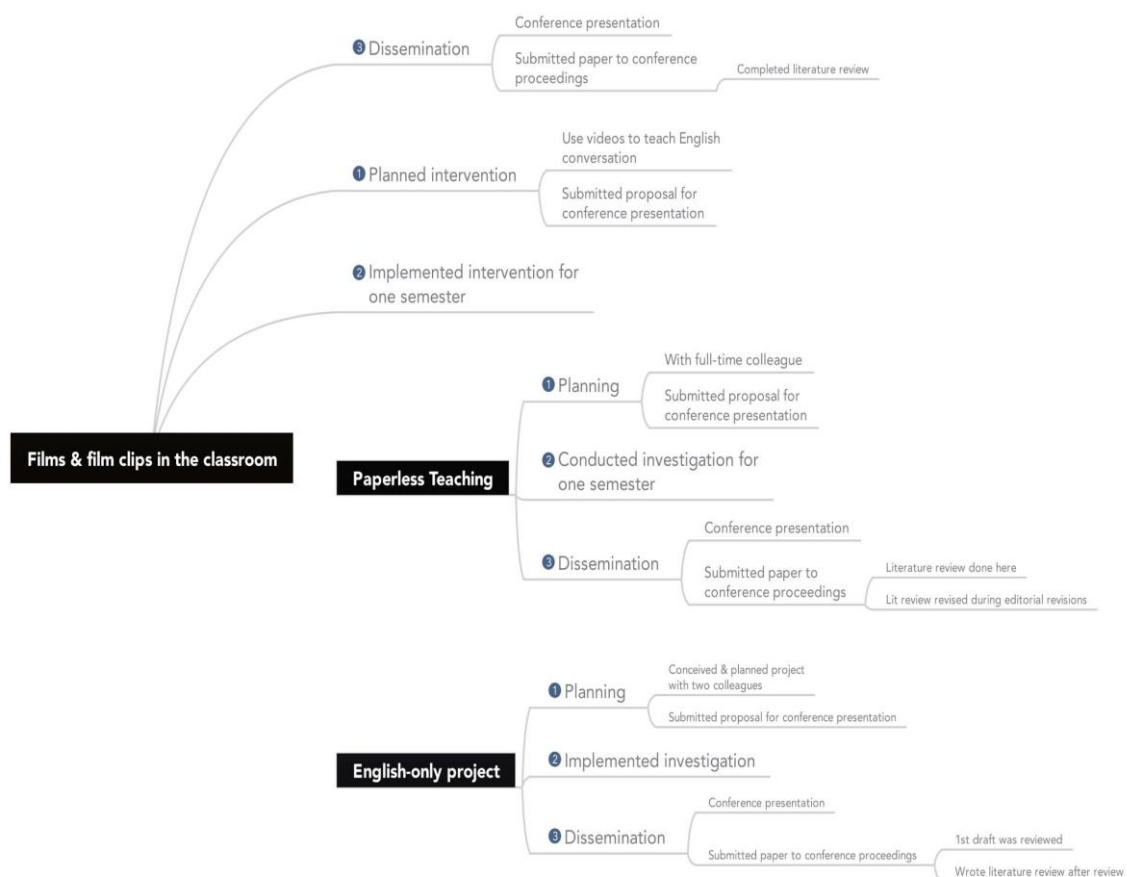
Appendix A: Map of Gerald's AR project



Appendix B: Map of Theron's writing investigation projects



Appendix C: Map of Kathy's three classroom-based research projects



6 The Influence of Experience as Non-Native Speakers on Beliefs of Native Speakers

Koichi Shimahara

Tokyo University of Foreign Studies, Japan

skoichi1221@gmail.com

Introduction

With the number of non-native speakers of Japanese increasing in Japan, some researchers have claimed that Japanese as a lingua franca (JLF), or Japanese as spoken by non-native speakers of Japanese, has different norms from Japanese as used by natives (Okazaki 2000, etc.). Therefore, they argue, it is necessary for native Japanese speakers to be aware that the norms of Japanese language usage in conversations between native speakers differ from those in conversations between native and non-native speakers. In theory, the acceptance of JLF in conversations means the Japanese language is no longer inherent to native Japanese speakers, and the distinction between the communication of native versus non-native speakers breaks down. In that case, there would be no power balance issues between them in dialogue, issues which have often pointed out in past research. In addition, using JLF could lead to equal participation from and an equal relationship between both speakers in a dialogue, regardless of their linguistic background.

The Ministry of Internal Affairs and Communications has proposed such an equal relationship between native and non-native speakers to be ideal for the coming multicultural society in Japan. If JLF becomes the most common language used in conversations between native and non-native speakers in the coming multicultural society, both types of speakers might be required to learn it. However, we currently do not have a clear idea of how native speakers acquire JLF. A study conducted by Yanagida (2011) analyzed how native speakers' linguistic features change as they interact with non-native speakers over a long period. However, their mindsets—i.e., what they pay attention to during conversations—have rarely been the focus of empirical research. If JLF comes to be used in Japanese society and learned by native speakers, not only should their linguistic features be analyzed, but also their mindsets, with special focus on how they change.

In previous research, native speakers' mindsets were analyzed mostly in follow-up interviews, which are interviews conducted after conversations have taken place (Tsuda 2003, etc.). Through follow-up interviews, each utterance expressed in actual conversations can be scrutinized with very specific questions: for example, "Why did you talk like this?" and "What made you say this sentence?". However, what can be analyzed by follow-up interviews is only what participants pay attention to during the conversations; how their mindsets change cannot be determined.

This study aims to reveal the mindsets of native Japanese speakers when interacting with non-native Japanese speakers and how their mindsets change through such interactions. To achieve this, semi-structured interviews were conducted with seven native speakers where they were questioned about their past experiences with non-native speakers. The application of a semi-structured interview format was expected to be able to reveal how their mindsets change and what factors influence these changes.

The influence of experience as a non-native speaker of a foreign language shall especially be focused on in this paper. In other words, this paper will illustrate how the experiences of Japanese native speakers as non-native speakers of other languages can influence their mindsets. With the impact of globalization in Japan, conversations between Japanese citizens and non-citizens could conceivably be in any language. Therefore in the coming multicultural society, it is expected that native Japanese speakers will have more and more chances to become non-native speakers of other languages. Therefore, the results of this paper could contribute to the discussion of how native speakers and non-native speakers can mutually learn and grow in a society.

Methodology

Semi-structured interviews were conducted with seven native Japanese speakers to reveal how their mindsets change due to interactions with non-native Japanese speakers. Because one's mindset during a conversation is at times very personal, establishing a good rapport with interviewees is an important aspect of data collection when using this methodology. In order to facilitate the development of good rapport, the participants were selected through snowball sampling. Under this process, each interviewee was asked to introduce another candidate interviewee who had previously interacted more in dialogue with non-native speakers. Milroy and Gordon (2003) described this sampling method as advantageous for allowing "a new subject

not as a complete outsider but more in the role of a ‘friend of a friend’” (p. 32) to become a part of the study.

Basic information on the seven interviewees is presented below.

Name	Age	Sex	Main Experiences with Non-Native Speakers	Main Experiences as a Non-Native English Speaker
Iwata	21	Female	Friends at university	1 year study abroad in Australia, International Activity Circle at university
Itai	22	Female	Friends at university	1 month study abroad in United States, International activity circle at university
Ichimaru	27	Female	Friends at university	1 year study abroad in United States, Tutor experience at international dormitory
Komoto	22	Male	Friends at university	No special experience
Sasaki	22	Female	Friends at university, Japanese language teaching training	No special experience
Kaji	25	Male	Japanese language teaching training	No special experience
Saika	22	Female	Japanese language teaching training	No special experience

Table 1: Basic information of interviewees

As mentioned above, this paper especially focuses on the influence of native speakers’ experiences as non-native speakers. Three of the seven interviewees were thought to have much more experience as non-native English speakers than the other four. Therefore, in the latter part of this paper, the mindsets of those three are compared with the mindsets of the others.

In the semi-structured interviews, questions similar to the ones below were used to reveal participants’ mindsets during conversations.

- Do you remember when you talked with non-native speakers of Japanese for the first time in your life? Please describe the situation and what you thought during or after the conversation.
- In your elementary school (junior high school, high school, university, etc.) days, did you have a chance to talk with non-native Japanese

speakers, regardless of which language you used? Please describe each situation in detail and what you thought or noticed during each conversation.

- Do you have any memorable moments from interactions with non-native speakers? Please describe them.
- Do you have any experience as a non-native speaker of any language? If you have, please describe each interaction, each speaker, and what you noticed during each interaction.
- What do you usually pay attention to during conversations with non-native Japanese speakers? Do you think your mindset changes according to the situation and the speakers? If so, please mention a few examples from your experience.
- What kind of difficulties do you think non-native speakers of Japanese have during conversations with native speakers? Have you taken specific actions to reduce those difficulties for them?

Although this paper focuses on the influence of the interviewees' experience as non-native speakers of English, other factors were also considered in interviews, such as their experiences of interacting with non-native speakers of Japanese and of receiving Japanese courses as a second language. Semi-structured interviews were conducted with each interviewee from May to July in 2014. Each interview lasted sixty minutes and was conducted twice with each interviewee in order to establish good rapport. For the analysis, all sound data was transcribed with phonetic information. The transcription was coded and categorized, focusing on what interviewees paid attention to during conversations and the factors thought to have influenced their mindsets.

Although its ability to reveal changes in interviewees' mindsets could be stated to be an advantage of the semi-structured interview methodology, some possible disadvantages should also be mentioned here. First, although the interviewer gave participants enough time to remember each conversation and to answer each question, it is still possible for them to have not accurately remembered their previous mindsets. Second, because only what can be put into words can be expressed in interviews, there could have been other, untold elements that influenced their mindsets.

In the next section, the interviewees' mindsets will be sorted into three categories, and presented with actual narratives resulting from the analysis.

Findings

As a result of qualitative analysis, the mindsets expressed by participants in the interviews were categorized into three groups: “taking care to speak intelligibly”, “wanting to improve their partners’ Japanese language skills”, and “wanting to continue the conversation”. After looking at these three categories with the context of actual narratives, we shall focus on how the mindsets of native speakers with much experience as non-native speakers differ from those of native speakers with little such experience.

Taking care to speak intelligibly

All of the native speakers reported they take care to speak intelligibly when they talk with non-native speakers. Related linguistic and para-linguistic behaviours mentioned in interviews were to speak slowly, to speak with short and simple sentences, and to not use difficult vocabulary. Below is a quote from Itai.

1. I speak slowly and clearly: otherwise, they wouldn't understand. (from Itai)

Quote 1 was the most common characteristic of the mindsets of the interviewees, especially in their conversations with non-native speakers with limited Japanese proficiency. However, even in conversations with those with high Japanese proficiency, some native speakers expressed that they still take care to speak intelligibly. Below is another quote from Itai.

2. I don't know why, but I did speak slowly with a German friend. His Japanese was exceptionally good and I'm sure he would have understood what I meant even if I didn't speak slowly. But still, I did speak slowly with him. (from Itai)

Past research on foreigner talk—defined as “the particular register used primarily to address foreigners: i.e, people who do not have full native competence in one’s language” (Ferguson 1981, p.10)—has revealed that non-native speakers of a language, especially those with high proficiency in it, often negatively evaluate foreigner talk by native speakers, including behaviours such as speaking slowly and using basic vocabulary (Sakamoto et al. 1989). Although Itai stated her German friend did not comment about the way she spoke, it is possible that he evaluated it negatively.

As for the changes in the way they speak, most native speakers expressed that they were able to speak slowly and clearly from the beginning of their recollections, even in their elementary school days. However, they also

mentioned that the way in which they choose words and adjust sentences has changed, attributing this to various factors. Quote 3 is from Komoto.

3. I could speak slowly and clearly from the beginning, but I came to be able to choose my words and adjust my sentences through the experience of interactions with non-native speakers.
(from Komoto)

Some native speakers actually reported being able to choose their words and adjust their sentences according to their audience. However, it is thought to be difficult for native speakers to judge which words are easy or difficult for non-native speakers to understand. The next quote was Itai's answer to the question, "How do you judge which words are easy or difficult?"

4. Since they would understand what primary Japanese school kids use, I used those words.
(from Itai)

Some research has revealed differences in which vocabulary is perceived as difficult native versus non-native speakers, and to what extent. What participants report in interviews and what they actually do can be different. Whether or not Itai could actually speak clearly and choose her words and sentences carefully cannot be proven based on self-reporting alone: therefore, actual conversations with her as an interlocutor would need to be analysed.

In addition to interactions with non-native speakers, Japanese language teaching coursework was also mentioned in interviews as a factor influencing speech patterns. The quote below is from Sasaki.

5. I know how I can speak Japanese intelligibly because my major is Japanese Language Teaching. If it were not, my speech would likely be naturally more difficult for non-native speakers to understand.
(from Sasaki)

Here in quote 5, Sasaki reported how she became able to choose her words and sentences according to her listener's Japanese proficiency. Because her major is Japanese Language Teaching, she felt very confident to judge her interlocutors' proficiency in Japanese and to adjust her vocabulary appropriately.

In summary, for the factor "Taking care to speak intelligibly", two factors influencing participants' mindsets and the way they speak were revealed: experience interacting with non-native speakers and experience taking Japanese language teaching coursework.

Wanting to improve their partners' Japanese language skills

Some native speakers reported that they wanted to improve their partners' Japanese language skills. The Japanese language has two main speech styles, polite style and plain style, the difference between which "has traditionally been described as the distinction between formality and informality respectively" (Miyamoto, 2008, p.132). Related linguistic behaviours mentioned in interviews were to correct mistakes actively and to use slang and plain style. A quote below refers to correcting mistakes.

6. I always correct their mistakes because I'm sure every international student wants to improve their Japanese.
(from Iwata)

All native speakers in this study were university students. Therefore, most of the non-native speakers mentioned by native speakers were international students attending university. Because most international students enrolled in Japanese universities are thought to study Japanese, it might be true that they "want to improve their Japanese". However, whether or not they want to learn through daily conversations is questionable. Some non-native speakers might not be willing to recognize and accept their mistakes in daily conversations. Moreover, correcting mistakes is often pointed out to be a negative face violation of the hearer from the perspective of Brown & Levinson's politeness theory (1987). It could be claimed that too many corrections in conversations can lead to interpersonal problems.

Mindsets regarding slang and plain style are depicted by the quotes below from Iwata.

7. I use slang actively because international students wouldn't know it unless they use slang with friends, and that's how I learned English slang.
(from Iwata)

8. I always use plain style with them because I guess they usually use polite style in classrooms or with teachers.
(from Iwata)

In quote 7, the influence of Iwata's experience as a non-native English speaker was explicitly mentioned. Iwata emphasized her experience studying abroad for a year throughout the interview. It seemed that her mindset in conversations with non-native Japanese speakers was based on what she experienced as a non-native English speaker.

In quote 8, Iwata also referred to her partners' Japanese proficiency. There is extensive research on how native versus non-native Japanese speakers use

polite style and plain style in different contexts. Kato (2006) revealed that native and non-native speakers share the norm of using “polite style in conversations of first encounters”. Therefore, if native Japanese speakers use plain style from the beginning, as mentioned in quote 8, non-native speakers could get confused by the discrepancy between what they learned in their Japanese classes and how interlocutors outside the classroom actually express themselves. In other words, it could be claimed that even if Iwata wants to foster a partner’s Japanese proficiency, such behavior could actually work to its detriment with regard to acquiring the sociolinguistic norms of Japanese.

Although their specific mindsets were not put into words, Itai and Ichimaru, who likewise have much experience as non-native speakers of English, expressed that they care about the development of their partners’ language skills. The influence of experience as non-native English speakers will be considered later.

Wanting to continue the conversation

Some native speakers reported their desire to continue conversations with non-native speakers. Related linguistic behaviours mentioned in interviews were overlooking errors, choosing comfortable topics, and asking many questions. With these mindsets, native speakers play the roles of language hosts in conversations with non-native speakers, who accordingly act as language guests. What Komoto expressed in quote 9 is in fact contradictory to what was described by Iwata in quote 6.

9. *I often used to correct their mistakes, but recently not so much. It might make them unwilling to speak.*
(from Komoto)

Komoto reported that he stopped correcting mistakes after he noticed that corrections could make non-native speakers unwilling to speak. He cited the reason for his recent behavior as because he wanted to engage in conversation with non-native speakers as long as possible without causing them stress. From the perspective of JLF, it could be argued that correcting mistakes should be avoided in order for there to be an equal relationship between native speakers and non-native speakers. However, at the same time, there are non-native speakers who are willing to have their mistakes corrected and to improve their language skills. Although the matter should be discussed further, one could argue that the best general principle is for native speakers to be flexible with regard to if and when to correct mistakes.

The next quote below concerns topics Iwata chose in conversations with non-native speakers.

10. *When I talk with beginners, I choose familiar topics and ask what happened, not what they think.* (from Iwata)

Iwata raised some examples of questions which she often asks non-native speakers with low Japanese proficiency, including “What did you do yesterday?” and “What did you eat this morning?”. Such specific questions are thought to be easier to answer than abstract questions, such as “What do you think about the news?”. Similar mindsets were heard from other participants.

As for the desire to continue conversations, no factors influencing participants’ mindsets were explicitly mentioned in the interviews.

The influence of experience as non-native English speakers

Of the factors revealed above, this section focuses on interviewees’ experiences as non-native speakers of English. This factor was only explicitly mentioned in discussions of the category “wanting to improve their partners’ Japanese language skills”. Because three native speakers with considerable experience as non-native speakers studied abroad as international students in their high school days, it is possible they see elements of themselves as students and learners in non-native Japanese speakers. Ichimaru and Iwata showed their envisioned role as a supporter as below.

11. *Every international student is like a big child.* (from Ichimaru)
12. *In order for them to grow up, I do...* (from Iwata)

When asked the reason for her answer in quote 11, Ichimaru expressed that she thinks the experience with her host mother in America has greatly influenced the way she interacts with non-native speakers. Moreover, the three individuals with more experience had clear ideas of how to talk and demonstrated them very confidently in the interviews. They often used confident and assertive phrases, such as “I always do...” and “Unlike other native speakers, I do...”. In addition, their attitudes seemed to be very supportive yet fixed at the same time. They basically assigned non-native speakers of Japanese the role of Japanese learners who are weaker than

native speakers in their conversation skills yet want to learn Japanese nonetheless.

On the other hand, the other participants, who had no experience as non-native speakers of any language, expressed few indisputably clear ideas and lacked confidence when expressing them. In the interviews, they often answered questions with lots of caveats, such as “I’m not sure if it’s OK, but...”, and “This is not what I always do, but...”. However, because they were not confident, they seemed to be more flexible depending on the situation and interlocutor in questions than those interviewees with much experience as non-native speakers. Below are quotes from native Japanese speakers without much experience as non-native speakers of a language.

13. *Only when I’m a tutor do I try to speak intelligibly.* (from Kaji)

14. *If he seems confident with his Japanese, I speak with no adjustment.*
(from Komoto)

Because there are many kinds of non-native speakers with different characteristics, native speakers should be flexible and adapt to each situation with a non-native speaker individually.

Conclusion

The purpose of this paper was to illustrate the mindsets of native speakers when interacting in conversations with non-native speakers, while taking into consideration the influence of their past experience as non-native speakers of other languages. The application of follow-up interviews as the research method allowed changes in those mindsets to be revealed. Mindsets mentioned in interviews were categorized into three groups, “taking care to speak clearly”, “wanting to improve their partners’ Japanese language skills”, and “wanting to continue the conversation”. Presenting actual narratives within each category should be valuable for future research.

However, the possible gap between what participants said they do and what they actually do needs to be considered. Moreover, the result cannot be generalized due to the small sample size. Nonetheless, the results of this research make a contribution to discussions about JLF and how native Japanese speakers can learn it.

References

- Penelope Brown and Stephen C. Levinson (1987). *Politeness: some universals in language usage*. Cambridge University Press: Cambridge, UK.
- Charles A. Ferguson (1981). ‘Foreigner talk’ as the name of a simplified register, In *International Journal of the Sociology of Language*, 28, pp9-18.

- Yoshitaka Kato** (2006). 接触場面における文体・話題の社会言語規範 [Sociolinguistic Norms about Speech Styles and Topic in Contact Situation]. In *Bulletin of International Student Education Center Tokai University*, 26, pp1-17.
- Ministry of Internal Affairs and Communications** (2006). 多文化共生の推進に関する研究会報告書 地域における多文化共生の推進に向けて [Research Group concerning the Promotion of Multicultural Coexistence: Towards the Promotion of Multicultural Coexistence in Local Communities]. Available at: http://www.soumu.go.jp/kokusai/pdf/sonota_b5.pdf [last accessed: 17 December 2015].
- Lesley Milroy and Matthew Gordon** (2003). *Sociolinguistics: Method and Interpretation*. Blackwell Publishing: Oxford, UK.
- Caltabiano, Yuri and Yuri Miyamoto**. (2008). Consequences of Shifting Styles in Japanese: L2 Style-Shifting and L1 Listeners' Attitudes. In Melissa Bowles, Rebecca Foote, Silvia Perpiñán, and Rakesh Bhatt (eds), *Selected Proceedings of the 2007 Second Language Research Forum*. Cascadia Proceedings Project: Somerville, USA, pp131-143.
- Toshio Okazaki** (2000). Japanese Language Education: Towards Multilingual and Multicultural Symbiosis In Mackie, V., Skoutarides, A. and A. Tokita (eds), *Japanese studies: communities, cultures, critiques Volume four: New directions in Japanese linguistics*. Monash Asia Institute: Clayton, Australia, pp11-23.
- Tadashi Sakamoto, Misao Kozuka, Machiko Hasatani, Akie Kosaki, Midori Inaba and Chieko Harada** (1989). 「日本語のフォリナー・トーク」に対する日本語学習者の反応 [Learners' Reaction to Japanese Foreigner Talk], In *Journal of Japanese Language Teaching*, 69, pp121-146.
- Sanae Tsuda** (2003). Interpersonal Functions of the Polite Forms Desu/Masu in Japanese Conversations. In *Intercultural Communication Studies XIX*, 3, pp81-89.

7 Unconscious Attitudes toward L2 Speakers

Ali H. Al-Hoorie
Nottingham University, UK
The English Language Institute, Jubail Industrial
College, Saudi Arabia
hoorie_a@jic.edu.sa

The most common instrument used in attitude and motivation studies is the questionnaire. The researcher asks a group of respondents some questions, and the answers to these questions form the empirical basis for theory development (a similar logic applies to qualitative interviews). The underlying assumption of this approach is that the respondent is able to recognize and articulate the determinants of their behavior. This view is in line with Gordon Allport's (1953, p. 114) view that "simple, conscious report is the whole truth. It can be taken at its face value."

However, motivational psychologists have been increasingly moving away from this view. It is now taken for granted by most researchers that there are more influences than those reportable in questionnaires, and more than meets the eye. If this is the case, then researchers need tools that can tap into these unconscious (or implicit) influences. One of the most commonly used tools is the Implicit Association Test.

What is the Implicit Association Test?

The Implicit Association Test (IAT, Greenwald, McGhee, & Schwartz, 1998) is a computerized reaction-time measure. Participants classify a series of stimuli either to the right or left of the screen depending on their semantic meaning. As an illustration, imagine that a study was designed to measure attitudes toward English and Arabic. In this case, words like *Pleasant* and *Unpleasant* are paired with each of *English* and *Arabic*. In the first condition, *Arabic* may be paired with *Pleasant* and *English* with *Unpleasant* (see Figure 1). In the second condition, the pairing is reversed so that *Arabic* is now paired with *Unpleasant* while *English* with *Pleasant*. The score is calculated by comparing the reaction time in each condition. Most Arabic speakers find it easier to perform the task when *Arabic* is paired with *Pleasant*, whereas most English people find the opposite easier.

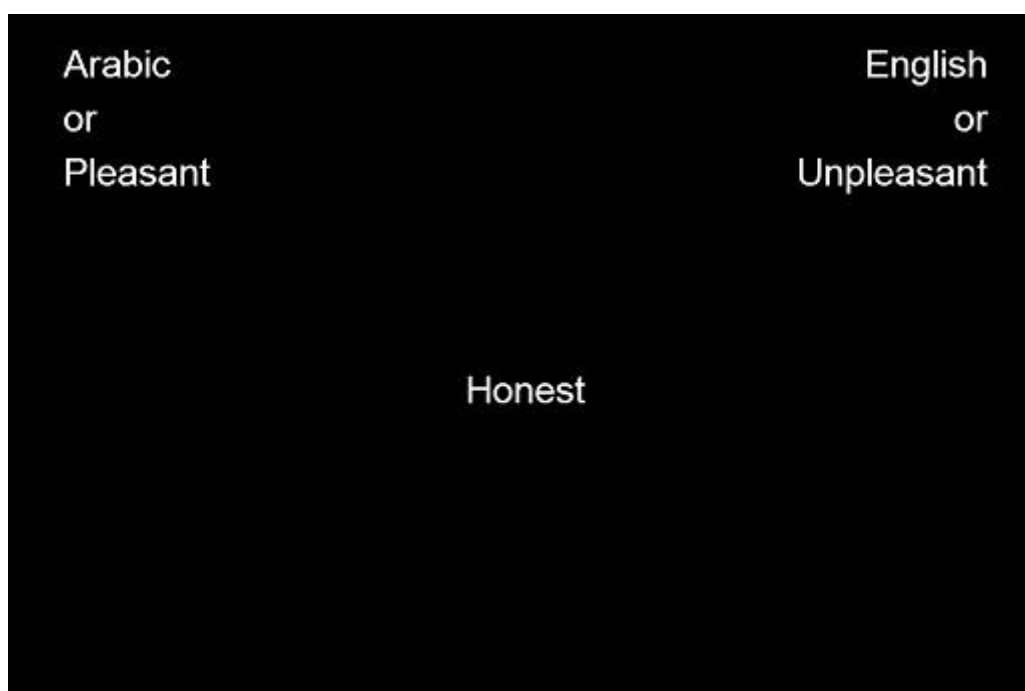


Figure 1: A trial of the IAT. The correct answer here would be the left pair because the stimulus *Honest* belongs to *Pleasant*.

Just as its name suggests, the IAT measures an ‘association’ between two words. It is ‘implicit’ because most people are unaware of these associations and usually find their own results surprising. Whether these associations also qualify to be considered an attitude has been a subject of debate. However, an increasing number of empirical studies are showing that implicit associations indeed predict meaningful outcomes and behaviors that are consistent with the claim that these associations are attitudes (e.g., see Rudman, 2008, for a review).

More specifically, the most interesting findings come from tests of known-groups validity and predictive validity. As for known-groups validity, research shows that the IAT is capable of correctly distinguishing among members of different groups in accordance with our a priori knowledge of them, such as reliably determining the participant’s gender, nationality, and even affiliation to a group artificially created in the laboratory (for a review, see Lane, Banaji, Nosek, & Greenwald, 2007). As for predictive validity, a range of criterion variables—including behavioral, judgmental, and physiological measures—have been successfully predicted by the IAT in various areas including consumer references, political preferences, personality traits, sexual orientations, and close relationships.

In one meta-analysis, the predictive validity of the IAT averaged $r = .274$, and the IAT and self-report measures displayed incremental validity, in that each predicted criterion variance over and above that predicted by the other measure alone (Greenwald et al., 2009). In another meta-analysis (Oswald et al., 2013), the average predictive validity was lower ($r = .15$) in the area of racial discrimination. However, the latter meta-analysis was criticized for including correlations that were not theoretically expected, which suppressed the resulting correlation (Greenwald, Nosek, & Banaji, 2015).

As an illustration of this type of research, consider a study by Glaser and Knowles (2008). The researchers examined the relationship between the implicit stereotype associating Blacks with guns and the tendency to shoot Blacks faster than Whites in a computer simulation game (known as the ‘Shooter Bias’). The participants viewed pictures of Blacks or Whites holding either a weapon or a benign object. Although the task was to shoot individuals holding a weapon regardless of their skin color, participants with a stereotype associating Blacks with guns (especially those with low motivation to control their prejudice) misidentified Blacks as holding guns more frequently.

Implicit attitudes in second language learning

In the remainder of this chapter, I will give a brief summary of two initial studies I conducted using the IAT in the area of second language (L2) learning. In the first study, Arab learners of English ($N = 365$) who were studying in the UK performed the IAT and completed a questionnaire. The results showed that learners with explicit–implicit congruence (i.e., with favorable attitudes toward English both explicitly and implicitly) were significantly more open to the L2 group than were learners with explicit–implicit incongruence (i.e., favorable at the explicit level only). These results showed that the IAT scores behave in a theoretically meaningful way.

The first study did not include a measure of actual L2 achievement because the participants came from various universities and language institutes in the UK. Therefore, no standardized measure of their achievement was available. However, the implicit test would become much more useful if it can also predict who would (not) be successful in the L2. The second study therefore included a measure of L2 achievement in order to test the hypothesis that implicit attitudes are also associated with L2 achievement.

The second study involved learners ($N = 311$) studying English at one institution in Saudi Arabia. The participants performed the Single-Target

Implicit Association Test (ST-IAT, Wigboldus, Holland, & van Knippenberg, 2005), which is a variation of the IAT that can use one category only (e.g., *English* without having to compare it with *Arabic*). In addition to successfully replicating the first study, the results of the second study showed that learners who scored higher in the implicit test did achieve higher in the L2 course.

Conclusion

Going back to Gordon Allport's above-mentioned claim that simple self-report is 'the whole truth', the results of these two studies demonstrate that self-report may not constitute the whole truth after all. This is because expressing positive attitudes toward at the explicit level does not always reflect positive attitudes at the implicit level as well. Researchers should therefore examine implicit attitudes in order to find out whether they are in harmony or in conflict with explicit attitudes, and this cannot be done by self-report measures alone.

When it comes to L2 learning, the results presented above show that an implicit dimension exists. They also demonstrate that explicit-implicit harmony and conflict are related to both attitudes and achievement in L2 learning. These findings are encouraging, as they open up a whole new line of inquiry with ample research opportunities for L2 motivation researchers.

References

- Gordon W. Allport (1953).** The trend in motivational theory. In *American Journal of Orthopsychiatry* 23:1, pp107-119.
- Jack Glaser & Eric D. Knowles (2008).** Implicit motivation to control prejudice. In *Journal of Experimental Social Psychology* 44, pp164-172.
- Anthony G. Greenwald, Debbie E. McGhee & Jordan L. K. Schwartz (1998).** Measuring individual differences in implicit cognition: The Implicit Association Test. In *Journal of Personality and Social Psychology* 74:6, pp1464-1480.
- Anthony G. Greenwald, Brian A. Nosek & Mahzarin R. Banaji (2015).** Statistically small effects of the Implicit Association Test can have societally large effects. In *Journal of Personality and Social Psychology* 108:4, pp553-561.
- Anthony G. Greenwald, T. Andrew Poehlman, Eric Luis Uhlmann & Mahzarin R. Banaji (2009).** Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. In *Journal of Personality and Social Psychology* 97:1, pp17-41.
- Kristin A. Lane, Mahzarin R. Banaji, Brian A. Nosek & Anthony G. Greenwald (2007).** Understanding and using the Implicit Association Test: IV. What we know (so far) about the method. In Bernd Wittenbrink & Norbert Schwarz (eds), *Implicit Measures of Attitudes*. Guilford Press, New York, pp59-102.

- Frederick L. Oswald, Gregory Mitchell, Hart Blanton, James Jaccard & Philip E. Tetlock (2013).** Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. In *Journal of Personality and Social Psychology* 105:2, pp171–192.
- Laurie A. Rudman (2008).** The validity of the Implicit Association Test is a scientific certainty. In *Industrial and Organizational Psychology* 1:4, pp426–429.
- Daniël H. J. Wigboldus., Rob W. Holland & Ad van Knippenberg (2005).** *Single Target Implicit Associations*. Unpublished manuscript.

8 (Re)thinking the Interface Hypothesis and its Implications for Language Teaching

Joana Teixeira

CLUNL, Universidade NOVA de Lisboa

joana.v.teixeira@gmail.com

Introduction²

The interface between syntax and other domains has recently become a key area of interest in Generative Second Language Acquisition (GenSLA). Much of the research on linguistic interfaces has been influenced by the Interface Hypothesis (IH), which was proposed by Sorace & Filiaci (2006) in an attempt to explain the non-target-like behaviour found at very advanced stages of second language (L2) acquisition. Originally, the IH claimed that narrow syntactic properties are acquirable, whereas properties at the interface between syntax and other domains may not be fully acquirable (cf. Sorace & Filiaci, 2006). In its most recent version, the IH specifies that properties at grammar-external interfaces (i.e. interfaces which connect the grammar to external domains) are more likely to be a source of problems in end-state grammars than properties involving grammar-internal interfaces (i.e. interfaces which link different modules within the grammar) (cf. Sorace & Serratrice, 2009; Tsimpli & Sorace, 2006; Sorace, 2011). The IH is, therefore, an account of non-target patterns at the level of ultimate attainment.

While some studies have supported the IH's predictions (e.g. Belletti & Leonini, 2004; Belletti, Bennati & Sorace, 2007; Bohnacker & Rosén, 2007; Lozano, 2006; Madeira, Xavier & Crispim, 2009; Sorace & Filiaci, 2006; Wilson, 2009; among others), others have produced results which disconfirm them (e.g. Domínguez & Arche, 2014; Donaldson, 2011, 2012; Ivanov, 2012; Slabakova, Rothman & Kempchinsky, 2011; Slabakova, Kempchinsky & Rothman, 2012; Slabakova, 2013; among others). Despite their relevance, some of the latter studies have been overlooked in the IH-related literature (e.g. Sorace, 2011; White, 2009, 2011a). To date, the research on or inspired by the IH has exclusively examined questions pertaining to linguistic theory.

² The present article is based upon work funded by Fundação para a Ciência e a Tecnologia (grant no. SFRH/BD52263/2013).

No attempts have been made to explore the pedagogical implications of the findings on L2 acquisition at the interfaces. Therefore, it remains unclear if and how these findings might be useful for L2 pedagogy.

With a view to addressing these issues and contributing to advance the field of Applied GenSLA – a new strand within applied linguistics which explores the interface between GenSLA and L2 pedagogy (cf. Whong, Gil & Marsden, 2013) –, the present paper seeks to (i) discuss the IH in the light of recent findings on L2 acquisition at the interfaces, including those which have been overlooked in previous review articles, and (ii) propose potential pedagogical implications of GenSLA findings on linguistic interfaces. The paper is organised as follows: Section 2 summarises recent research findings on L2 acquisition at grammar-internal and grammar-external interfaces. In section 3, I examine the consequences of those findings for the IH. Section 4 explores how research findings on the IH can inform L2 teaching. Finally, in the last section, the main conclusions of the paper are presented.

Research on L2 acquisition at the interfaces: An overview

In recent years, the focus of research in GenSLA has been on structures at linguistic interfaces, i.e. structures that are subject to conditions of varying nature, requiring the integration of knowledge from syntax and other domains, such as grammar-internal modules, like semantics and morphology, or grammar-external domains, like discourse and pragmatics³. Research findings have consistently shown that some interfaces tend to be problematic to L2 learners/speakers (L2ers), while others do not cause significant problems to them. Nevertheless, the status that each interface has in interlanguage grammars is still a matter of ongoing debate and research. In order to present a balanced review of the evidence for and against the IH, this section will briefly summarise the most significant findings of recent work on L2 acquisition at the interfaces, considering both internal and external interfaces.

Grammar-internal interfaces

For the IH as currently formulated, the dissociation between grammar-internal and grammar-external interfaces is vital to explain well-attested difficulties of L2ers such as optionality and fossilisation. According to this

³ Underlying this line of inquiry is the assumption that the language faculty is composed of various modules which interact in specific ways (cf. Rothman & Slabakova, 2011, and White, 2009, for further details).

hypothesis, the interface structures that are especially prone to optionality and fossilisation in end-state grammars are those which involve external interfaces. Thus, internal interfaces are not expected to pose widespread problems at highly advanced levels of proficiency.

Confirming the latter prediction, the research carried out over the past decade has consistently shown that structures at internal interfaces are not a typical locus of fossilisation and optionality in end-state grammars. Only in exceptional cases do these structures create difficulties to L2ers at a near-native level (e.g. Hopp, 2007, 2009). Although internal interface structures do not cause significant problems at the level of ultimate attainment, they are not always easy to acquire. In fact, while some internal interfaces, like the syntax-semantics interface, do not typically pose significant developmental challenges to L2ers (e.g. Dekydtspotter, Sprouse & Anderson, 1997; Dekydtspotter, Sprouse & Swanson, 2001; Dekydtspotter, Sprouse & Thyre, 2001; Madeira, Xavier & Crispim, 2010; Slabakova, 2003, 2009; among many others)⁴, others, most notably the syntax-morphology interface, do create problems. For example, L2ers tend to have difficulties in realizing bound inflectional morphology (e.g. tense marking), as well as free function words (e.g. articles) in a target-like manner, even in non-elementary levels (cf. Lardiere, 1998, 2007; White, 2003, 2009; among many others). Due to these well-attested difficulties, Slabakova (2009: 280) proposes that functional morphology is “the bottleneck of L2 acquisition”.

Crucially, the results of GenSLA research on internal interfaces are fully compatible with the IH, because this hypothesis merely claims that properties at external interfaces are the primary locus of permanent optionality and fossilisation in end-state grammars, which does not necessarily mean that near-native speakers (near-NSs) will perform target-like with respect to all properties at internal interfaces or that no problems will occur at developmental stages regarding these interfaces.

Grammar-external interfaces

As the interfaces that are predicted to be vulnerable to optionality and fossilisation in end-state grammars are the so-called grammar-external interfaces, it is this type of interface that has been most investigated in the IH-related literature. Research on external interfaces has mainly analysed the

⁴ Note that there are a few attested cases of problems in the acquisition of properties at the syntax-semantics interface (cf. Bruhn de Garavito & Valenzuela, 2006; Guijarro-Fuentes & Marinis, 2007).

interaction between syntax and discourse/pragmatics, a domain typically treated as being outside the grammar (White, 2009). The best researched property at this interface is the distribution of null and overt subjects in null subject languages (NSLs). Given that it was on the basis of evidence from studies on null and overt subjects that the IH was formulated, the focus of the present review of the literature on external interfaces will be on these structures. Research findings on the acquisition of other structures relevant to the syntax-discourse/pragmatics interface will also be discussed, though in less detail.

The availability of null subjects is a purely syntactic property that distinguishes NSLs, like European Portuguese (EP), Italian and Spanish, from non-null subject languages (NNSLs), like English. Even though the syntax of NSLs can generate both null and overt subjects, there is no true optionality between the two forms in these languages, as their distribution is determined by discourse/pragmatic conditions. More specifically, null subjects typically occur when there is topic maintenance, whereas overt subjects are produced when there is a change of topic. NSLs also display different preferences in the interpretation of null and overt pronominal subjects. According to the Position of Antecedent Hypothesis (Carminati, 2002), null subject pronouns are assigned to the antecedent in the canonical subject position (which tends to be interpreted as topic), while overt subject pronouns are generally assigned to antecedents in other positions, as the EP sentences in (1) illustrate⁵. Hence, it can be concluded that the production and interpretation of null and overt subjects involves the interaction between syntax and discourse/pragmatics.

- (1) a. A Maria_i disse à Joana_j que Ø_i vai ganhar o jogo.
(Lit. The Maria_i told to the Joana_j that Ø_i goes win the game)
b. A Maria_i disse à Joana_j que ela_j vai ganhar o jogo.
(Lit. The Maria_i told to the Joana_j that she_j goes win the game)

One of the first studies to investigate the L2 acquisition of the division of labour between null and overt subjects in NSLs was carried out by Sorace & Filiaci (2006). These researchers examined the interpretation of pronominal subjects in intrasentential contexts by a group of native speakers (NSs) of English who had attained a near-native level in Italian. Their results revealed that the near-NSs and the control group of NSs of Italian had similar patterns of preferences with respect to the interpretation of null subjects. In contrast,

⁵ In EP, these preferences can be cancelled by certain factors, such as the occurrence of the subjunctive mood.

near-NSs behaved differently from NSs in the interpretation of overt subject pronouns. They had a higher preference for interpreting the subject of the matrix clause as an antecedent of overt subject pronouns, even when the context did not favour that interpretation. Significantly, Belletti, Bennati & Sorace's (2007) subsequent work on the production and interpretation of null and overt subjects in L2 Italian produced similar results.

These findings are regarded as evidence that near-NSs have acquired the syntax of null subjects and exhibit divergent behaviour only at the interfaces. On the one hand, the near-NSs' target-like performance with respect to the production and interpretation of null subjects suggests that they have successfully reset the null subject parameter to the Italian value. On the other hand, the fact that near-NSs produce overt subjects in the absence of a shift of topic, and interpret overt subject pronouns as correferential with a pragmatically inappropriate subject antecedent indicates that it is at the interface between syntax and discourse/pragmatics that they manifest persistent difficulties. Therefore, these results lend support to the IH.

The asymmetry in the interpretation and production of null and overt pronominal subjects reported in these studies on L2 Italian-L1 English has also been found with other language combinations (e.g. Madeira, Xavier & Crispim, 2009; Tsimpli & Sorace, 2006, among others) and in other areas of language development, such as L1 attrition (e.g. Tsimpli, Sorace, Heycock & Filiaci, 2004) and L1 bilingual acquisition (e.g. Serratrice, Sorace & Paoli, 2004; Sorace, Serratrice, Filiaci & Baldo, 2009). The latter fact suggests that, as Sorace (2011) supports, the IH applies not only to the very advanced stages of L2 acquisition, but also to simultaneous bilingualism and L1 attrition.

To explain the difficulties which are found at external interfaces in these different areas of language development, at present, the proponents of the IH appeal to processing-based factors (cf. Sorace, 2011)⁶. Drawing on the concept of inhibitory control which has long been considered a crucial factor in bilingual performance in the field of Psycholinguistics (e.g. Green, 1986, 1998), the proponents of the IH claim that all speakers who have more than one grammar represented in the mind (i.e. L2ers, L1 attriters and simultaneous bilinguals) need to inhibit one language when using the other. According to them, the inhibition of the language not in use is a costly operation which may take attentional resources away from other linguistic

⁶ For a summary of the previous accounts developed within the IH framework, see Sorace, 2011.

tasks, leading to problems in processes that impose high demands on processing resources, like the ones which involve the integration of syntactic and grammar-external information (cf. Sorace, 2011). In other words, the proponents of the IH do not claim that speakers who have more than one grammar represented in the mind are unable to attain native-like representations for structures at grammar-external interfaces; rather, they propose that, even if these bilingual speakers' underlying linguistic representations are native-like, some level of residual, but permanent, divergence is expected in performance as a by-product of bilingual processing. Since the problems at external interfaces are explained as cognitive consequences of bilingualism, it is predicted that some level of divergence will arise with respect to these interfaces, independently of L1–L2 pairings.⁷

This prediction is not confirmed by all studies on the distribution of null and overt subjects. In a study on the interpretative preferences of elementary and advanced L2ers of EP concerning pronominal subjects, Madeira, Xavier & Crispim (2012), for example, show that Italian NSs exhibit native-like antecedent assignment preferences for null and overt subjects in EP, whereas Chinese L2ers of EP behave like the near-NSs of Italian investigated by Sorace and colleagues (i.e. they exhibit problems in the interpretation of overt pronominal subjects, but no difficulties regarding null subjects). These results suggest contra the IH that the integration of syntactic and discourse information does not necessarily cause difficulties to L2ers. In line with these findings, Rothman (2009) demonstrates that advanced L2ers of Spanish who are NSs of English generally display native-like knowledge of the distribution of pronominal subjects in the target language. In addition, a series of studies by Zhao (2011, 2012, 2014) on the interpretation of null and overt subjects in L2 Chinese, a language where the distribution of subjects is different from that found in Romance NSLs like Italian and EP (cf. Zhao, 2011, 2012, 2014 for details), shows that L2ers successfully acquire properties at the syntax-discourse interface, which is not consistent with the IH's predictions. Taken together, these findings on the distribution of null and overt subjects indicate that it may be a little premature to generalise about the impossibility of complete convergence at the syntax-discourse/pragmatics interface.⁸

⁷ Note that the occurrence of L1 influence effects in addition to processing effects is not ruled by this hypothesis.

⁸ Note that the latest formulations of the IH make claims that are best tested via online measures, but most of the studies which are not in line with this hypothesis use offline measures. These studies are, however, comparable to the ones by Sorace & Filiaci (2006)

Like the studies on the L2 acquisition of overt and null subjects, the research which focuses on other external interface phenomena has produced mixed results with regard to the validity of the IH. In fact, while some studies confirm the IH (e.g. Hopp, 2004, who examined scrambling in L2 German; Belletti & Leonini, 2004, Belletti, Bennati & Sorace, 2007, and Lozano, 2006, who investigated subject-verb inversion in NSLs; Bohnacker & Rosén, 2007, who studied the prefield in L2 German; Wilson, 2009, who investigated personal and demonstrative pronouns in German anaphoric dependencies), others disconfirm its predictions by demonstrating that L2ers display native-like knowledge of structures at the syntax-discourse/pragmatics interface, such as *c'est* clefts and *avoir* clefts in L2 French (Donaldson, 2012), left and right-dislocation in L2 French (Donaldson, 2011ab), clitic-doubling in L2 Bulgarian (Ivanov, 2009, 2012), focus fronting (FF) in L2 English (Slabakova, 2013), and FF, clitic left dislocation (CLLD) and clitic right dislocation in L2 Spanish (Domínguez & Arche, 2014; Slabakova, Rothman, Leal Mendez, Campos & Kempchinsky, 2011; Slabakova, Rothman & Kempchinsky, 2011; Slabakova, Kempchinsky & Rothman, 2012; Slabakova, 2013).

Crucially, the studies which demonstrate that target L2 ultimate attainment is possible at the syntax-discourse/pragmatics interface simultaneously show that knowledge of how discourse/pragmatics factors constrain syntax tends to emerge quite late in the course of L2 development and generally after the acquisition of purely syntactic properties (e.g. Ivanov, 2009, 2012; Rothman, 2009; Slabakova, Rothman, Leal Mendez, Campos & Kempchinsky, 2011; Slabakova, Kempchinsky & Rothman, 2012). It appears that discourse/pragmatic conditions can only be acquired with relative ease when they are the same in the L1 and the L2 (e.g. Lozano & Mendikoetxea, 2010; Madeira, Xavier & Crispim, 2012). Thus, the existing data on L2 acquisition at grammar-external interfaces generally supports the following prediction which the IH indirectly makes about L2 development: grammar-external interfaces present developmental challenges for L2 acquisition.

To sum up, it seems that linguistic interfaces behave in a less homogeneous way than assumed by the IH. From the results of the studies conducted to date, three conclusions can be drawn. First, grammar-internal interfaces are not generally a problem area at the level of ultimate attainment (even though they may be subject to developmental delays), which is in line with the IH's predictions. Second, properties at the syntax-discourse/pragmatics interface

and Belletti, Bennati & Sorace (2007), because they also used offline measures. Hence, differences in results cannot be merely attributed to methodological differences.

generally pose developmental problems to L2ers, as the IH indirectly predicts. Third, properties at the syntax-discourse/pragmatics interface are potentially, but not necessarily, problematic in ultimate attainment, which is not consistent with the IH, as currently formulated, in that it predicts that the integration of syntactic information with grammar-external information imposes an extra processing load that should ultimately lead to some level of divergence in performance at the end-state of L2 acquisition.

Implications for the IH

Given that the IH as currently formulated is not entirely supported by GenSLA research findings on linguistic interfaces, it needs to be reformulated to have greater descriptive and explanatory power. In a new formulation of the IH, two key questions must be (re-)examined: (i) Is the IH only concerned with ultimate attainment or does it also apply to developmental stages? (ii) Under what circumstances are structures at the syntax-discourse/pragmatics interface prone to permanent fossilisation?

The first question has generated some debate in the literature. On the one hand, the proponents of the IH claim that this hypothesis is an account of non-target patterns at the level of ultimate attainment, which is not intended to apply to developmental stages (Sorace, 2011), although they concede that it may indirectly make predictions about them (Sorace, 2012). On the other hand, some researchers, most notably Lardiere (2011) and White (2011b), argue that the IH does make predictions about L2ers at a non-near-native level that can and should be tested. Indeed, if external interface properties are residually difficult for the most proficient near-NSs, it is reasonable to expect L2ers at lower stages of acquisition to have similar or even more pronounced problems concerning those properties. As the studies carried out to date indicate that external interfaces do pose developmental challenges, in the future, the IH should be extended to developmental stages.

Unlike question (i), question (ii) has no straightforward answer. At present, it is still not well understood why some external interface structures fossilise while others do not. Nevertheless, recent work by Slabakova (2013) and Domínguez & Arche (2014) offers some clues as to why this happens. In her study on the acquisition of CLLD and FF in L1 English – L2 Spanish and of topicalization and FF in L1 Spanish – L2 English, Slabakova (2013) tentatively proposes that when the constructions at the syntax-discourse interface are the same in the L1 and the L2, as in the case of FF, they do not pose difficulties to highly advanced L2ers. If they are different in the two languages and transfer is misleading, then construction frequency in the

input becomes a decisive factor in determining the success of the process of L2 acquisition: frequent constructions will be successfully acquired, but infrequent ones will not. Domínguez & Arche's (2014) work on the acquisition of CLLD and S(ubject) V(erb) – VS contrasts in L2 Spanish further indicates that it is not just the frequency of the target construction that matters, the lack of transparency in the evidence available in the input may also be a source of indeterminacy in non-native grammars. Thus, it appears that what may ultimately determine whether or not an external interface structure will fossilise in end-state grammars is a combination of L1 and input factors. However, further research is needed.

In conclusion, the research that has been conducted over the past decade in the field of GenSLA clearly indicates that the IH needs to be reformulated, but its findings do not point the exact direction in which the hypothesis should go to become descriptively and explanatorily adequate. Researchers still have much work to do to develop a version of the IH offering an articulated and robust account of how different factors lead to (non-) convergent behaviour at external interfaces. In this respect, the investigation of the role of L1 and input factors and of how they interact with processing factors both at developmental and final stages of L2 acquisition may prove a fruitful line of inquiry for future IH-related research.

Implications for L2 teaching

Even though there are still some questions on L2 acquisition at the interfaces which remain unanswered in the literature, what is currently known about this process has significant implications for teaching practice that can and should be examined. Nevertheless, to date no attempt has been made to explore the relevance and potential applications of GenSLA findings on interfaces to L2 classrooms. As a result, interface structures tend to be overlooked by L2 teachers, material designers and specialists in L2 pedagogy. Yet, given that GenSLA research has consistently shown that at least some interface phenomena are problematic for L2ers, these phenomena should deserve attention in L2 classrooms. With a view to bridging the gap that currently exists between GenSLA and L2 pedagogy, in this section, I will show how research findings on the IH can inform L2 teaching.

At present, the only interface structures that tend to receive attention in L2 classrooms are those which involve the syntax-morphology interface. Nonetheless, this is not the only problematic interface for L2ers. In line with the IH, GenSLA research indicates that the syntax-discourse/pragmatics interface is also difficult to acquire and possibly more vulnerable to

permanent fossilisation than the syntax-morphology interface. For this reason, and assuming that what is difficult to acquire requires attention in the classroom, this external interface should not be ignored in L2 teaching. In particular, when an external interface structure is not similar in the L2ers' L1 and L2, and the input on that structure is ambiguous and/or scarce, some form of explicit teaching may be beneficial. L2 teachers, hence, need to teach not only grammar contents relevant to the syntax-morphology interface, but also constructions at the syntax-discourse/pragmatics interface.

Since the overall findings from Instructed SLA suggest that instruction can have a substantial impact on the rate of acquisition and ultimate attainment, but cannot alter the route of L2 acquisition (see Spada & Lightbown, 2012, for a brief review), the teaching of interface structures must take into account what is known about the development of these structures in interlanguage grammars. As previously noted, GenSLA research findings reveal that structures at the syntax-discourse/pragmatics interface are generally acquired late and, in certain cases, may not even be fully acquired. Furthermore, research findings suggest that, when the discourse/pragmatics factors which constrain syntax are not the same in the L1 and the L2, the syntactic properties of these structures are acquired before their discourse/pragmatic properties. In the light of these facts, it can be hypothesised that, in the cases where discourse/pragmatics conditions are not transferable from the L1, the teaching of structures at the syntax-discourse/pragmatics interface is more likely to be effective if teachers focus on their purely syntactic properties before teaching their discourse properties. As L2ers only seem to be ready to learn the latter properties at intermediate to advanced levels, it is advisable to focus on such properties only at those levels of proficiency. In the cases where discourse/pragmatics conditions are transferable from the L1, no explicit focus on these conditions seems to be needed throughout the process of L2 development.

On the basis of the findings from GenSLA research and of what is currently known about instructed SLA, it is possible to propose the following general guidelines as to how structures at the syntax-discourse/pragmatics interface should be taught:

(i) *L2ers need to be exposed to linguistic input with multiple instances of the target interface structure, and where its syntactic and discourse/pragmatic properties are transparent and unambiguous.* This is because recent research suggests that, when the evidence to which L2ers are exposed is ample and unambiguous, they may overcome their developmental difficulties at the

syntax-discourse/pragmatics interface (e.g. Slabakova, 2013; Domínguez & Arche, 2014).

(ii) *Teachers need to lead L2ers to notice external interface structures in the input and to raise their awareness of the syntactic and discourse/pragmatic rules that govern these structures.* Underlying this proposal is the assumption that noticing is a prerequisite for L2 acquisition (cf. Schmidt, 1990, 1993, 2001), and that L2ers who are aware of the rules underlying a linguistic structure are more likely to notice it when they subsequently encounter it (cf. Ellis, 1994, 2003, among others). Thus, awareness can facilitate and trigger acquisition.

(iii) *Teachers can use the same types of tasks which are used in GenSLA experiments, in order to raise L2ers' awareness of structures at the syntax-discourse/pragmatics interface and/or provide them with opportunities for practice.* The tasks which may be used for teaching purposes include: a) *contextualized preference task*, which is a task that presents L2ers with a linguistic context and requires them to read alternative sentences and decide which of them is more appropriate in that specific context, as in (2); b) *picture verification task*, which is a task where L2ers are asked to decide which picture from a set of pictures with minor differences, like (3), is the most appropriate illustration of the sentence; and c) *translation task*, which is a task where L2ers have to translate sentences from their L1 into the L2, as in (4).

(2) **Contextualized preference task: Sample item on VS/SV orders in Spanish**

- | | |
|----------------|---------------------|
| ¿Quién llegó? | (Lit. who arrived?) |
| a. Juan llegó. | (Lit. Juan arrived) |
| b. Llegó Juan. | (Lit. arrived Juan) |
| c. Both | |

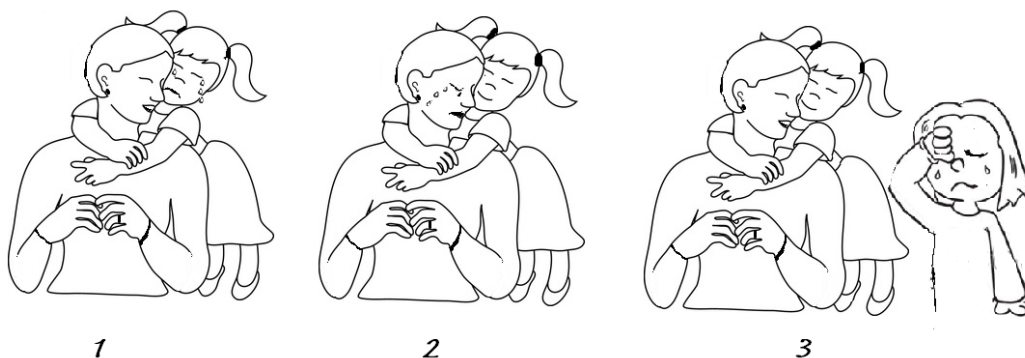
Answer: b⁹

(3) **Picture verification task: Sample item on anaphora resolution in EP**

A filha abraça a mãe enquanto chora.

⁹ In Spanish, when the subject is the narrow focus of the sentence (i.e. when it is the only constituent that introduces new information), VS orders are preferred over SV orders.

(Lit. The daughter hugs the mother while cries)¹⁰



Answer: *I*¹¹

(4) **Translation task: Sample item on CLLD in Spanish**

[Context: We need to buy a toy and sweets for Maria.]

The toy, I will buy before work.

Answer: *El juguete, lo compraré antes de ir al trabajo*¹²

Even if all the guidelines just described are followed, teachers should not expect L2ers to acquire external interface structures straight away, and should also be aware that at least some of those structures may be particularly difficult to acquire.

To sum up, what is currently known about linguistic interfaces suggests that the syntax-discourse/pragmatics interface must receive more attention in L2 classrooms. GenSLA findings on the IH can inform the teaching of structures relevant to this interface in a number of ways: (i) in the selection and sequencing of grammar contents, (ii) in making teachers' expectations of L2ers more realistic and informed, and (iii) in the development of learning tasks. Although the existing data on L2 acquisition at the interfaces allow us to propose general guidelines for teachers, further research on the role of the quantity and quality of the input and on the roles of positive and negative evidence in L2 acquisition at external interfaces will be needed before we

¹⁰ The three images were created on the basis of an image downloaded from http://allcoloringpictures.com/download/Girl_Hugging_Her_Mom1.jpg [retrieved on August 2015].

¹¹ For details on why this is the right answer, see the discussion on null and overt subjects in section 2.

¹² Spanish CLLD and English topicalization are similar with respect to discourse function: in both structures, the left-dislocated phrase is the topic of the sentence. They are, however, different at a syntactic level in that the former construction, but not the latter, requires the left-dislocated phrase to be doubled by a clitic.

can satisfactorily answer the following question: How can one teach properties at external interfaces in an effective way?

Conclusion

In conclusion, over the past decade, the IH has become a very influential generalisation in the field of GenSLA. The research inspired by this hypothesis has produced mixed results, with some studies confirming its predictions and others disconfirming them. Taken together, research findings indicate that external interface structures are problematic for L2ers at developmental stages – at least when they are different in the L1 and L2 – and potentially, but not necessarily, at the level of ultimate attainment. To account for these facts, the IH must be extended to developmental stages and its predictions about ultimate attainment must be reformulated. Instead of claiming that all external interfaces are prone to permanent fossilisation, the IH must clearly state under what circumstances fossilisation is likely to (not) occur. This requires further research.

Even though researchers still have much work ahead of them to piece together the puzzle of L2 acquisition at the interfaces, it is possible to propose implications of what is currently known about this process for L2 teaching practice. As shown in the present article, the findings from GenSLA research on the IH can be useful for L2 teachers in various ways, namely in shaping their expectations of L2ers, in the selection and sequencing of grammar contents and in the development of approaches to teaching interface structures. Thus, contrary to what is often assumed in and outside the field of GenSLA, the research conducted on the IH is relevant not only for L2 acquisition theory, but also for the theory and practice of L2 teaching. Future research on the IH should, therefore, explore two types of interfaces: (i) a linguistic interface – the syntax-discourse/pragmatics interface – and (ii) an interdisciplinary interface – the interface between GenSLA and L2 pedagogy.

References

- Adriana Belletti & Chiara Leonini (2004).** Subject Inversion in L2 Italian. In Susan Foster-Cohen, Michael Sharwood Smith, Antonella Sorace, & Mitsushiko Ota (eds), *EuroSLA Yearbook 4*. John Benjamins: Amsterdam, the Netherlands, pp95-118.
- Adriana Belletti, Elisa Bennati & Antonella Sorace (2007).** Theoretical and Developmental Issues in the Syntax of Subjects: Evidence from near-native Italian. In *Natural Language and Linguistic Theory* 25, pp657-689.
- Ute Bohnacker & Christina Rosén (2007).** Transferring Information-Structural Patterns from Swedish to German. In Alyona Belikova, Luisa Meroni & Mari

- Umeda (eds), *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)*. Cascadilla Proceedings Project: Somerville, USA, pp27-38.
- Joyce Bruhn de Garavito & Elena Valenzuela (2006)**. The Status of *Ser* and *Estar* in Late and Early Bilingual L2 Spanish. In Carol A. Klee & Timothy L. Face (eds), *Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*. Cascadilla Proceedings Project: Somerville, USA, pp100-109.
- Maria N. Carminati (2002)**. *The Processing of Italian Subject Pronouns*. (Doctoral dissertation). GLSA Publications: University of Massachusetts, USA.
- Laurent Dekydtspotter, Rex A. Sprouse & Bruce Anderson (1997)**. The Interpretive Interface in L2 Acquisition: The process-result distinction in English-French interlanguage grammars. In *Language Acquisition* 6, pp297–332.
- Laurent Dekydtspotter, Rex A. Sprouse & Kimberly Swanson (2001)**. Reflexes of Mental Architecture in Second-Language Acquisition: The interpretation of *combien* extractions in English-French interlanguage. In *Language Acquisition* 9, pp175-227.
- Laurent Dekydtspotter, Rex A. Sprouse & Rachel Thyre (2001)**. The Interpretation of Quantification at a Distance in English-French Interlanguage: Domain specificity and second-language acquisition. In *Language Acquisition* 8, pp265–320.
- Laura Domínguez & María J. Arche (2014)**. Subject Inversion in Non-native Spanish. In *Lingua* 145, pp243-265.
- Bryan Donaldson (2011a)**. Left-dislocation in Near-native French. In *Studies in Second Language Acquisition* 33, pp399-432.
- Bryan Donaldson (2011b)**. Nativelike Right-dislocations in Near-native French. In *Second Language Research* 27, pp361-390.
- Bryan Donaldson (2012)**. Syntax and Discourse in Near-native French: Clefts and focus. In *Language Learning* 62, pp902-930.
- Rod Ellis (1994)**. A Theory of Instructed Second Language Acquisition. In Nick Ellis (ed), *Implicit and Explicit Learning of Languages*. Academic Press: San Diego, USA, pp79-114.
- Rod Ellis (2003)**. *Task-based Language Learning and Teaching*. Oxford University Press: Oxford, UK / New York, USA.
- David W. Green (1986)**. Control, Activation and Resource. In *Brain and Language* 27, pp210–223.
- David W. Green (1998)**. Mental Control of the Bilingual Lexico-Semantic System. In *Bilingualism: Language and Cognition* 1, pp67–81.
- Pedro Guijarro-Fuentes & Theodoros Marinis (2007)**. Acquiring Phenomena at the Syntax/Semantics Interface in L2 Spanish: The personal preposition *a*. In Leah Roberts, Ayse Gürel, Sibel Tater & Leyla Marti (eds), *EuroSLA Yearbook* 7, John Benjamins: Amsterdam, the Netherlands, pp67-87.
- Holger Hopp (2004)**. Syntactic and Interface Knowledge in Advanced and Near-native Grammars. In Susan Foster-Cohen, Michael Sharwood Smith, Antonella Sorace, & Mitsushiko Ota (eds), *EuroSLA Yearbook* 4. John Benjamins: Amsterdam, the Netherlands, pp67-94.

- Holger Hopp (2007).** Ultimate Attainment at the Interfaces in Second Language Acquisition: Grammar and processing. (Unpublished doctoral dissertation). University of Groningen, the Netherlands.
- Holger Hopp (2009).** The Syntax–Discourse Interface in Near-native L2 Acquisition: Off-line and online performance. In *Bilingualism: Language and Cognition* 12, pp463–483.
- Ivan Ivanov (2009).** Topicality and Clitic Doubling in L2 Bulgarian: A test case for the Interface Hypothesis. In Melissa Bowles, Tania Ionin, Silvina Montrul & Annie Tremblay (eds), *Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)*. Cascadilla Proceedings Project: Somerville, USA, pp17-24.
- Ivan Ivanov (2012).** L2 Acquisition of Bulgarian Clitic Doubling: A test case for the Interface Hypothesis. In *Second Language Research* 28, pp345-368.
- Donna Lardiere (1998).** Case and Tense in the Fossilized Steady State. In *Second Language Research* 14, pp1-26.
- Donna Lardiere (2007).** *Ultimate Attainment in Second Language Acquisition*. Lawrence Erlbaum: Mahwah, USA.
- Donna Lardiere (2011).** Who Is the Interface Hypothesis About? In *Linguistic Approaches to Bilingualism* 1, pp48–53.
- Cristóbal Lozano & Amaya Mendikoetxea (2010).** Interface Conditions on Postverbal Subjects: A corpus study of L2 English. In *Bilingualism: Language and Cognition* 13, pp475-497.
- Cristóbal Lozano (2006).** The Development of the Syntax–Discourse interface: Greek learners of Spanish. In Vincent Torrens & Linda Escobar (eds), *The Acquisition of Syntax in Romance Languages*. John Benjamins: Amsterdam, Netherlands, pp371–399.
- Ana Madeira, Maria Francisca Xavier & Maria de Lurdes Crispim (2009).** A Aquisição de Sujeitos Nulos em Português L2. In *Estudos da Língua(gem) / Estudos em Aquisição de L1 e L2* 7 no 2, pp163-198.
- Ana Madeira, Maria Francisca Xavier & Maria de Lurdes Crispim (2010).** Interpretação Semântica e/ou Pragmático-Discursiva de Sujeitos na Aquisição de Português L2. In Ana Maria Brito, Fátima Silva & João Veloso (eds), *Textos Seleccionados do XXV Encontro Nacional da Associação Portuguesa de Linguística*. Edições Colibri: Lisbon, Portugal, pp513-529.
- Ana Madeira, Maria Francisca Xavier & Maria de Lurdes Crispim (2012).** Uso e Interpretação de Sujeitos Pronominais em Português L2. In Ana Costa, Cristina Flores & Nélia Alexandre (eds) *Textos Seleccionados do XXVII Encontro da Associação Portuguesa de Linguística*. APL: Lisbon, Portugal, pp376-397.
- Jason Rothman (2009).** Pragmatic Deficits with Syntactic Consequences? L2 pronominal subjects and the syntax–pragmatics interface. In *Journal of Pragmatics* 41, pp951–973.
- Robert Schmidt (1990).** The Role of Consciousness in Second Language Learning. In *Applied Linguistics* 11, pp129-158.
- Robert Schmidt (1993).** Awareness and Second Language Acquisition. In *Annual Review of Applied Linguistics* 13, pp206-226.
- Robert Schmidt (2001).** Attention. In Peter Robinson (ed), *Cognition and second language instruction*. Cambridge University Press: Cambridge, UK, pp. 3-32.

- Ludovica Serratrice, Antonella Sorace & Sandra Paoli (2004).** Subjects and Objects in Italian-English Bilingual and Monolingual Acquisition. In *Bilingualism: Language and Cognition* 7, pp183–206.
- Roumyana Slabakova (2003).** Semantic Evidence for Functional Categories in Interlanguage Grammars. In *Second Language Research* 19, pp42-75.
- Roumyana Slabakova (2009).** What is Easy and What is Hard to Acquire in a Second Language? In In Melissa Bowles, Tania Ionin, Silvina Montrul & Annie Tremblay (eds), *Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)*. Cascadilla Proceedings Project: Somerville, USA, pp280-294.
- Roumyana Slabakova (2013).** The Effect of Construction Frequency and Native Transfer on Second Language Knowledge of the Syntax–Discourse Interface. In *Applied Psycholinguistics*, pp1-29.
- Roumyana Slabakova, Paula Kempchinsky & Jason Rothman (2012).** Clitic-Doubled Left Dislocation and Focus Fronting in L2 Spanish: A case of successful acquisition at the syntax-discourse interface. In *Second Language Research* 28, pp319-343.
- Roumyana Slabakova & Jason Rothman (2011).** The Mind-Context: On acquisition at the linguistic interfaces. In *Lingua* 121, pp568-576.
- Roumyana Slabakova, Jason Rothman & Paula Kempchinsky (2011).** Gradient Competence at the Syntax-Discourse Interface. In Leah Roberts, Gabriele Pallotti & Camilla Bettoni (eds), *EuroSLA Yearbook* 11, John Benjamins: Amsterdam, the Netherlands, pp218-243.
- Roumyana Slabakova, Jason Rothman, Tania Leal Mendez, Gonzalo Campos & Paula Kempchinsky (2011).** Pragmatic Features at the L2 Syntax-Discourse Interface. In Nick Danis, Kate Mesh & Hyunsuk Sung (eds), *BUCLD 35: Proceedings of the 35th annual Boston University Conference on Language Development*. Cascadilla Proceedings Project: Somerville, USA, pp564-576.
- Antonella Sorace & Francesca Filiaci (2006).** Anaphora Resolution in Near-native Speakers of Italian. In *Second Language Research* 22, pp339-368.
- Antonella Sorace & Ludovica Serratrice (2009).** Internal and External Interfaces in Bilingual Language Development: Beyond structural overlap. In *International Journal of Bilingualism* 13, pp195-210.
- Antonella Sorace (2011).** Pinning Down the Concept of ‘Interface’ in Bilingualism. In *Linguistic Approaches to Bilingualism* 1, pp1–33.
- Antonella Sorace (2012).** Pinning Down the Concept of ‘Interface’ in Bilingualism – A reply to peer commentaries. In *Linguistic Approaches to Bilingualism* 2, pp209-216.
- Antonella Sorace, Ludovica Serratrice, Francesca Filiaci & Michela Baldo (2009).** Discourse Conditions on Subject Pronoun Realization: Testing the linguistic intuitions of older bilingual children. In *Lingua* 119, pp460-477.
- Nina Spada & Patsy M. Lightbown (2012).** Instructed Second Language Acquisition. In Peter Robinson (ed), *The Routledge Encyclopedia of Second Language Acquisition*. Routledge: London, UK / New York, USA, pp319-327.
- Ianthi Maria Tsimpli, Antonella Sorace, Caroline Heycock & Francesca Filiaci (2004).** First Language Attrition and Syntactic Subjects: A study of Greek and Italian near-native speakers of English. In *International Journal of Bilingualism*, 8, pp257–277.

- Ianthi Maria Tsimpli & Antonella Sorace (2006).** Differentiating Interfaces: L2 performance in syntax–semantics and syntax–discourse phenomena. In David Bamman, Tatiana Magnitskaia & Colleen Zaller (eds), *BUCLD 30: Proceedings of the 30th Annual Boston University Conference on Language Development*. Cascadilla Press: Somerville, USA, pp653–664.
- Lydia White (2003).** Fossilization in Steady State L2 Grammars: Persistent problems with inflectional morphology. In *Bilingualism: Language and Cognition* 6, pp129–141.
- Lydia White (2009).** Grammatical theory: Interfaces and L2 knowledge. In William C. Ritchie & Tej K. Bhatia (eds), *The New Handbook of Second Language Acquisition*. Emerald Group: Leeds, UK, pp49–68.
- Lydia White (2011a).** Second Language Acquisition at the Interfaces. In *Lingua* 121, pp577–590.
- Lydia White (2011b).** The Interface Hypothesis – How far does it extend? In *Linguistic Approaches to Bilingualism* 1, pp108–110.
- Melinda Whong, Kook-Hee Gil & Heather Marsden (eds) (2013).** *Universal Grammar and the Second Language Classroom*. Springer: Dordrecht.
- Frances Wilson (2009).** Processing at the Syntax–Discourse Interface in Second Language Acquisition. (Unpublished doctoral dissertation). University of Edinburgh, UK.
- Lucy Zhao (2011).** The Syntax and Interpretation of Embedded Null Subjects in Chinese and their Acquisition by English-speaking Learners. In Leah Roberts, Gabriele Pallotti & Camilla Bettoni (eds), *EuroSLA Yearbook* 11, John Benjamins: Amsterdam, the Netherlands, pp191–217.
- Lucy Zhao (2012).** Interpretation of Chinese Overt and Null Embedded Arguments by English-speaking Learners. In *Second Language Research* 28, pp169–190.
- Lucy Zhao (2014).** Ultimate Attainment of Anaphora Resolution in L2 Chinese. In *Second Language Research* 30, pp381–407.

9 L2 Pragmatic Problems: Do Short-Term Study Abroad Experiences Help Learners to Reduce Them?

Yoko Sato
Hosei University
yoko@hosei.ac.jp

Pragmatics has long been considered to play a crucial role in L2 communication. In fact, the author's 2008 study revealed that Japanese university students' communication problems in English were mainly caused by pragmatically inappropriate performance features (Sato, 2008, 2013). However, the effect of study abroad (SA) experiences on L2 pragmatic development has not been researched adequately. In particular, hardly any studies have included the paralinguistic and non-verbal facets, despite the fact that these have been reported as having a grave impact on communication (e.g. Fujimoto, 2003; Gumperz, 1982, 1995, 1996; Hawrysh & Zaichkowsky 1990; Sato, 2008, 2013; Wilson 2004).

A series of studies were set out in 2010 in order to fill this gap. These studies have investigated the impact of SA experiences on the communicative competence of Japanese university students, including the aforementioned under-researched areas. Each study has examined a different group of students who joined an SA programme in a particular year. As the number of students in each group is small (approximately 25), it has been suggested that more data need to be accumulated by replicating the methodology in order to draw a clearer picture of the effect of SA experiences (see Sato, 2013a and 2015a, b). There have also been some modifications to the SA programme and the method of data collection, which have shed new light on the issues addressed and have called for the examination of new data. Until 2012, the students were sent to two US universities, whereas from 2013, the students have been sent to one US university and one Australian university. This resulted in different living and learning environments for half of the students. In addition, a new interviewer-examiner was employed in 2014, following the possible Halo effect observed in the previous year (Sato, 2015a).

This paper reports the study investigating the student group who went on the SA programme in 2014. More specifically, it focuses on findings related to pragmatic problems as perceived by the interlocutor¹³.

Method

The participants were 24 second-year Business students between the ages of 19 and 21. The students were at a pre-intermediate to intermediate level of English proficiency as assessed by TOEFL ITP (423-540 [$M = 464.2$, $SD = 30.55$]; $n = 23$ ¹⁴). In 2014, they joined the Faculty's optional, credit-bearing SA programme and enrolled on ESL courses offered at two universities: at University M in Australia, the course lasted for 11 weeks and the students lived with local host families; at University R in the United States, the course lasted for 16 weeks and each student lived in an on-campus dormitory for both international and local students, sharing a room with another Japanese student. While abroad, the students submitted online monthly reports about their SA experiences. Seventeen students had had experiences abroad prior to joining the programme, fourteen of whom in English-speaking countries. However, such experiences were typically limited to family or school trips of up to two weeks. (See Table 1 for the breakdown of the participants.)

Placement	Length	M	F	Total
University M (Australia)	11 weeks	7	6	13
University R (US)	16 weeks	7	4	11
Total		14	10	24

Table 1: Breakdown of the participants by placements and genders

Before and after the programme, one-to-one, face-to-face oral proficiency interviews were conducted by an experienced native speaker (NS) assessor. The interview format was based on the IELTS Speaking test (Cambridge ESOL, 2007), with some modifications drawing on the author's previous study (Sato, 2008). Each interview lasted approximately 12 to 14 minutes. The interviewer assessed the students' performances using the public version of the IELTS Speaking band descriptors (IELTS, n.d.). The descriptors had nine bands and consisted of the four criteria of Fluency & Coherence, Lexical Resource, Grammatical Range & Accuracy, and Pronunciation. (It should be noted that the setting and content of the interviews were somewhat different from those of the official IELTS test. Therefore, this study does not claim the same degrees of reliability and validity as the original IELTS test.)

¹³ See Sato (2015b) for some other findings of the present study.

¹⁴ One female student who went to University M did not have the relevant score.

After each interview, a questionnaire was administered in order to explore pragmatic problems perceived by the interlocutor. It asked the interviewer to rate the negative affective impact that the students' performances had on him, such as irritation or unpleasantness, on a 5-point scale ranging from "1: very serious" to "5: none". The interviewer then answered an open-ended question regarding the causes of such an impact. This subjective questionnaire aimed to collect data on pragmatically inappropriate features in the students' performances, which have often been reported as causing negative affective impacts on the interlocutors (e.g. Gumperz, 1982, 1995, 1996; Sato, 2008, 2013; Thomas, 1983; Wolfson, 1983).

Results

The interviewer ratings indicate that the students' performances improved significantly after SA as a group ($p < .001$). The mean score improved from 2.94 ($SD = 1.10$) in the pretest to 3.98 ($SD = 0.70$) in the posttest. (See Figure 1.)

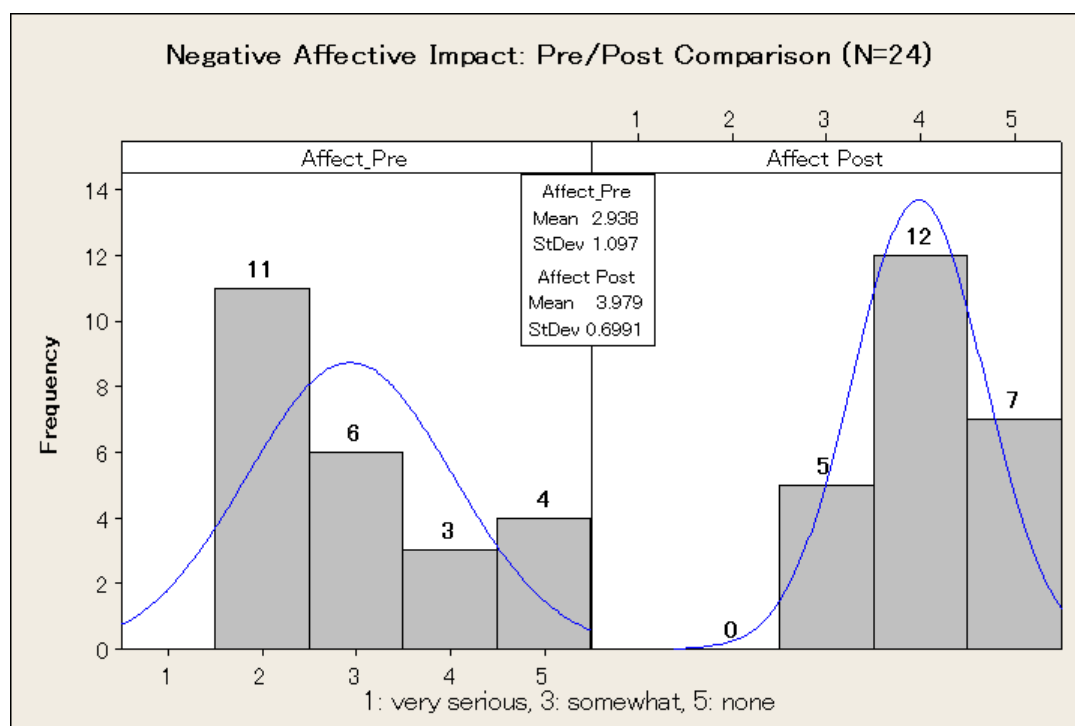


Figure 1: Negative affective impact rating: Group data

However, there were great differences amongst individuals in the pre/post score changes ($M = 1.04$, $SD = 0.97$). Two students received lower scores in the posttest, and the scores of three students stayed the same. (See Figure 2.)

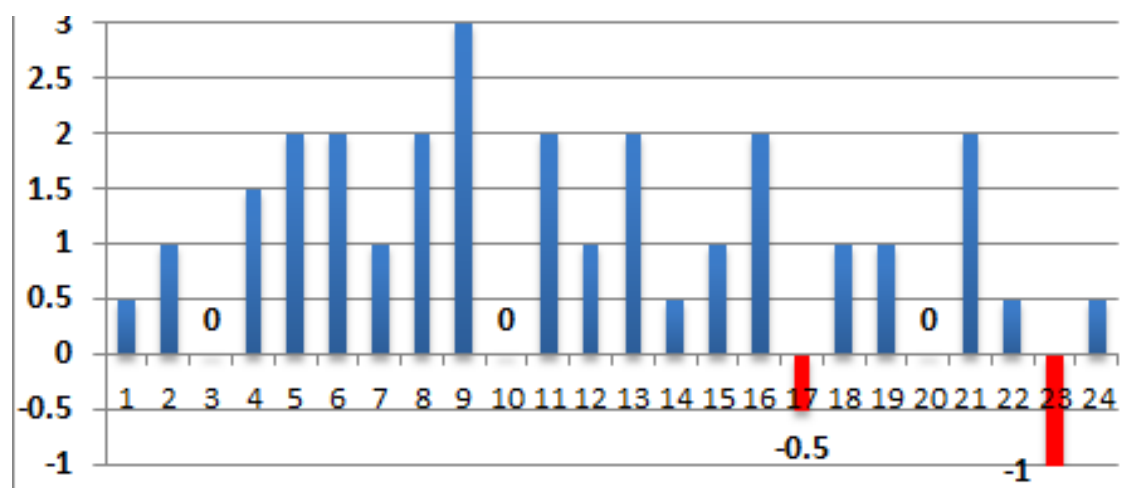


Figure 2: Negative affective impact rating: Pre/Post changes: Individual data N=24

Spearman's correlation analysis showed a very strong, significant, negative correlation between the pretest scores and the pre/post changes ($r = -.828$, $p = .000$). This means that those students who received lower scores in the pretest made greater improvement. This in turn led to the much smaller individual differences in the posttest scores; all students reached the middle score "3" (see Figures 1 and 3).

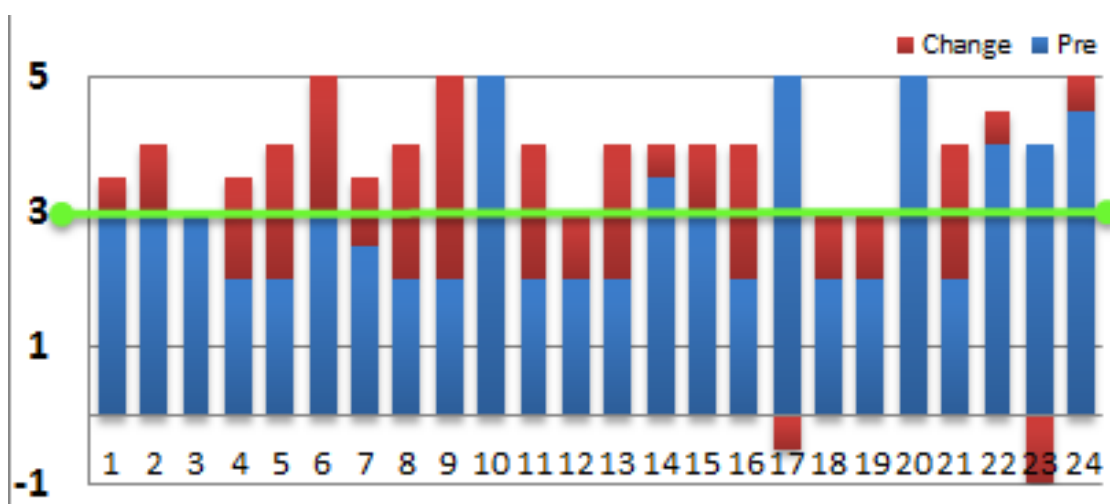


Figure 3: Negative affective impact rating: pre/post individual data N=24

A Mann-Whitney Test showed no gender-related difference in the change of scores. However, there was a significant difference between the two groups who went to different places ($p = .0304$). No significant difference was observed between these groups before SA, but the Australian group made greater improvement than the US group; the median of the former was 1.5 whereas that of the latter was 0.5.

As to the interlocutor comments on the causes of the negative affective impact, the total number decreased from 44 in the pretest to 39 in the posttest (multiple comments allowed). Comments related to dysfluency (e.g. unfilled pauses and slow speech) and inappropriate use of L1 (including fillers and self-talk) decreased to nearly one-third after SA. Conversely, incoherent utterances, which were not commented on before SA, appeared as the third most frequent cause of problems after SA. No or little change was observed for inappropriate body language (e.g. lack of eye contact), under-elaboration and suprasegmental features of pronunciation (e.g. flat intonation, low volume) (see Figure 4).

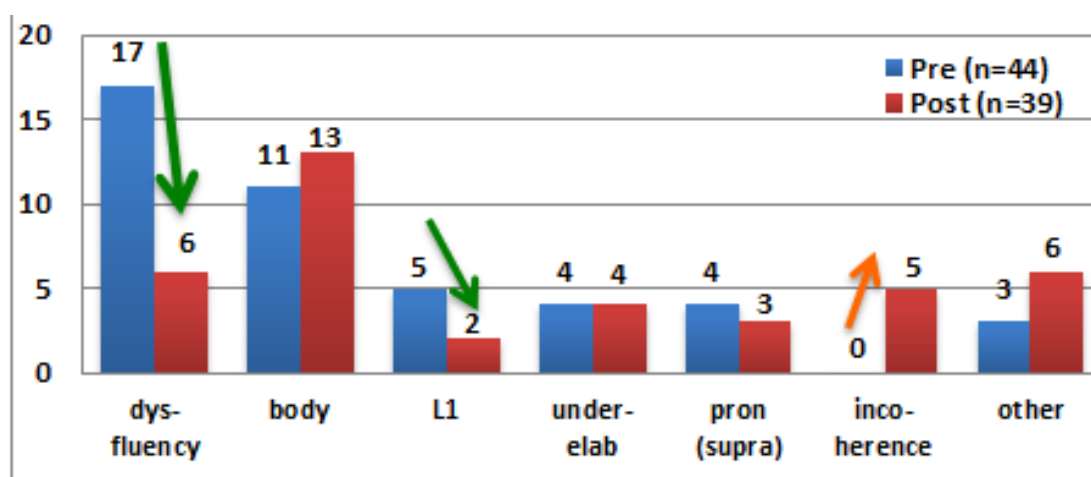


Figure 4: Causes of negative affective impact Pre/Post Comparison (Frequency)

Case studies

Case studies were conducted with two students whose scores improved and regressed the most. The score of Student 9 was much lower than that of Student 23 before SA, whereas the scores reversed after SA. In the proficiency rating, Student 9 also received a much lower score than Student 23 before SA, but her score improved greatly after SA. By contrast, Student 23 received a slightly lower score in the posttest than in the pretest. (See Figure 3 above and Table 2 below.)

Student (Sex)	Place- ment	Affective Impact Rating (1: very serious – 5: none)			Proficiency Rating (Average of 4 analytical scores)		
		pre	post	change	pre	post	change
9 (F)	Univ M (11 wks)	2.0	5.0	+3.0	3.63	5.38	+1.75
23 (M)	Univ R (16 wks)	4.0	3.0	-1.0	5.75	5.63	-0.12

Table 2: Profiles of Student 9 and Student 23

The interlocutor's comments on the causes of negative affective impact showed that, in the pretest, inappropriate use of prosody and vowel epenthesis in Student 9's speech, seemingly influenced by her L1, tired the interlocutor despite her positive body language and communication strategy. In the posttest, however, none of these were mentioned while her positive body language and communication strategies were maintained. As a result, she received the highest score "5" (see tables 3 and 4). By contrast, Student 23 received more positive than negative comments in the pretest. The only negative feature was his slow speech, which was potentially tiring and irritating. His good language ability, communication strategy, elaboration and body language together seem to have led to a relatively high score "4". In the posttest, however, he received negative comments on his under-elaboration, low volume, flat intonation, and inappropriate body language, which together created the impression that he was uninterested in the interaction (see tables 3 and 4.) As a result, he received the middle score "3".

	Pre	Post
Student 9	Intonation and word stress/rhythm does not follow English stress. Is TIRING to listen to. Japanese pronunciation intrudes (<i>isU</i> , <i>itsU</i> , <i>eh</i>). Intonation rises at the end of sentences - sounds very questioning/unsure.	A little nervous/shy but very little negative impact.
Student 23	Uses slow speech which some listeners may find tiring and might have caused irritation to some, especially when he was talking about more unfamiliar topics.	Especially during Part 1, he seemed unwilling to extend, and trailed off towards the end of sentences. He played with his watch throughout and didn't seem interested. He has a quiet, low voice and flat intonation, seemed uninterested at times and this was heightened by his playing with his watch.

Table 3: Interlocutor comments on the causes of the negative affective impact

	Pre	Post
Student 9	She is 'open' - smiles and looks interested. Really tries to answer. She's not afraid to ask for clarification. Does not have long pauses.	Eye-contact was good and she smiled throughout. She asked questions when not sure. She picked up on my (non verbal) puzzlement and explained vocab etc/reformulated.
Student 23	Language ability (vocab and grammar) good. Willing to paraphrase at most times and offers reasons without being asked. Eye contact and smiles help - likeability factor.	He had the language to explain/paraphrase and he gives examples.... Asks for clarification.

Table 4: Interlocutor comments on the strengths of learner performance

In order to explore possible factors that contributed to the above differences, the two students' monthly reports were examined. Their reports showed some distinct characteristics. Student 9's reports showed her positive attitudes throughout. She initially experienced some problems, but tried to overcome them and see the positive side of her experiences. (Note: The following excerpts were originally written in Japanese and translated into English by the present author.)

There have been a lot of problems in my life here. Only 1 week after I arrived in Australia, I had to stay with a new host family because the uncle of my original host father passed away and he had to be away for 4 days. It was a bit difficult for me because the new host family had strict rules and was fussy in a strange way. However, I think I was lucky to have experienced two host families in one study abroad programme. The original host family is very kind. They always repeat when I don't understand. They also listen to me when I tell them about what happened on the day, so I think I'm in a very good environment. (October)

She seems to be highly motivated and doing her best to study in and out of class.

.... I'm doing my best to study in class. I also started studying the basics of English again using my smartphone while commuting by bus and train. I try and read English books, too. (October)

She also sought out-of-class opportunities for mixing with locals.

I've been actively trying various extra-curricular activities. Language exchange, for example. I participate in 'Japanese Club' and 'Japaneasy', where students who want to study Japanese and those who want to study English get together and communicate. I've made a few friends there but

we're not very close yet. I'll try and email them more often so that we can get closer. (October)

This month, I participated in a volunteer activity for the first time in my life. I can confidently say it was a good experience for me. The name of the volunteer group is 'Conservation Volunteers'. A Japanese student who started studying here before us gave us information about this group, and I contacted them through their website and made a booking. I'm glad I was able to get involved in a local activity. (November)

Student 23 also sought and appreciated opportunities for intercultural communication and using English out of class. However, he lost study motivation at an early stage of SA and became lazy.

During the first month here I felt everything was new every day, but recently I've been feeling I'm a bit in a rut. After class, I eat, go back to the dorm, take a nap and study as a matter of routine. (September)

.... My conversation partner is a very nice person. We go out to a nearby town by car, go to the cinema, joke in the library and so on. I also ask her for some advice. I feel just meeting her has made participating in this programme worthwhile. I also made Brazilian friends. I got a bit bored with classes, but I'm having a very good time with them. (September)

He also spent a lot of time with other Japanese students speaking Japanese.

.... I've made very good friends with students from Nagasaki and have been spending more and more time with them. I know I should avoid speaking Japanese because I'm participating in the Study Abroad Programme to improve my English, but I think some kind of fate brought us together, so I'm in two minds whether to prioritise them or English. (September)

Provisional Conclusion, Limitations & Future Directions

The results suggest that the short-term SA experiences helped most of the Japanese university students to reduce their pragmatic problems, as measured by the affective impact their performances had on the NS interlocutor. There was a very strong, significant, negative correlation between the pre-SA scores and the pre/post changes. This means that those students who received lower scores before SA made greater improvement. After SA, much smaller differences were observed amongst individuals, and all students reached Score 3 or above out of 5. It may be that the SA experiences facilitated the development of pragmatic competence to the extent that was minimally required in the particular setting. An interesting

finding was the significantly greater improvement the Australian group made despite the fact that their stay was shorter (11 weeks) than the US group (16 weeks). It may be the result of living with local host families, which provided them with ample opportunity to communicate not only with their peers but also adults.

As for the causes of pragmatic problems perceived by the interlocutor, dysfluency and use of L1 greatly decreased after SA. Conversely, incoherent utterances, which were not commented on before SA, appeared as the third most frequent cause of problems after SA. Little change was observed for inappropriate body language and prosody. The number of comments on under-elaboration did not change. One interpretation of these findings would be that the communicative pressure and social context in the target-language environment have facilitated the development of fluency and awareness of the importance of avoiding L1. This in turn might have resulted in more prompt but less coherent utterances. Also, after SA, the students may have attempted to convey messages which were more complex than their discourse competence allowed them to express in a coherent manner.

The brief case studies of the students whose scores improved and regressed the most also suggest that the length of stay in the target environment does not necessarily differentiate the development of pragmatic competence. Rather, the attitudes towards the SA experiences, such as study motivation both in- and out-of-class, and involvement in local communities, seem to be the key factor that contributed to the difference.

There are some methodological shortcomings, which may limit the generalisability of the findings. The small sample size calls for further studies and accumulation of data. Also, the fact that only one interlocutor-examiner was used could have amplified the impact of idiosyncratic tendencies and subjectivity in the rating and comments. However, an effort was made to enhance the reliability of the ratings by selecting an experienced interviewer-examiner. Using a panel of video raters may enhance the generalizability of the rating¹⁵, although their perception may be different from that of the actual interlocutor this study set out to investigate. Further studies incorporating more qualitative investigation of attitudinal factors will provide a clearer picture of the impact of SA experiences on L2 pragmatic development.

¹⁵ The author would like to thank Rod Ellis for his suggestion at the 2014 AILA congress in Brisbane that video rating should be added to improve the generalisability of the affective impact rating.

Acknowledgements

This work was supported by a Grant-in-Aid for Scientific Research (C) (#26370710).

References

- Donna Fujimoto (2003).** Nonverbal communication and pragmatics. In *The Language Teacher*, 27:5 May 2003. Retrieved 15 July 2013 from http://www.jalt-publications.org/old_tlt/articles/2003/05/fujimoto
- John J. Gumperz (1982).** *Discourse Strategies*. Cambridge University Press: Cambridge, UK.
- John J. Gumperz (1995).** Mutual inferencing in conversation. In Ivana. Markova, Carl F. Graumann & Klaus Foppa (eds), *Mutualities in dialogue*. Cambridge University Press: Cambridge, UK, pp101-123.
- John J. Gumperz (1996).** The linguistic and cultural relativity of inference. In John. J. Gumperz & Stephen C. Levinson (eds), *Rethinking linguistic relativity*. Cambridge University Press: Cambridge, UK, pp374-406.
- Brian M. Hawrysh & Judith L. Zaichkowsky (1990).** Cultural approaches to negotiations: understanding the Japanese. In *European Journal of Marketing* 25:10, pp40-54.
- IELTS (n.d.).** *IELTS Speaking band descriptors (public version)*. Retrieved from http://www.ielts.org/PDF/UOBDS_SpeakingFinal.pdf
- Yoko Sato (2008).** *Japanese University Students' Problems and Communication Strategies in EFL Speaking: a descriptive study using retrospective verbal reports*. Unpublished doctoral dissertation, University of Reading, Reading, UK.
- Yoko Sato (2013a).** Retrospective verbal reports as a way to investigate cross-cultural pragmatic problems in oral interaction. In Bert Peeters, Kerry Mullan & Christine Béal (eds), *Cross-culturally Speaking, Speaking Cross-culturally*. Cambridge Scholars Publishing: Newcastle upon Tyne, UK, pp11-45.
- Yoko Sato (2013b).** Short-term ESL study abroad: Are Japanese students “more pleasant” to talk to after the programme? In *the Proceedings of the 45th Annual Meeting of the British Association for Applied Linguistics*, pp219-224.
- Yoko Sato (2015a).** Do short-term study abroad experiences help L2 learners to become pragmatically more appropriate? In *the Proceedings of the 47th Annual Meeting of the British Association for Applied Linguistics*, pp155-164.
- Yoko Sato (2015b).** The linguistic impact of short-term ESL study abroad experiences: The 2014 group. In *Conference of the International Journal of Arts & Sciences* 8:2, pp83–90.
- Jenny Thomas (1983).** Cross-Cultural Pragmatic Failure. In *Applied Linguistics* 4 February 1983, pp91-112.
- University of Cambridge ESOL Examinations (2007).** *IELTS handbook 2007*. Retrieved from http://www.ielts.org/pdf/IELTS_Handbook.pdf
- Andrew Wilson (2004).** When contextualization cues mislead: misunderstanding, mutual knowledge, and non-verbal gestures. In *California Linguistic Notes* 29, pp1-4.
- Nessa Wolfson (1983).** Rules of Speaking. In Jack. C. Richards & Richard W. Schmidt (eds), *Language and communication*. Longman: London, UK, pp61-87.

10 The implicit role of classroom silence in Japanese EFL contexts

Seiko Harumi
SOAS, University of London

Introduction

Despite the enthusiastic promotion of internationalisation and active oral interaction in Japanese EFL contexts by the Japanese ministry of education, learners' reticence in class is still considered deeply problematic (King, 2013). As reflected in various studies (Sacks et al., 1974), the labels 'silence' or 'silent student' may have primarily negative connotations when it comes to spoken interaction. In order to deepen our understanding of its use and shed light on other dimensions of silence as an interactional resource in L2 learning, this study suggests that it is also important to re-examine its role from an emic point of view in more depth and evaluate the positive functions of classroom silence in interaction in contexts where 'silence' has a culturally significant meaning in society (Lebra, 1987).

The following English teacher's dilemma expresses attempts to capture the role and function of specific silences in Japanese EFL contexts from a student's point of view.

I illustrate the problem as a teacher and regard learner silence as a wall, but also sympathize with students' frustration with teachers who, rather than understanding their responses, interpret them as a lack of initiative or a refusal to participate.
(Teacher A)

As can be seen, the functions of silence appear to be interpreted differently by both teachers and students. This study tries to reply to this question of the function of silence, which is frequently raised by teachers in Japanese contexts. It therefore explores implied meanings expressed by students who see silence in a positive light as a feature of the learning space.

The study was conducted through the use of multiple data analysis; a questionnaire survey, 20 hours' video-analysis of classroom interaction and prospective interviews with two students, focusing on the pragmatic role of

silence in question and answer sequences between the teacher and student in class. Research findings report various types of 'facilitative' silences from students' points of view, along with others which expanded interaction over successive turns. It also attempts to outline some pedagogical inferences which emerged from the voices expressed by Japanese EFL learners at tertiary level.

Investigations into classroom silence

In recent years, the role of silence in classroom interaction has been re-evaluated in various ways though it is undeniable that it can still have a negative influence on spoken interaction. In order to understand to what extent the positive role of silences has been recognised and promoted for its use in L2 learning, this section outlines three key positive roles of silence in classroom interaction, as well as the often discussed nature of the Japanese.

First, silence has been more widely identified in L2 classrooms as a space for learning, providing opportunities for listening, thinking, concentration and the reformulation of ideas. Its positive use has mainly been witnessed in multilingual settings where Asian students are present. For example, Goldstein (2003) reports on silent Asian students in Canadian mainstream secondary schools. Further, Japanese primary school students' classroom behaviour in UK mainstream classes has been examined by McPake and Powney (1998). What is also important to note here is that those Asian background students' reticence was initially considered obstructive to local teachers' participatory learning styles. However, teachers' in-depth long-term observations came around to a positive evaluation, considering silence as a process of learning, emphasising the role of listening.

Further, from the perspective of conversation analysis, both Gardner (2007) and Lee (2007) argue that silent reply in classroom interaction should be considered as a process of communication rather than the final product of verbal expression. Referring to the difficulties L2 learners anticipate in L2 interaction, Gardner (2007, p. 58) observes that 'broken' starts and 'delay' in answering quickly become accountable actions in themselves and claims that this needs to be understood in L2 interaction. Similarly, Lee (2007) argues that 'collective silences' in classroom interaction can function as facilitative turns, giving an example from an American tertiary educational context. The third type of silence, which has attracted attention more recently functions as a waiting-time (Ingram & Elliott, 2014, Maroni, 2011). Walsh (2002, p.5) specifically explains its role as a moment during which 'it promotes opportunities for self-expression'.

When it comes to socio-culturally oriented silence in Japan, silence has been identified as a ‘preparation time’ before verbalising one’s own ideas or replying to the questions. For example, Barnlund (1975, p.78) defines this type of silence as preparation time, that is, a ‘pre-mediation’ or ‘rehearsal’ stage, observing what other participants are thinking before actual verbalisation occurs. Similarly, La Forge (1983, p.73) also points out the importance of a complete consensus of feeling among the in-group members before taking a course of action.

So far, this study has explored points made by previous studies which emphasise and value aspects of silence involved in L2 learning, seeing them as important processes or elements within interaction or learning. It has also discussed the value of silence in Japanese contexts. The following section will ask whether, and if so how, these positive interpretations of silence are reflected in their actual performance and contribute to L2 oral interaction.

Silence in classroom interaction: methodology and data analysis

Research context

In order to depict the complex nature of silence, multiple and qualitative data analyses were adopted. The data comprise a questionnaire survey on the value of silence responded to by 85 Japanese EFL learners at tertiary level with the mean age of 21.3, 20 hours’ video recorded data and semi-structured follow-up interviews with two key participants. The questionnaire was administered at three different universities. First year students were asked open-ended questions to elicit their views on silence both outside and inside the classroom. The use of open-ended questions was considered more appropriate than providing pre-categorised answers as learners’ thoughts or beliefs could be elicited in more depth. As for the video-analysis of classroom interaction, the interaction specifically presented in this study was a comparatively relaxed conversation class with only 16 students, seven male and nine female students, seated in a square configuration with their American male teacher, who acted as one of the participants or a supporter of their conversation class. The class itself was solely devoted to oral interaction. Topics and questions were pre-submitted by learners before the class. So, in this sense, the class itself had a student-centred organization to ease the unnecessary pressure to speak. Nevertheless, the researcher initially observed quite rigid or patterned turn-taking exchanges between pairs of individuals, with no interruption or participation from others apart from the teacher.

Analysis of the questionnaire survey

The question consisted of open questions on positive and negative uses of silence both inside and outside classes. That is, this part of the enquiry aimed to elicit respondents' views on silence in their own language, Japanese, in society at the macro level and also at the institutionalised and micro level, specifically in L2 EFL classes. It intended to see how far, or whether, their views on the use of silence in their native language is reflected in their L2 learning.

Interaction-oriented reasons		Context- oriented reasons	
listening	46%	depending on topics	11%
thinking	6%	public context (inc. classroom)	8%
disagreement	1%	depending on participants	1%
astonishment	3%	Others	
depending on other's reaction	23%	should be valued in interaction in general	1%

Table 1: Positive use of silence in L1

As for the positive use of silence outside the class, when students communicated in their own language, Japanese, their answers largely fell into two categories. 79 per cent of the responses indicated interaction-oriented positive use of silence, as a listening, thinking and interactional resource such as disagreement or astonishment. However, the rest of the responses mentioned the more context-specific elements, referring to the participant background and the context of interaction itself. What is interesting here is the fact that classrooms were included within the public context category and students thought silence should be respected as a positive interactional resource in class. Another notable pattern was students' dependence on other participants' reaction to their responses. Nearly one fourth of respondents considered their use of silence positively, as an appropriate response to others' reactions. Nevertheless, when they were asked whether silence can be impolite, 69 per cent of respondents considered it extremely rude not to answer questions when asked to do so. These responses indicate that, at macro level, students are fully aware of both positive and negative uses of silence in their own cultural setting.

Students' silent responses imply	
linguistic support	56%
time-support	7%
modification of questioning	20%
psychological support	11%
consideration on turn-allocation	6%

Table 2: The use of silence in EFL class and its function

As far as the classroom context is concerned, students' responses to the specific question 'When do you use silence in English class and what do you want the teacher to understand as its role?' fell into five categories. These categories, listed in Table 2 above, included linguistic, cognitive, interactional psychological, and turn-taking uses found in students' socio-cultural backgrounds. In contrast to the use of silence outside the classroom, students' answers were more request-oriented, reflecting their need for support to promote their oral interaction rather than claiming their own positive use of silence. In addition, students' explanations for their use of silence appear to be more varied and specific in certain contexts, such as the requests to change the questioning techniques and turn allocation procedures. In this sense, although students' views on silence in class have some links with its use at a macro level, they also clearly reflect the elements of the institutionalised use. The data from the questionnaire present a generalized picture of the use of silence in class, but specific behaviour and performance needs to be investigated further, with additional observations on its use from the students themselves.

Analysis of classroom interaction

This study examines the types of silence where individual students are required to answer teacher's questions in class. In these referential questions, students are expected to give their own opinion according to their specific context. The initial turn-taking sequence examined here is as follows.

Turn 1: Teacher questions

Turn 2: Students' silent response

Turn 3: Teacher replies to silence or student self-repairs

The types of silence used in class, based on the analysis of 20 hours' video-recorded classroom interaction, were examined quantitatively using categories set out by Saville-Troike (1985, p.145). According to this categorisation, the channel of silence can be either vocal (spoken language) or a non-vocal resource (written language). Silence can be also categorized as a code, which is either verbal (filled-pauses such as fillers and noise) or non-verbal (non-vocalized cues). In this study, vocalised-verbal noise, for example, fillers were considered as filled-pauses uttered in L1 and L2. Unfilled pauses involve no vocal sound and are accompanied by non-vocal cues such as eye or head movement or facial expressions.

The recorded data shows that the types of silence used by Japanese learners usually involve vocally unfilled pauses. The mean score for unfilled pauses

recorded in four different classes within this group of students was 96.6 per cent. This clearly indicates that the students in the lower-to upper-intermediate proficiency level (CEFR B1-B2) lack the appropriate expertise to use fillers in interaction and that their oral proficiency level actually is much lower than their overall ability. This is demonstrated by the script data in the following section. This also shows that silence can lead to extreme cases of anxiety and frustration for co-participants, especially teachers and other peers from different cultural backgrounds, when they do not understand, what is happening during vocally unfilled silence.

In the conference presentation, the following examples (Excerpts 1- 3) were presented as transcribed extracts in which students used repeated silences as interactional resources and afterwards explained why they did so when interviewed. These interactions exemplify the types of silence prevalent in Japanese EFL classes.

This extract starts with an interaction between students with occasional lexical support from the teacher (Excerpt 1). This was the first in a series of three excerpts. Here, Student 2 (S2) is the key participant whose use of silence is the main focus.

Student (1), Student (2), Teacher

- 1 S1: What did you do last weekend, Takashi?
- 2 S2: (1.2) I went to (1.2) Chinese (0.3) town.
- 3 T: (0.1) China town.
- 4 S2: (1.2) with my friends
- 5 S1: Why did you go?
- 6 S2: (1.2) I eat (0.3) China (0.5) Chinese...
- 7 T: I ate Chinese food.
- 8 S1: Was it (0.3) delicious?
- 9 S2: (2.1) Yes. ((smiling))

Excerpt 1: First stage of interaction: collaborative turn exchanges

Student 1 picked up the topic about the activity at the weekend and nominated Student 2. Although there are frequent pauses within sentences, S2 managed to complete the sentence and soon after the teacher correct S2's lexical mistakes. Then S2 actually completed his sentence by adding further information in line 4. This shows Student 2's great willingness to collaborate and communicate. In the case of the second question, asked by S1 again, S2 slightly struggled with the choice of vocabulary once more and was supported by the teacher. This very brief three-way verbal interaction

demonstrates collaborative exchanges and the key student, S2 appeared to be relatively comfortable to engage. In his follow-up interview, he mentioned how he struggled to choose a suitable lexical item and the frequent lapses within his sentences strongly indicated that he was searching for the right word.

Student (2), Teacher

- 10 T: What's your favourite Chinese food?
11 S2: (2.3) ((smiling slightly))
12 T: What's your favourite Chinese dish?
13 S2: (2.9) ((leaning his head to the left))
14 T: Which Chinese dish is your favourite?
15 S2: (4.4) ((slightly looking down))
16 T: What kind of Chinese food do you like?
17 S2: (3.3) *Kakuni**

* *Kakuni* = simmered pork dish

Excerpt 2: Second stage of the interaction: prolonged use of unfilled silence

In contrast to the opening turns exchanged in Excerpt 1, since the teacher asked one particular question in line 10, the second pair of students' responses involved non-vocal silences accompanied by body and eye movement. However, the student finally responded to the fourth question asked by the teacher, in which he scaffolded by changing the lexical items and the word order of the sentences. Later, the student explained that he could not comprehend the first question as it was suddenly asked by the teacher, and that during the silences in which followed, he was trying to grasp the meaning of the question as well as attempting to listen carefully.

Student (2), Student (3), Teacher

- 18 T: What's that?
19 S2: (2.6) ((looking away))
20 T: What's that in English?
21 S2: (7.8) ((leaning his head to the left and at 4.3 seconds, starting to look at student next to him))
22 S3: (3.3) ((smiling while leaning his head to the left))
23 S2: (2.3) I don't know. ((very soft voice))
24 T: Is that a dessert? Meat? Vegetables?

Excerpt 3: Third stage of interaction: fear of alienation from the class

At this stage, Student 2 still struggled to answer the question in Excerpt 2 at line 15, as his non-verbal behaviour suggests. At line 21, the student then

asked for help from his peer, Student 3. Then, he finally verbalised the answer. For this last stage, the student commented that he simply did not know how to explain the particular dish although he was trying to think. He also commented that he could not find the right word or expression and felt awkward to say 'I don't know' straight away in class, so he asked for a help from his peer. As his peer also could not help, he finally decided to utter the word, 'I don't know.' The teacher's scaffolding started.

In this analysis of three stages of interaction, the learners' use of silence had various functions. At the first stage, his attempt to answer the question was evident as he uttered words, though with many lapses, and the meaning of these types of silences has a linguistic origin. It was a specifically lexical problem and it can be more easily understood than the entirely non-vocal silences predominant at the second stage. In this particular instance, the teacher's scaffolding (Foley, 1994) greatly helped him to gain time and opportunities to comprehend and finally answer, although in his L1. The use of silence at the last stage originated in the peer pressure he felt as he explained. It can thus be seen that the meaning and function of silence is changeable and that it is the teacher's role to determine what type of help needs to be offered as this is a key factor in successful verbal interchanges. In other words, if this kind of support is to be offered, sequences of silence can function as important turns to expand discussion over several turns. For example, even though Student 2's answer was in L1, further interaction took place. Asking for help from his peers did not produce a satisfactory answer for the teacher in this case, but it helped the student to verbalise his ideas. What is noticeable here is the absence of a request for clarification from the key student, even though he was clear about the type of difficulties he was dealing with. Instead, the silence adopted served as a form of surrogate medium.

Findings

The types of facilitative silence in classroom oral interaction can be summarised as follows. Referring back to Teacher A's comment in the introductory section above, when students have difficulty verbalising words or their own thoughts, we, as teachers, may jump to the conclusion that students merely lack proficiency or linguistic knowledge. Yet, in the follow-up interviews with the students, they often suggested that silence acted as a 'desirable space' for learning. This mirrors the results of the questionnaire in which students generally asked for some form of linguistic help. That is, they thought non-vocal responses allowed time for them to listen, comprehend and formulate ideas. The observation was the difficulty students experienced

in verbalising their own needs or ideas in front of peers. This type of silence seems to have more psychological and culturally-enhanced origins, such as the pressure of being in the limelight, alienation of one's ideas from others, quiet searching time for consensus decisions or silent requests for the teacher to give prompts for the next step. This also mirrors the responses from the questionnaire survey, but this observation illuminates the way turns are, or can be developed along with a specific form of culturally-oriented silence.

Pedagogical implications

Although the use of lengthy silence with very little or no vocal cues can be frustrating or a challenge to communication, this study has shown that we can consider its use as a possible means of facilitating further interaction. As a possible pedagogical corollary, demonstrated by the experts discussed, a number of useful approaches have been observed. First, in the case of linguistically-oriented silence, the teacher's use of scaffolding devices such as question modification and longer-waiting time can facilitate learners' successive turns. But second, in the case of psychologically and culturally oriented silence, longer waiting times can invite counter-effects as they may increase pressure on the student. In this case, changing the pace of interaction, allowing learners to interact with peers before the student replies to the original speaker may help to facilitate further interaction. As Gu and Maley (2008, p.228) claim, 'awareness of the taken-for-granted, hidden values and phenomena' is necessary in order to explore the possibility for further facilitation of interaction. Although this study has focused on understanding students' perspectives on classroom silences, it is evident that teacher's use of questions can be a key factor in promoting or terminating interaction. This role needs to be further investigated, especially in order to interact with silence as a facilitative interactional resource.

Transcription Symbols

- . (full-stop) falling intonation
- (.) micro-pause of less than 0.1 second
- (1.2) length of a silence (e.g. a pause of 1.2 seconds)
- [] beginning or end of simultaneous or overlapping speech
- (()) description of non-verbal behaviour
- Italics* foreign word (Japanese)

References

- Dean C. Barnlund (1975). *Public and private self in Japan and the United States*. The Simul Press:Tokyo, Japan.
- Joseph Foley (1994). Scaffolding. In *ELT Journal*, 48:1, pp101-102.

- Rod Gardner (2007).** Broken starts: Bricolage in turn starts in second language Talk. In Zhu Hua and Paul Seedhouse (eds.), *Language teaching as a social interaction*. Palgrave MacMillan: London, UK, pp58-71.
- Tara Goldstein (2003).** *Teaching and learning in a multilingual school: choices, risks and dilemmas*. Laurence Erlbaum Associates: Mahwah, NJ, USA.
- Quing Gu & Allan Maley (2008).** A study of Chinese students in the UK. In *Language and intercultural communication*, 84, pp224-245
- Seiko Harumi (2011).** Classroom silence: voices from Japanese EFL learners. In *ELT Journal*. 65:3, pp260-269.
- Jenni Ingram & Victoria E. Elliott (2011).** Turn-taking and 'wait time' in classroom interactions. In *Journal of Pragmatics*. 62, pp1-12.
- Jim King (2013).** *Silence in the second language classroom*. Palgrave Macmillan: London, UK.
- Paul La Forge (1983).** *Counselling and culture in second language acquisition*. Oxford University Press: Oxford, UK.
- Takie Lebra-Sugiyama (1987).** The cultural significance of silence. In *Multilingua*, 6:4, pp343-358.
- Yo-An Lee (2007).** Third turn position in teacher talk: contingency and the work of teaching. In *Journal of Pragmatics*, 39, pp190-206.
- Barbara Maroni (2011).** Pauses, gaps and wait time in classroom interaction in primary schools. In *Journal of Pragmatics*. 43, pp2081-2093.
- Joanna McPake & Janet Powney (1998).** A mirror to ourselves? The educational experience of Japanese children at school in the UK. In *Educational Research*, 402, pp169-179.
- Harvey Sacks, Emanuel. A. Shergloff & Gail Jefferson (1974).** A simplest systematics for the organization of turn-taking in conversation. In *Language*, 50:4, pp696-735.
- Muriel Saville-Troike (1985).** The place of silence in an integrated theory of communication. In Deborah Tannen & Muriel Saville-Troike (eds.), *Perspectives on silence*. Ablex: Norwood, UK, pp3-18.
- Steve Walsh (2002).** Construction or obstruction: teacher talk and learner involvement in the EFL classroom. In *Language Teaching Research*, 6:1, pp3-23.

11

Some conversational challenges of talking to children on the radio: The case of the Chris Evans Breakfast Show

Ian Nakamura
Okayama University
iannaka@okayama-u.ac.jp

Not only must the participants talk generally in ways that are acceptable to that audience (or else they will switch off), but also the talk has several specific features which are designed for that audience to relate to.

(Andrew Tolson, 2008, p. 28)

Background

My initial interest in talk with children began many years ago with a featured spot on Art Linkletter's radio and TV shows. This well-known program host interviewed children in a portion of his program called, 'Kids say the darndest things'. He would ask children questions about what they did in their daily lives. Inevitably, the kids would say something 'cute' that would draw laughter from the live audience which consisted mainly of adults. For a few minutes, the child was the center of attention as people hung on their every word.

A related personal experience was attending a weekly after school 'show and tell' group in elementary school led by a counselor. In a similar way, she would ask us to talk about any special event in our lives. Several years later I found out that this group consisted of children (in the first grade) who were considered (by their teachers) to be quiet in class. Actually, it turned out to be a good experience to have an adult give me her undivided attention and ask me about what I did (e.g., ride around on my new bicycle).

The adults in both cases would draw out meaning (both intended and unintended) of what the child would say. Besides facilitating talk in what could be an intimidating situation for a child, I now realize that there were some conversational skills on display worth noting as a teacher-researcher. For example, in both cases, the adult would help the child speaker make

his/her telling of the experience clear to the audience through the use of certain types of question, timely responses, and paraphrasing.

Introduction

Examining the organizational details of how institutional talk occurs can draw attention to a range of issues involved in facilitating talk in interactions. More specifically, I argue that a new direction in applied linguistics and pedagogy as a collaborative enterprise can be explored when broadcast talk with children and classroom talk with students are seen as sharing a similar conversational challenge: how to keep the talk going. One common way to ensure some continuity of turn taking is to sequentially organize the turn taking around question-answer exchanges. Both display and inferential questions are asked. In addition, the inquirers are conscious of time constraints when opening, developing, and closing this type of dyadic exchange in the presence of an ‘overhearing audience’ (Heritage, 1985).

The focus of this study is finding out what language teachers can learn about talking with students by investigating how talk-in-interaction occurs in another institutional setting. By focusing on the actions of the ‘adult’ speaker, this study highlights how teachers can take the initiative in a variety of ways. How talk is co-managed and how teachers contribute to it is a question that is not asked nearly enough.

As a language teacher as well as a conversation analyst, I believe that going through a systematic process of collecting, transcribing, and analyzing data of talk-in-interaction outside my usual area of teaching and research in university (see Nakamura 2006, 2008, and 2010, for studies of teacher-student talk) has revitalized my efforts to facilitate talk-in-interaction, both in and out of the classroom, with students individually as well as collectively. One upshot of this kind of study for professional development is that the same steps of research inquiry (collecting data and analyzing it) can be followed in different contexts. Regardless of the type of class, interaction between teacher and students is the core medium of instruction, understanding, and action. The benefit of such inquiry is in noticing more details of how interactions unfold (than depending on recall alone) and how details matter in heightening awareness and deepening understanding of how talk is co- accomplished.

Review of some of the issues in institutional talk

I will mention a few studies in the conversation analysis informed literature of interviews to highlight some of the issues involved in the sequential

organization of institutional talk based on question-answer exchanges. The inquiry includes: (1) How is talk opened? (2) How does the 'facilitator' keep the talk going when there is a problem? (3) What kinds of problems occur? (4) How is the talk closed? By looking at some of the conversational challenges noted in various types of institutional talk, the challenges and possible solutions become apparent.

Interactional troubles

Suchman and Jordan (1990) identify some problems in using interviews to administer a questionnaire: (1) There is an unresolved tension between the interview as an interactional event and a neutral measurement. (2) Getting a valid response means deciding questions beforehand and asking the question without any variation. (3) Meanings of questions and answers (intended and otherwise) can have multiple interpretations. (4) Interactional resources that usually clarify and repair conversations might not be mentioned. (5) Talk outside the focus area may be seen as irrelevant. Thus, issues in institutional talk become more apparent when comparing its organization with ordinary conversation.

Differences between interview and conversation: A question of constraints

Suchman and Jordan (1990) also mention the following features in co-constructing talk: (1) Local vs. external control, (2) Recipient design of questions, (3) Requirement to give an answer, (4) Establishing relevance, (5) Clarification of meaning (and whose meaning), and (6) Detection and repair of any misunderstanding. These 'problems' draw attention to institutional constraints. Koole (2003) points out a central issue in any interactional practice. The very neutrality sought can work against establishing interpersonal relationships. One way to address this tension between being objective and subjective is to study 'interactions'. Reflecting upon training and experience can help. In this direction, Richards (2011) recommends the following steps:

- (1) Refine one's interview technique: Consider the impact of the interviewer's responses on subsequent actions by the recipient.
- (2) Develop one's awareness of features of talk: Give more attention to careful transcription and study of each event.

- (3) Improve one's analytical sensitivity: Learn to focus on the talk through the actions displayed rather than on just the content.

Once an interview or more generally a talk based on the exchange of questions and answers is completed, there are some contingent problems in the reporting of the discourse. Often the focus is on the 'subject' talked about rather than the interaction. Potter and Hepburn (2005) identify four possible issues: (1) No mention of the interviewer in any reported speech. (2) No transcript conventions or reference given. (3) Lack of explanation or details of analytic observations. (4) No description of the interview set-up. (5) No acknowledgment of the interactional nature of the talk given. Making transcripts and discussing them in some detail with colleagues can alleviate some of the problems above to some degree.

Dilemma of asking questions

While questions are a common and familiar way to open and extend institutional talk with children and students, as well as patients and clients, issues about the extent of preparation and the openness to spontaneous talk need to be taken into account. For example, Puchta and Potter (1999) believe the talk should be planned in advance and be 'highly focused on predefined topics and issues'. On the other hand, it should also be 'spontaneous and conversational'.

Comparison of radio talk with children and language classroom talk

We begin with a look at the institutional talk of what Chris Evans, the host (or disc jockey) of the radio program, refers to as 'talk with the kids'. In review of specific features of the discourse event, I will discuss some basic similarities and differences present in the organization of talk in social interaction between talking with kids on the air and teaching students in the language classroom. In both situations, talk is being organized around question-answer exchanges where both display and inferential questions are asked.

Within this framework, both the radio host and teacher are aware of such institutional constraints as time limit, expected topic, recognizable transition markers, and type of response. Likewise, the radio guest and student also orient to the talk by waiting for the question before speaking. Finally, there is an overhearing audience in both contexts though the degree of recipient-

design being addressed to the listenership is more pronounced in broadcasted talk than in the classroom.

How this radio talk show is different from how language classroom interactions typically occur starts with the purpose of the talk. There are pedagogic issues that teachers will want to address with L2 user/learners. Since the majority of the radio talk participants are native English speaking children pedagogic-oriented repair seldom appears. Other ways in which the radio talk is different stem from the conventions established and followed over time as the regular listenership knows the format.

Another characteristic of radio talk is the guest is introduced without any visuals. Knowledge of the guest (e.g., age, gender, and hometown) is gained solely by listening to the opening talk. Other features are the dyadic talk occurs for the benefit of an overhearing audience (not in the studio, but wherever they happen to be listening to the show) and the three parties (host, guest, and audience) are physically situated in different locations. Technology brings them all together.

Logistical features of this particular radio talk (known to the regular followers) are the guests appear on two consecutive days (except on holidays). Openings, development, and closings of each encounter are designed to promote growing familiarity with 'just enough' redundancy to allow new listeners to join in on the way. Unlike in teaching a group of students, here the disc jockey gives his undivided attention to a single guest for a certain amount of time. The goal of broadcasted talk is to carry out an entertaining chat for the listeners.

A conversational challenge faced by both the disc jockey and teacher is continuity of talk. A look at the literature of institutional talk or 'talk at work' (Drew and Heritage, 1992) shows concern for keeping sequential organization of actions intact.

- (1) How is the talk made understandable for both the guest and other listeners? (Heritage, 1985, producing talk for an 'overhearing audience')
- (2) How do speakers know when there is a problem? How do they address it? (Schegloff, Sacks, Jefferson, 1977, 'organization of repair').

- (3) What does the host do to get the guest to elaborate on the topic-in-progress? (Hutchby, 2005, 'formulation and elicitation of feelings')

My data analytic interest is in how the use of applied linguistics (in the form of conversation analysis) can help heighten awareness and deepen understanding among practitioners of how talk-in-interaction can be co-managed for extended and meaningful social discourse. My interest in this particular radio talk is to discover what classroom teachers would like to know: How to make the conversation not only understandable and engaging for the nominated speaker, but also for the others present.

Based on recordings and transcripts, the presenter will examine moments when questions seem to 'work' in elaborating the topic-in-progress. On the other hand, there are times when responses such as 'I'm not sure' or silence need to be addressed. The next turn action becomes critical to ensure the flow of talk. Possible choices include changing the topic or rephrasing the question.

Data and analysis

As exemplified in Drew and Heritage (1992) in their seminal work, *Talk at Work*, there are general features that characterize institutional discourse such as constraints and speaker roles. At the same time, specific contexts have particular forms and routines which display and reinforce how actions take place. Talk with the kids on 'The Chris Evans Breakfast Show on BBC Radio 2' follows a routine that is familiar to regular listeners.

Background information

Just before the 7:30 am news (on weekdays), Chris Evans talks to two children (ages 7-14) by turn. The conversation revolves around the question: 'What you are doing for the first time ever?' The time slot (1.5-2.5 minutes) is divided into two portions: The first child returns to the show (from yesterday) to talk about what he/she did yesterday. After the completion of that talk, a second child appears on the show for the first time and says what he/she is going to do today. This conversation will continue the next day.

Mapping out the routine

Looking at the structure of the talk sequentially, the following steps are taken:

- (1) Greetings exchanged between host and guest.

- (2) Introduction by host of the guest to the listeners.
- (3) Telling his/her experience is guided by host's questions and remarks.
- (4) Pre-closing: 'Mark yourself out of ten'.
- (5) Closing remarks by the host.
- (6) Introduction of the next guest by the host
- (7) Greetings exchanged.
- (8) Focus question asked and answered briefly (as a plan to be carried out).
- (9) Pre-closing request: 'Come back tomorrow and tell us about it'.

To make the organization even easier to follow, in addition to the focus question about what they did (or rather what they will do), other set phrases are always used as in steps 4 and 9. These standard instructions help the audience as well as the recipient know how to act. Having an expected trajectory can be important for the juvenile participants (current speaker, listeners, and even potential future speakers). The clarity of the procedure also helps relieve some of the anxiety of performing live. It should be remembered that they are talking on the phone with the talk being broadcasted. Finally, it should be noted that music is strategically inserted with a fanfare to mark such key transitional moments as heralding the arrival of a new guest, concluding this portion of the show, and leading into the news (which will be given by a different person).

Accounts of how the talk occurred

Out of a collection of over fifty talks, I will discuss four transcribed excerpts from talks with two children in terms of sequential organization on a turn by turn basis where a response to a prior turn is the main mode of operation. These talks are broadcasted on the radio, so except for pictures of the host and the date of the programs in the archives on the BBC radio 2 homepage, there is no other visual information. Thus, the speakers and audience rely solely on what they hear.

Example no. 1: Opening the talk

The host and child talked briefly the day before to establish what she planned to do. The purpose of the talk today is to tell the host and the audience what she actually did. This second talk (with the same child) lasts about one minute. We join the conversation after the speaker has been re-introduced to the audience.

Note: Basic transcript conventions are taken from conversation analysis. See Wong and Warring (2010) for an explanation. The speakers are H: program host-disc jockey (Chris Evans), K: kid. She is seven years old.

- 6 H: So what did you do for the first time ever yesterday?
7 K: I went to Blackpool illuminations with my guides
8 H: Alright so how many of you went
9 K: Uh I'm not sure. hhh
10 H: Okay did you go on the train, bus, mini-bus=
11 K: =Went on the train
12 H: Okay excellent. So did you go walking toward the lights under an
13 umbrella?

A common way to mark a transition from introductions and greetings to the main question or from one question to the next one is through the use of 'So' as a preface to the question. The focus of the talk is stated through the question, 'What did you do for the first time ever?' The rest of the talk will be devoted to developing a conversation around what the recipient nominates as the relevant topic. Here it is about her trip to Blackpool to see the illuminations. The first question-answer exchange (lines 6-7 goes smoothly), but some trouble appears to occur in the response to the second question (line 8). It can be imagined that the first answer was prepared in advance as the opening question is always the same.

The next question, however, cannot be so clearly anticipated. Line 8 begins with an acknowledgment of the first answer and the use again of 'so' to mark the transition to the next question, 'How many of you went?' (with stress on the final sound). 'I'm not sure' is sandwiched between a delay token ('Uh') and brief laughter (hhh). This arrangement suggests trouble in providing an answer.

The interesting conversational challenge for the host in the third position of the sequence of turns after the question and answer is to decide what to do next. 'Okay' in line 10 works as acceptance or at least a closure to that line of inquiry as H asks a different question about transportation. Instead of asking about how she went there, H gives her a list of options to choose from. K picks up quickly (= indicates a latched turn without the usual pause present between change of speakers) on what is being asked. The host then gives an upgraded receipt from 'okay' to 'excellent' (line 12). Again 'so' marks that another question is coming. This one is in yes/no form which the host might have decided is easier to answer than a 'how' question.

Two observations can be made: (1) After K's 'I'm not sure' response to a wh-question, H has asked questions with built-in choices. (2) The questions, in lines 8, 10, and 12-13, demonstrate H's familiarity with going to the same place by providing knowledgeable and recognizable choices. He does confirm this idea of having direct experience (in his next extended turn not shown in this transcript) by telling her about going there with his own family when young.

As the same talk continues below, another wh-question is asked, 'How long did it take?' (line 20). K's response, in line 21, is the same as in line 9 above with a delay token, 'I'm not sure', and brief laughter. This time, in line 22, instead of moving on to a different question like in line 10, the host provides choices in how to answer as well as a yes/no option (line 22). K takes the last option and says, 'no'. The host appears to closely monitor what the child can answer or has difficulty with (not unlike what a teacher does). Sometimes the open-ended questions (e.g., information type) can be challenging for younger children, (e.g., 5-7 years old) like in this case, to understand the intent of the inquiry and how to answer. Below, the same conversation continues.

Example no. 2: Closing the talk

- 20 H: And how long did it take you to see what you saw in the end.
21 K: Uhm (.) I'm not sure hh=
22 H: =Okay an hour, two hours, the whole night? You still there?
23 K: Eh no.
24 H: All right okay alright mark yourself out of ten then please.
25 K: A million
26 H: A million out of ten you were brilliant thank you Loula Chinri (0.6)
27 for being on the show goodbye (0.5) have a nice day; have a nice life

Time is running down, so H then clearly marks the transition to his standard pre-closing question (of asking the child to give the experience a score on a scale from one to ten) prefaced with three discourse markers in rapid succession, 'all right', 'okay', 'alright' (line 24). It is beyond the scope of this paper to do more than mention that the children often rate their experience well over ten (e.g., 100, 1 million). The host closes the interaction with final remarks to the child. Possibly since there seemed to have been some hesitation by K to be on the show again (heard in pre-show talk), H is particularly positive and supportive of her performance.

These 'conversations' take place live on the radio. The children call in (probably with the help of their parents) several minutes earlier when the

host has announced that he is now taking calls, ‘to talk to the kids’. Someone on staff receives the call and provides the host with basic information about the guest. In the next example, after the obligatory greetings and introductions are made, the host asks the standard question and the child nominates a topic. In this case, he will visit the Tower of London ‘to see the poppies’.

Example 3: The first visit on the show

- 10 H: What are you doing for the first time ever today.
11 K: I’m going to the great Tower of London to see the poppies.
12 H: Alright you’re going to have a great time there. Will you come back
13 tomorrow to tell us if they’re still beautiful?
14 (1.7)
15 K: Definitely.
16 H: Thank you very much indeed. Lovely thing to do half term.
17 Half past seven now.

Again, H shows familiarity with the place where the child will visit. Immediately after K responds in lines 11 and 15, H comments favorably on the choice of going to the Tower of London. The design of these remarks (lines 12 and 16) appear to be intended for both the boy (who is 10 years old) and the audience.

In terms of what the host does with his two turns (lines 10 and 12) in this preliminary meeting (which lasts only 30 seconds), H seems to have a routine of starting off with a wh-question and then making a comment and adding a yes/no question. Finally, a clear closing of the talk comes with H’s announcement of the time, 7:30 a.m., and that the news will be next.

The talk the day after the boy’s visit to the Tower of London begins with a review of the question and answer. While these talks are live on the air and the participants seem to be interacting spontaneously on an orderly turn by turn basis, there are signs that preparation, technique, and strategy are involved in limiting yet elaborating the topic of seeing the poppies. Certainly there are many other things to see in the Tower of London which could be mentioned. Some speculation could be made that the focus is timely with many war memorial ceremonies taking place around that time.

Example no. 4: The second visit

- 4 H: Where did you go yesterday for the first time ever
5 K: I went to the Tower of London.
6 H: And you went to see the poppies ah: what did you think?

- 7 K: Pardon?
8 H: What did you think of the poppies?
9 K: They were amazing.
10 H: They are amazing aren't they. How many are there?
11 K: There are eight hundred and eighty eight thousand two hundred
12 and forty six.

The sequence, line 6 to line 12, demonstrates the coordinated accomplishment of the interaction. H retrieves the topic of 'poppies' from yesterday. However, some trouble appears to immediately occur either by H's slightly awkward wording of the question (line 6) or by K simply not fully hearing what was said. In response to 'pardon', H rephrases the initial question. While they both agree with the choice of 'amazing', the verb tense differs. The boy says 'were' and the host says 'are' (lines 9 and 10, respectively). There is some looseness in the talk which participants need to do some adjusting. That will happen in the next part.

What becomes clear in terms of the issue of spontaneous versus preparedness is the next question-answer pair. Asking how many poppies were/are there is not the usual question we would ask anyone unless there is a special significance attached to the exact amount. The boy without hesitation gives an exact number. Later in this same talk (not included in the transcript), when K is asked to rate his experience, he cleverly repeats the same number, '888,246'.

Preliminary findings

I will mention three issues that were noticed with relevance to classroom teaching. I pose the observations as questions that I believe merit further discussion for how institutional talk is conducted. These issues are commonly encountered by the initiator-facilitator of institutional talk such as interviewers, examiners, medical and legal consultants, and instructors.

- (1) What kind of responses did not seem to answer the question? Examples found in the data included: 'I'm not sure', 'No' in reply to wh-question, long pause, 'Pardon?'
- (2) What did the host do in response? At least three different actions were noticed: Changed the topic or reworded question to give a choice, moved on to the next question, and repeated the question then added more information.
- (3) What kind of questions appeared to cause trouble? I speculate that children may recall or remember events in different ways than adults. The first two questions (in example 1, line 8 and

example 2, line 20), ‘How many?’ and ‘How long?’ are numerically oriented. The girl’s response in both cases is ‘Uhm, I’m not sure’. When the host attaches a fourth choice to clarify the initial question of how long she stayed (‘You still there?’), it carries a sense of humor that does not appear to be picked up on. The host then gives a statement with a question (example 3, lines 12-13 and example 4, line 6). Both occur within the same turn. The boy’s responses are a pause or delay and ‘Pardon?’ respectively.

Pedagogic implications and applications: Clarifying issues in decision-making

More specifically for teacher training and continuous professional development, the data and analysis allow teacher trainers and educators to identify and discuss issues that practitioners in the classroom commonly encounter.

- (1) What do teachers mean when they say that the student did not answer the question asked? Do they mean the correct form, relevant content, too long of a pause, or something else?
- (2) What other choices besides speaking louder, slower, code switching to L1, repeating, paraphrasing, waiting, changing the topic, or calling on another student are available to teachers when the student does not answer the question?
- (3) How much of the talk between teacher and student should be prepared (even scripted) and how much should be open ended and spontaneous?
- (4) How much should teachers add their own opinion and experience to the topic-in-progress? Does it help or hinder students’ efforts to talk?

Taking the time to study the sequential organization of talk-in-interactions can make a difference the next time teacher-student talk takes place. Practitioners, however, can never be fully prepared for what will happen as social interactions are co-produced. Perhaps, the bigger point is that because conversation can be unpredictable (especially with children); whatever preparation that can be done in advance is even more important to help with fluidity and continuity not to mention meaning. Regardless of the type of adult-guided discourse with children, building in opportunities for the children to speak somewhat freely enriches the meaningfulness of the account being told.

Conclusion

While talk-in-interaction between the children and host is enjoyable in its own right, analysis of the discourse through questions, responses, and repair allows researchers and teachers to consider the array of conversational skills that are employed, the interplay between planning and reacting, and the sharing of knowledge. Stepping outside of routine classroom spoken discourse and looking at how a professional radio talk host skillfully navigates the conversation raises important issues and questions about how talk is socially and interactionally managed.

Since classroom interactions as well as the talk with kids on the radio are co-accomplished social exchanges through spoken language, any plans to improve communication should focus not only on what learners and children say (or do not say), but also on how the teacher and radio host sequentially organize the orderly turn taking through recipient-designed moves.

Methodologically speaking, a process of collecting data and analyzing it through recordings, transcripts, along with some form of discourse analysis allows researchers, practitioners, and other participants to examine and discover fresh insights into how talk is co-constructed and what it means for practitioners who are interested in improving how they interact. Who facilitates the talk does matter. Hutchby (1999) is helpful in showing one way forward. Think beyond the institutions themselves as shaping the talk speakers do. He explains what conversation analysts have been arguing.

[I]nstitutions do not define the kind of talk produced within them; rather participants' ways of designing their talk actually constructs the 'institutionality' of such settings (p. 41).

The co-management of the participants' actions as the interaction unfolds through questions and responses forms a trajectory of co-accomplished talk. This aspect of spoken discourse is just one area where further research and application are needed.

Transcript convention used (Wong & Waring 2010)

- . (period) falling intonation
- ? (question mark) rising intonation
- = (equal sign) latch or continuing speech
- (.) (period in parentheses) micro-pause of less than 0.2 seconds

(1.5) (number in parentheses) length of a silence in seconds
A::h (colon(s)) prolonging of sound
Hh (series of h's) aspiration or laughter

References

- J. Maxwell Atkinson & Paul Drew (1979).** *Order in Court: Verbal Interaction in Judicial Settings*. Macmillan: London, UK.
- Steven Clayman & John Heritage (2002).** *The News Interview: Journalists and Public Figures on the Air*. Cambridge University Press: Cambridge, UK.
- Paul Drew & John Heritage (1992)** (eds). *Talk at work: Interaction in institutional settings*. Cambridge University Press: Cambridge, UK.
- John Heritage (1985).** Aspects of the production of talk for an overhearing audience. In Teun van Dijk (ed), *Handbook of discourse analysis, vol. 3, Discourse and dialogue*. Academic Press: London, UK, pp95-119.
- Ian Hutchby (1999).** Frame attunement and footing in the organisation of talk radio openings. In *Journal of Sociolinguistics* 3, pp41-63.
- Ian Hutchby (2005).** 'Active listening': Formulations and the elicitation of feelings—talk in child counseling. In *Research on Language and Social Interaction* 38:3, pp303-329.
- Tom Koole (2003).** Affiliation and detachment in interviewer answer receipts. In Harry van den Berg, Margaret Wetherell & Hanneke Houtkoop-Steenstra (eds), *Analyzing race: Multidisciplinary approaches to the interview*. Cambridge University Press: Cambridge, UK, pp178-199.
- Ian Nakamura (2006).** Understanding 'stuckness': Descriptions and interpretations of how EFL speakers and a native speaker co-manage talk-in-interaction. Unpublished PhD thesis. Lancaster University, UK.
- Ian Nakamura (2008).** Understanding how teacher and student talk with each other: An exploration of how 'repair' displays the co-management of talk-in-interaction. In *Language Teaching Research* 12, pp265-283.
- Ian Nakamura (2010).** Formulation as evidence of understanding in teacher-student talk. In *ELT Journal* 64, pp125-134.
- Jonathan Potter & Alexis Hepburn (2005).** Qualitative interviews in psychology: Problems and possibilities. In *Qualitative Research in Psychology* 2, pp281-307.
- Claudia Puchta & Jonathan Potter (1999).** Asking elaborate questions: Focus groups and the management of spontaneity. In *Journal of Sociolinguistics* 3, pp14-335.
- Keith Richards (2011).** Using micro-analysis in interviewer training: 'Continuers' and interviewer positioning. In *Applied Linguistics* 32, pp95-112.
- Emanuel Schegloff, Gail Jefferson & Harvey Sacks (1977).** The preference for self-correction in the organization of repair in conversation. In *Language* 53, pp361-382.
- Lucy Suchman & Brigitte Jordan (1990).** Interactional troubles in face-to-face survey interviews. In *Journal of American Statistical Association* 85:4, pp232-241.
- Andrew Tolson (2008)** (ed). *Television talk shows: Discourse, performance, spectacle*. Routledge: New York, USA.
- Jean Wong & Hansun Warring (2010).** *Conversation analysis and second language pedagogy: A guide for ESL/EFL teachers*. Routledge: New York, USA.

12

A comparative study of evaluative acts in classical concert reviews in British and Hong Kong newspapers

Fong-wa Ha
University of Essex
fwha@essex.ac.uk

Introduction

Despite the growing interest in genre and discourse analysis in English academic review writing (e.g., Bhatia 2008, Hyland 2000, Lewin 2005), relatively little work has been undertaken to compare the rhetorical acts of English and Chinese discourse. The genre of music criticism has received even less attention in either English or Chinese. One existing study of music review genres (Ha, 2011) shows that over time English record reviews have become less informative and more evaluative. This paper reports on a part of a larger study on music criticism, which aims to fill the research gap by comparing the incidence and type of rhetorical acts employed in classical concert reviews in both British English and Hong Kong Chinese newspapers. Each corpus contains 150 reviews published within a 10-year period, from August 2003 to July 2013. General trends of evaluative and non-evaluative acts are compared and reported cross-culturally.

Literature review

Music criticism is defined as “the profession of writing about the aesthetics, history, and evolution of music and of reviewing musical composition and performance in newspapers, periodicals, books, and on radio and TV” (Scholes 1970, p.171), which aims to “establish a line of communication between the creative artist and the public” (Sadie 1980, p.44). Oscar Thompson (1979), a music critic for the *New York Sun*, emphasised in 1934 that the function of music criticism is to “present a clear picture of what the music is, with its good points and bad, or what the performance was, with its salient characteristics, meritorious or otherwise” (p.29). To mark a distinction between academic and journalistic music criticism, Schick (1996) specifically points out that music criticism in newspapers provides an up-to-

date account of musical events, which “discusses and summarises the musical experience” for a lay audience and/or professional musicians (p.3).

The concert review, a genre of music criticism in popular media, has remained unstudied. A related genre, academic book reviews, has been given much attention in recent years. Reviews are a useful tool for academics because they are crucial sites of “disciplinary engagement” with much higher “interpersonal stakes” than research articles (Hyland 2000, p.41).

Evaluation, the expression of a writer to convey their attitudes, beliefs or values to their readers (Hunt and Vipond 1986, Hyland and Tse 2009), plays a crucial role in the function of reviews (Bondi 2009). According to Hyland (2000), reviews are “centrally evaluative” (Hyland 2000, p.11), produced with an explicit aim to evaluate the work and to provide a platform for members of a discourse community to share ideas and analyses (Hyland and Diani 2009).

Nevertheless, early academic book reviews or medical reviews attempted to avoid manipulating readers’ judgement (Duncan 1773, in Salager-Meyer et al 2007) by containing no evaluation whatsoever (Hyland 2000, Salager-Meyer et al 2007). In music criticism, there have been debates on whether music criticism should “describe or evaluate a composition or performance” (Hoger 1992, p.31). It has been argued that music criticism should give a clear image of the music and should be primarily informative (Thompson 1979). However, most music reviews include evaluation in practice (Schick 1996).

Despite the arguments about the function of music criticism, a piece of writing in the review genres usually consists of evaluation (praise and/or criticism), and informative, non-evaluative parts which are regarded as “neutral descriptions of aims, organisation and content” (Hyland 2000, p.44). Praise, as according to Holmes (1988, p.44), is “an act which attributes credit to another for some characteristic, attribute, skill, etc., which is positively valued by the writer and suggests a more intense or detailed response than simple agreement”. Criticism, on the other hand, is described as “the expression of dissatisfaction or negative comment” (Holmes 1988, p.44).

Brown and Levinson’s (1987) politeness theory is a major theoretical framework on which many studies of evaluation are based (e.g. Johnson 1992; Gea Valor 2000; Hyland 2000; Lewin 2005; Mackiewicz 2007). The face model proposed by the theory is thought to contribute to writers’

strategies in giving praise and criticism in reviews. There are two types of face: positive face is a person's desire for approval and acceptance; negative face indicates the desire for one's actions not to be impeded. In Brown and Levinson's terms, bestowing praise addresses one's positive face and redresses face-threatening acts (FTAs). A criticism, on the other hand, is a risky act which undermines one's positive face. Studies in reviews suggest that critics often apply mitigation strategies to redress FTAs (Hyland 2000; Hyland and Hyland 2001; Ha 2011; Itakura 2012).

Brown and Levinson's influential face model, however, has been criticised for not being suitable to account for Chinese contexts (Gu 1990; Mao 1994). The arguments are that the notion of negative face is either different (Gu 1990) or plays an insignificant role (Mao 1994) in Chinese culture, and that politeness is not only instrumental but also normative in the Chinese context (Gu 1990). Ji (2000), on the other hand, asserts that Chinese people are highly sensitive to negative face, and there is no evidence that the face model is not able to be identified in Chinese culture. Nevertheless, Ji does admit that Brown and Levinson's model might not be adequate to explain all strategies in different cultures, and that the same criteria cannot identify all strategies across cultures.

With Ji's concerns about cultural differences in mind, this study draws on not only Brown and Levinson's, but also a number of studies to compile a balanced framework addressing the evaluative strategies applied in both English and Chinese concert reviews (Hyland 2000 and 2004, Hyland and Hyland 2001, Ha 2011, Culpeper et al 2003, Shaw 2004, Stotesbury 2006, Itakura and Tsui 2011, Rockwell 2006, Rogerson-Revell 2007).

The research question is as follows:

What are the similarities and differences in the overall trends of rhetorical acts between English and Chinese concert reviews?

Corpora and analysis

The data for this research consists of two corpora compiled from three English newspapers published in the UK and three Chinese newspapers published in Hong Kong. The British newspapers are: *The Times*, *The Guardian* and *The Daily Telegraph*. The Hong Kong Chinese newspapers are: *The Hong Kong Economic Journal* (信報), *Ta Kung Pao* (大公報) and *Ming Pao Daily News* (明報). Each corpus contains 150

reviews. The selection of newspapers was based on two criteria:

1. Broadsheets that published the most classical concert reviews over the 10-year period from August 2003 to July 2013;
2. Broadsheets that employ a relatively large number of reviewers which reduces the risk of results being skewed by the domination of a few reviewers and their individual styles of writing (See Table 1).

Newspapers	English	Chinese
Number of reviews	150	150
Total number of reviewers	25	26
Total number of words / characters	198,793	53,416
Number of words / characters per review	356	1,525

Table 1: The corpora – English & Chinese newspapers August 2003 – July 2013

The following selecting criteria of reviews to include in the corpus were set to minimise variables that might affect the results of comparison:

1. Only single concert reviews should be selected;
2. Reviews selected should only write about western classical music (no folk, jazz, pop, metal, new age, etcetera);
3. Reviews about Chinese classical music should be excluded from the corpus.

A hand-tagged qualitative analysis was conducted to code each concert review. Based on Hyland's (2000) analytical framework of evaluation on book reviews, codes were identified based on their rhetorical purposes:

- *Concert Evaluative acts*: praise or criticism of aspects of the concert;
- *Background comments*: evaluations of aspects other than the concert itself. To make a distinction with evaluations of aspects of the concert, background comments were categorised as *positive comments* and *negative comments*;
- *Non-evaluative remarks*: information about aspects of the concert which is neutral and contains no evaluation.

Both praise and criticism were categorised into global and specific evaluations. *Global Praise* and *Global Criticism* describe overall comments made about a concert. *Specific Praise* and *Specific Criticism*, on the other hand, target specific aspects of a concert, namely programme, composer or composition, performer or performance, acoustics, instruments, audience behaviour, concert management, programme notes, and other aspects. The analytical frameworks for praise and criticism on aspects of the concert respectively were established as in Figures 1 and 2.

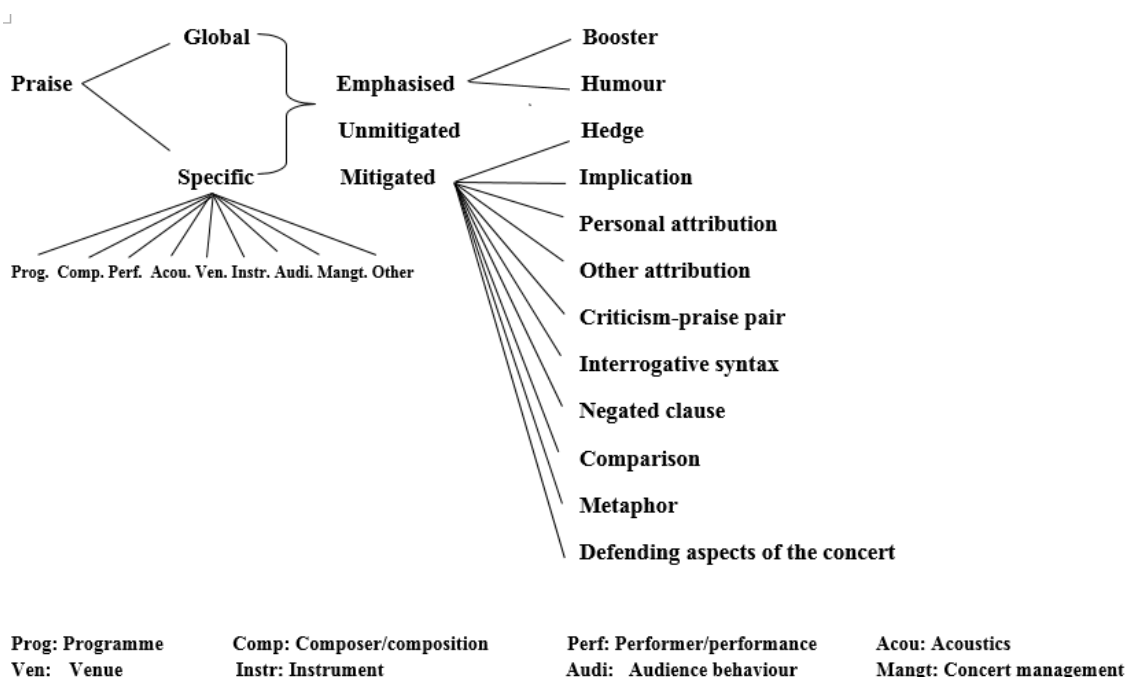


Figure 1: The analytical framework of praise

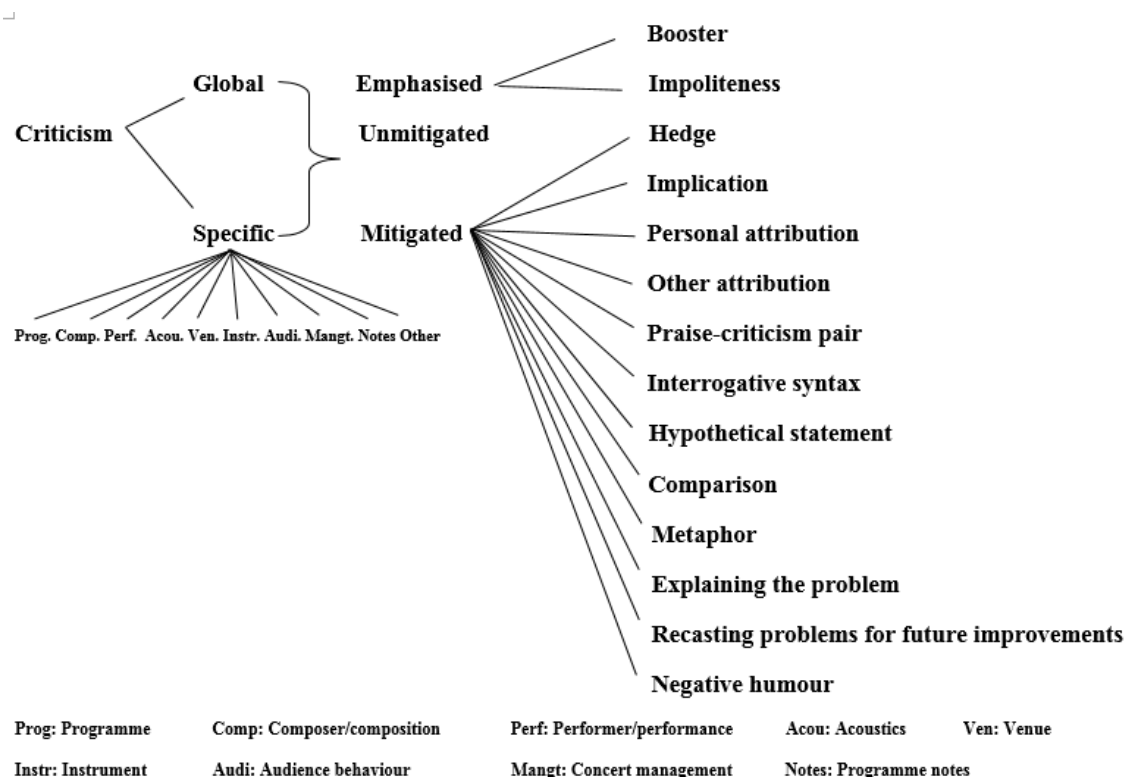


Figure 2: The analytical framework of criticism

An evaluation was further categorised into emphasised, unmitigated or mitigated praise or criticism (following Hyland 2000 and 2004, Hyland and Hyland 2001, Ha 2011, Culpeper et al 2003, Shaw 2004, Stotesbury 2006, Itakura and Tsui 2010, Rockwell 2006, Rogerson-Revell 2007). An

emphasised evaluation strengthens the illocutionary force of praise or criticism, thus making the evaluation even stronger. In contrast to emphasised evaluation, mitigation softens the evaluative act. A mitigated praise is applied to ‘make comments sound more objective’ (Ha 2011, p.357) or to work as “rephrased criticism” (Hyland/Hyland 2001, p.198). A mitigated criticism is used to limit the harshness of criticism (Hyland 2000, Hyland and Hyland 2001, Stotesbury 2006). An unmitigated evaluation is neither strengthened nor weakened in its illocutionary force. It usually contains explicit positive or negative lexical items to indicate praise or criticism (Hyland 2000). Sub-categories of emphasised and mitigated evaluations are listed in Figures 1 and 2.

Two principles were set when coding the reviews:

1. A clause/sentence containing more than one positive/negative semantic item will be coded as one instance of evaluation if it only refers to a single aspect of the concert under review. For instance:

她的聲線有力，咬字清晰

[Her voice was strong and the articulation was clear.]

In the above example, both “有力”(strong) and “清晰”(clear) are explicit positive semantic items referring to the performance of the vocalist, therefore the whole sentence is coded as one instance of unmitigated specific praise on performer/performance.

2. Every evaluative strategy will be counted if more than one appears in a single clause/sentence. For instance:

The music, all played far more expertly than it deserved by the London Symphony Orchestra under Daniel Harding (are they really that short of work?) was profoundly depressing.

The following codes are derived from the above example:

Emphasised praise (Booster) on performer/performance: *all played far more expertly than it deserved*;

Emphasised criticism (Impoliteness) on composer/composition: *are they really that short of work?*;

Emphasised criticism (Booster) on composer/composition: *profoundly depressing*.

To ensure reliability of the analysis, 10 reviews from each corpus were independently coded by a scholar from a British university (English) and a scholar from a Hong Kong university (Chinese). A very high reliability coefficient was obtained (English: 99.4%, Chinese: 99.54%).

After the analysis was completed, the frequencies of positive and negative evaluative acts were calculated and compared using SPSS. To enable a comparable analysis of the two corpora, the evaluative acts were calculated by percentage (%) and compared across the corpora. This is because of the inherent difference between English words and Chinese characters (Ge, Pratt et al. 1999; Teahan, Wen et al. 2000).

The similarities between English and Chinese reviews

Through analyses, a number of general trends have been found to be common to both English and Chinese concert reviews. They are:

- Reviews are predominantly evaluative;
- More reviews open and close with positive remarks than negative remarks;
- Critics prefer to comment on specific aspects of a concert than giving global evaluations;
- Global criticism is very rare in reviews;
- *Performer and Performance* is the most evaluated aspect of a concert;
- Critics praise more than criticise in their reviews;
- Criticism is largely indirect (mitigated).

Details of the findings are explained below.

Evaluation vs. non-evaluation

Comparison was first made between numbers of reviews that contained at least one instance of each of the main types of rhetorical acts, i.e. evaluative and non-evaluative acts. All the English and Chinese reviews contain evaluation. Evaluation can initially be divided into *Background Comments* and *Concert Evaluation*. Almost all reviews also contained non-evaluation, with the Chinese reviews (97.33%) containing slightly more than the English reviews (94%). It can therefore be said that both evaluative and non-evaluative information were common components of both English and Chinese concert reviews.

The above picture is somewhat simplified as it does not distinguish between reviews that contain a greater or lesser number of occurrences of any given type of act. Therefore the numbers of acts of each type in each review were examined. A total of 2952 rhetorical acts were recorded in the 150 English reviews sampled from August 2003 to July 2013, of which 2478 (83.94%) were evaluative including background comments and evaluation and only 474 (16.06%) were non-evaluative. A total of 4318 rhetorical acts in the

Chinese reviews were recorded, of which 3519 (81.50%) were evaluative acts and only 799 (18.50%) were non-evaluative. (Table 2)

		English Reviews		Chinese Reviews	
Evaluative acts	Background comments	145 (4.91%)	2,478 (83.94%)	456 (10.56%)	3,519 (81.50%)
	Concert evaluation	2,333 (79.03%)		3,063 (70.94%)	
Non-evaluative acts		474 (16.06%)		799 (18.50%)	
Total number of acts		2,952		4,318	

Table 2: The occurrences and percentages of rhetorical acts in English and Chinese reviews

As can be seen in Table 3, both English and Chinese reviews seem to be dominantly evaluative. The English reviews contain approximately four times more evaluative acts ($M = 16.52$, $SD = 5.58$) than non-evaluative acts ($M = 3.16$, $SD = 1.92$). The Chinese reviews also contain about four times more evaluative acts ($M = 23.46$, $SD = 9.97$) than non-evaluative acts ($M = 5.33$, $SD = 3.97$). Wilcoxon tests further indicates that the percentages of evaluative acts are significantly higher than those of non-evaluative acts for both English ($p < 0.001$) and Chinese reviews ($p < 0.001$). However, the result of a Mann-Whitney test indicates that there are no significant differences between the English and Chinese groups in their proportions of evaluative and non-evaluative acts.

	English concert reviews (N =150)				Chinese concert reviews (N = 150)			
Category of acts	Mean (per review)		Std. Deviation		Mean (per review)		Std. Deviation	
	No. of acts	% all acts	No. of acts	% all acts	No. of acts	% all acts	No. of acts	% all acts
Evaluative acts	16.52	82.89	5.58	11.59	23.46	81.62	9.97	10.25
Non-evaluative acts	3.16	17.11	1.92	11.59	5.33	18.38	3.59	10.25
Wilcoxon test	Evaluative acts vs. Non-evaluative acts: Z = -10582 <i>p</i> < 0.001				Evaluative acts vs. Non-evaluative acts: Z = -10.552 <i>p</i> < 0.001			
Mann-Whitney test	Chinese vs. English: <i>U</i> = 10101.500 <i>P</i> = .126							

Table 3: The occurrences of evaluative and non-evaluative acts in English and Chinese reviews

There have been ongoing debates regarding whether music criticism should be primarily evaluative or not (Thompson 1979; Levy 1987; Hoger 1992; Schick 1996), while musicological studies in music criticism have been largely descriptive/historical (Hoger 1992, p.14). However, it is apparent that music critics for both British and Hong Kong newspapers in the present day consider evaluation a core element in their reviews. As Schick (1996) stresses, “most criticism includes value judgements” in practice (p.22). Another possible explanation for a very low percentage of purely descriptive (non-evaluative) acts in both corpora is that it is difficult to separate evaluation and description (Levy 1987), as critics might embed evaluation in description and vice versa.

Positive and negative remarks in beginnings and endings

More reviews open and close with positive remarks than negative remarks (Table 4). Ha (2011) notices that in classical record reviews in English, most critics prefer to end the reviews on a positive note, and negative endings are rare in this review genre. Similar trends are observed in the study of English academic book reviews about business communication (Mackiewicz 2007). While criticisms at salient locations such as the beginning or ending of a review can be “particularly noticeable and memorable” (Mackiewicz 2007, p.202), a positive opening or closing remark can function as a mitigation strategy to soften the criticisms which appear in the review. Opening or closing a review positively can “address both ideational and interpersonal issues, expressing cognitive judgements... [and] redress the threat” the reviewed author might face (Hyland 2000, p.55). In the case of concert reviews, such a strategy might help save the face of the musicians/concert organisers and reduce the unfavourable effects caused by any criticism in the concert review.

		Number of English concert reviews (N = 150)	Number of Chinese concert reviews (N = 150)
Reviews opening remarks	Positive	95 (63.33%)	75 (50%)
	Negative	18 (12%)	7 (4.67%)
Reviews closing remarks	Positive	104 (69%)	94 (62.67%)
	Negative	40 (26.67%)	42 (28%)

Table 4: The occurrences and percentages of concert reviews opening or closing with evaluation

Dimensions of concert evaluation

Frequencies of *Concert Evaluation* commenting on aspects of the concert largely exceed frequencies of *Background Comments* in both corpora (Table 2). In fact, *Concert Evaluation* appears to be the most crucial component of the concert review, and can be found in every English and Chinese review in the corpora. It accounts for the biggest part of rhetorical acts in both English and Chinese reviews, occupying 79.03% and 70.94% of the total acts in each corpus respectively.

Praise and criticism of specific aspects of the concert significantly outnumber global comments. More than 98% of the English and Chinese concert evaluative acts are on specific aspects, and there is more global praise than criticism in the English and Chinese corpora. (See Figure 3)

According to Brown and Levinson's (1987) politeness theory, criticisms are potentially risky face-threatening acts, and researches reveal that writers of various disciplines have a tendency to redress their negative comments (Myers 1989; Gea Valor 2000; Hyland 2000; Hyland and Hyland 2001; Ha 2011). Condemning the entire work with global criticism is regarded as "a particularly face-threatening act" (Hyland 2000, p.48), and is therefore sparsely applied in concert reviews. As opposed to the findings of Hyland (2000) on academic book reviews, praise is most concerned with specific aspects of the concert (Figure 3). In a study of a similar genre to concert reviews, the evolution of record reviews (Ha 2011), a tendency of praise has been shown to be becoming more specific in recent years. Thus it is possible that in the present day, it is common practice in the music criticism industries of both cultures to criticise and praise the specific aspects of concerts.

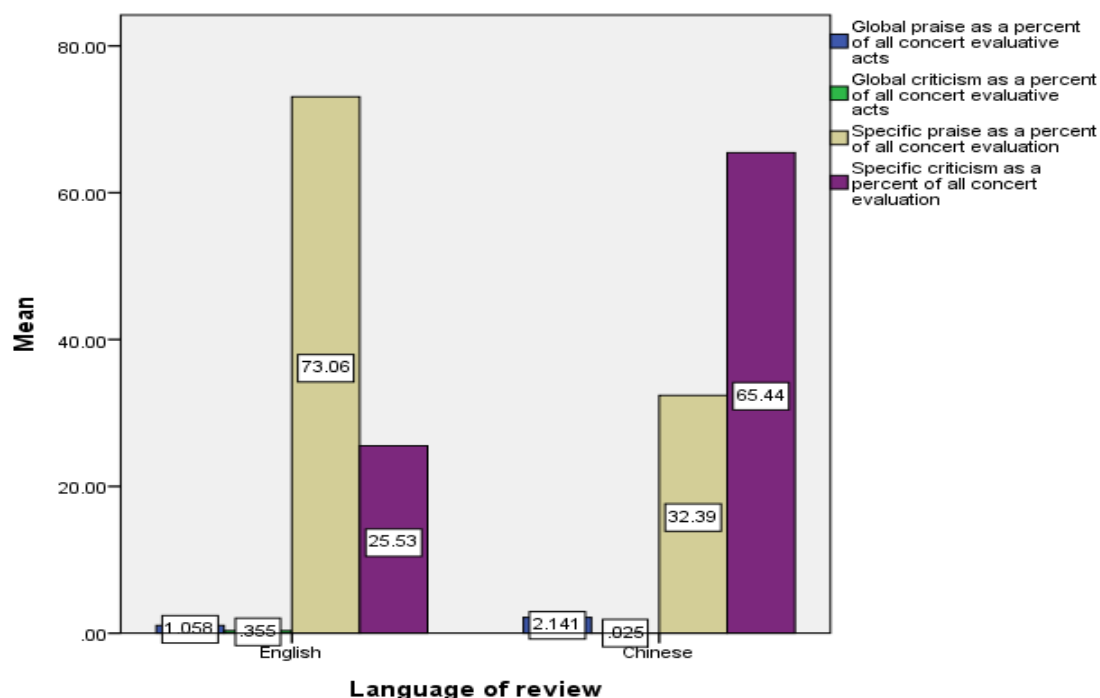


Figure 3: The distributions of global and specific evaluation in English and Chinese reviews

Regarding aspects of the concert, critics of both writing cultures seem to consider music to be their evaluative priority. Within this category, most attention has been given to the standard of performance, followed by the quality and selection of compositions (Figures 4 & 5).

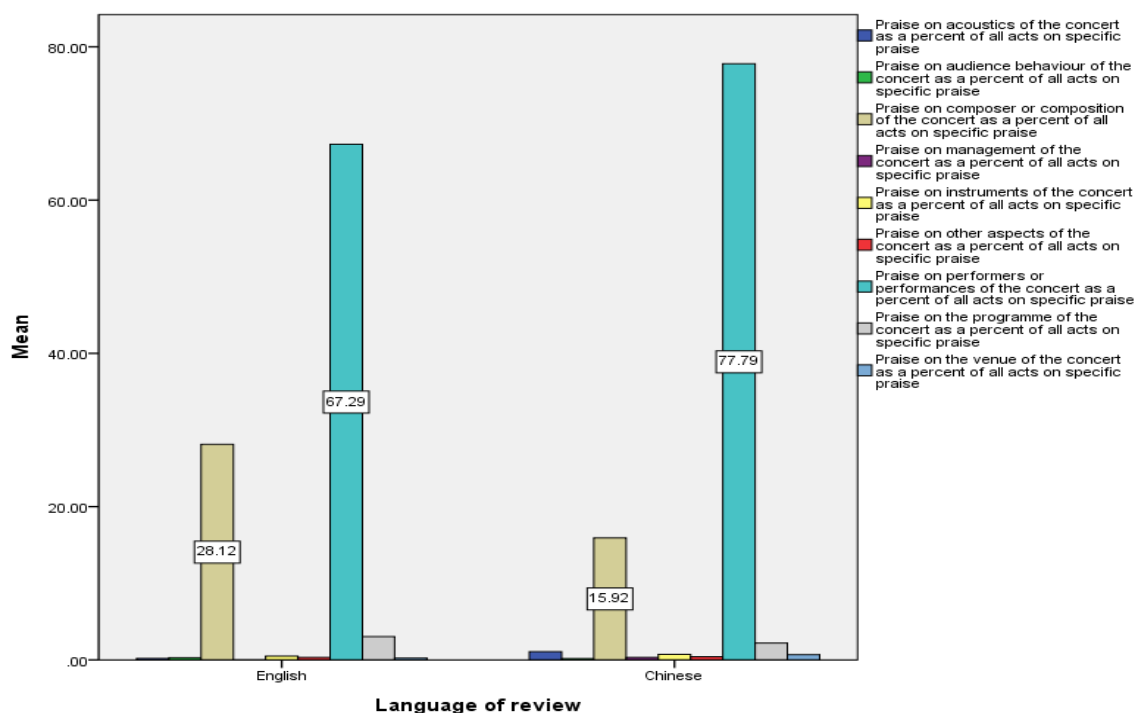


Figure 4: Distributions of praise on aspects of the concert

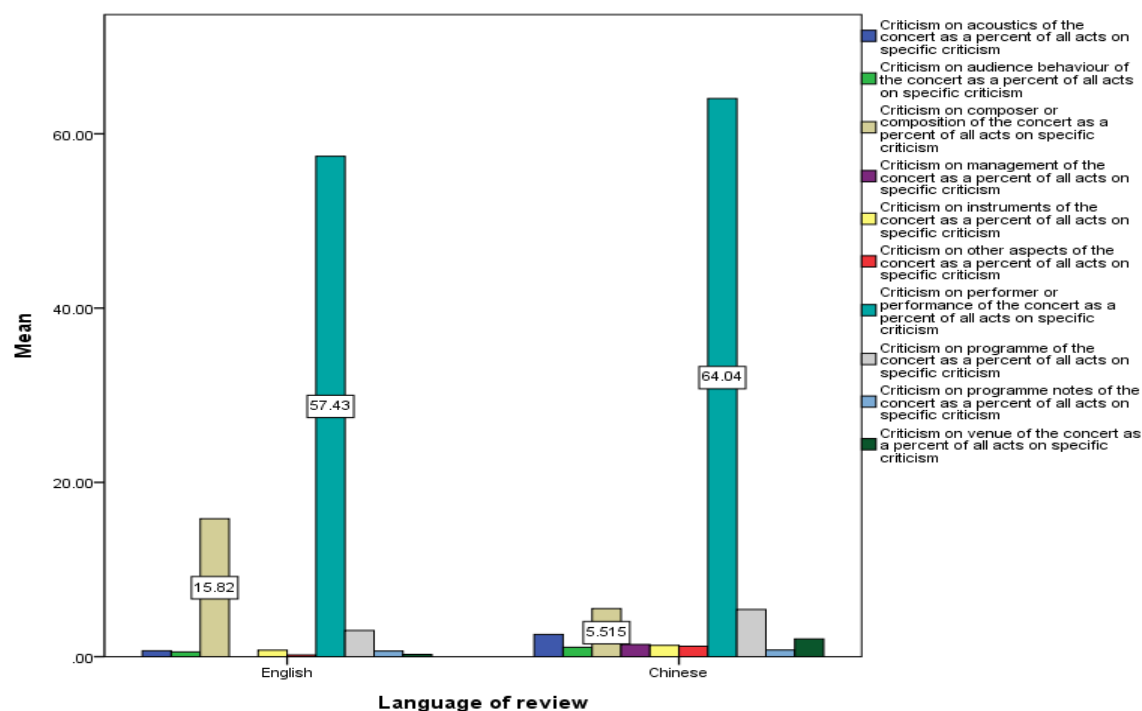


Figure 5: Distributions of criticism on aspects of the concert

Use of evaluative acts

Regarding details of praise and criticism acts, a number of similarities are observed across cultures: praise appears in more reviews than criticism; praise outnumbers criticism on average; praise is more straight-forward while criticism is more reserved.

First of all, praise of the concert appears in all English reviews. A single Chinese review contains no praise, focusing exclusively instead on criticism. Criticism, on the other hand, appears in fewer reviews than praise in both English and Chinese. 32 English reviews and 22 Chinese reviews do not contain criticism on aspects of the concert. In other words, approximately 20% English reviews and 15% Chinese reviews contain only praise about the concert.

Apart from the above phenomenon, as shown in Table 4, there is more praise than criticism in both English and Chinese reviews. On average, an English review contains 11.2 praise acts, which is nearly three times the average amount of criticism acts (4.4 per review). There are almost twice as many praise acts as criticism acts per Chinese review (13 and 7.4 respectively). It seems that music critics across cultures prefer to praise the positives rather than criticising shortcomings of concerts.

When taking a closer look at the concert evaluative acts, it is evident that both British and Hong Kong critics are rather generous in giving their praise. Almost three quarters of the praise acts in English and two-thirds of those in Chinese are either emphasised or unmitigated (Figure 6). On the other hand, critics in English and Chinese seem to be more reserved in giving criticism. Around two-thirds of the criticism acts in English and Chinese reviews are mitigated. Emphasised criticism, which reinforces the face-threatening act, appears least in both corpora (Figure 7). This indicates that both British and Hong Kong critics prefer to tone down their criticism than to give direct and strong criticism.

Brown and Levinson's (1987) theory of face-maintenance provides an insight into the above findings. Criticism is a face-threatening act and mitigating the criticism can lessen the potential FTA effects. Praise, on the other hand, plays into one's positive face of wanting to be accepted and approved of. Though there are risks in giving out praise, as "not everyone is entitled to compliment, and conveying praise implies an authority to appraise and make public one's judgements" (Hyland 2000, p.45), critics seem to be less reluctant in presenting positive comments about the concert.

In Moreno and Suárez's (2008) comparative study between English and Spanish reviews, some Spanish scholars do not review books that they think are very bad. It is possible that some music critics also hold the same attitude in reviewing concerts, which could be a reason for the relatively fewer instances of criticism in concert reviews. As part of a larger study, semi-structured interviews are scheduled to find out critics' perspectives towards evaluating concerts.

		English	Chinese
Number of concert evaluative acts per review		15.6	20.4
Number of praise acts vs. criticism acts per review	Praise	11.2	13.0
	Criticism	4.4	7.4
Most common praise act		Unmitigated	Mitigated
Most common criticism act (concert evaluation)		Mitigated	

Table 4: Comparing features regarding concert evaluative acts in English and Chinese reviews

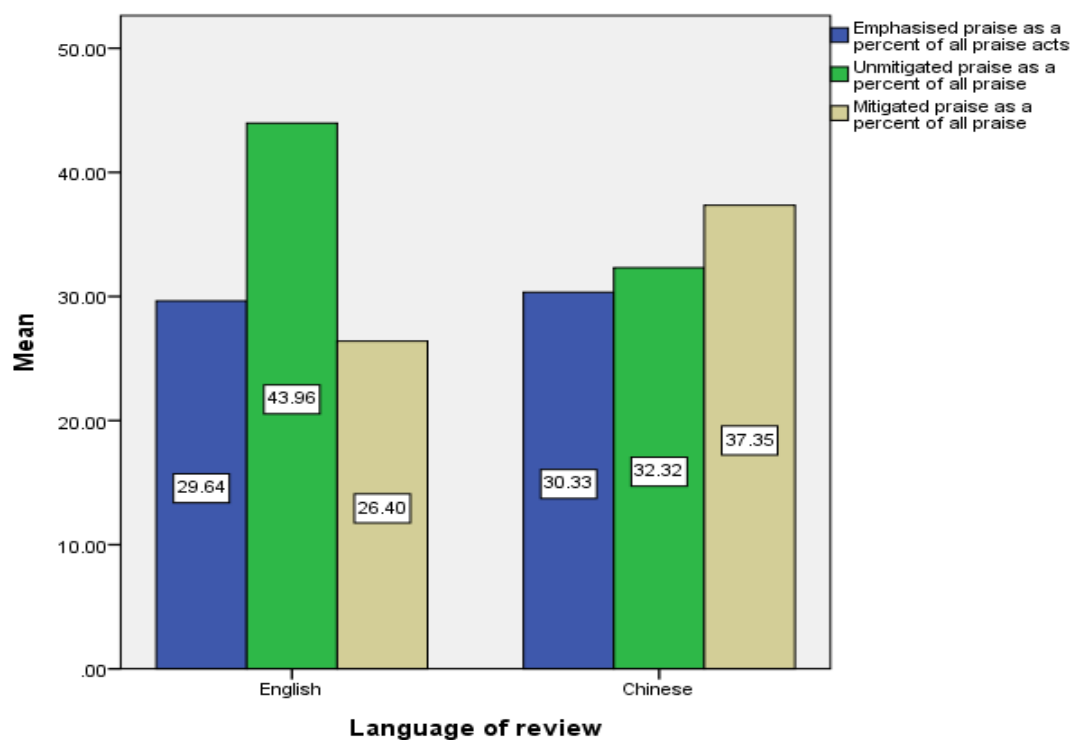


Figure 6: Distribution of types of praise acts in English and Chinese concert reviews

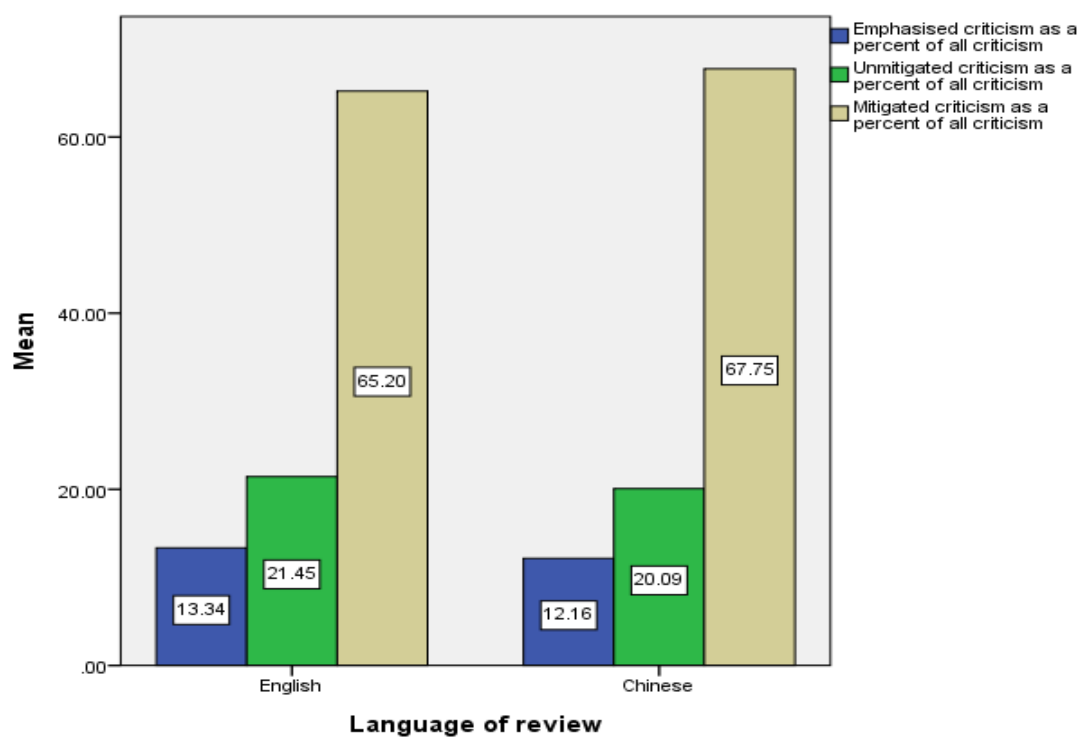


Figure 7: Distribution of types of criticism acts in English and Chinese reviews

The differences between English and Chinese reviews

Compared with the similar traits between English and Chinese reviews, differences are relatively few in their general trends. One difference is that Chinese reviews contain more instances of evaluation than English reviews (Table 4). Another difference is that only Chinese reviews contain evaluation on *Concert Management*, and that comments in this area are exclusively on the negative side (Figures 4 & 5). On average an English review has 356 words and a Chinese review has 1525 characters. There are ambiguities in segmenting Chinese words (Ge, Pratt et al 1999; Teahan, Wen et al. 2000), however, as single-character and two-character words are most frequent in Chinese language. Therefore, it might be possible to attribute the differences to a longer length in Chinese reviews, which allows more space for critics to evaluate various aspects of the concert.

British critics tend to praise more than criticise on both global and specific aspects of the concert. Hong Kong critics, though they praise more than criticise the concert as a whole, tend to give more criticism than praise on specific aspects of the concert (Figure 3). It might seem that Hong Kong critics are harsher than British critics. Nevertheless, as most criticism acts in Chinese reviews are mitigated (Figure 7), more analysis needs to be conducted before the conclusion can be drawn. In-depth interviews can be a useful tool to integrate the corpus-driven quantitative study, figure out critics' attitudes in evaluation and differences across the writing cultures.

Conclusion

The present study presents the quantitative textual analysis of British English and Hong Kong Chinese concert reviews. There are more similarities than differences in the overall rhetorical trends between the writing cultures. Both English and Chinese reviews appear to be more evaluative than informative. Quality of music at the concert is the focus of evaluation across both cultures. Critics in both writing cultures try to be positive in their reviews, one observed result of which is that a high quantity of reviews begin and end with compliments. There are also more incidences of praise than criticism in both reviews. The majority of the praise in both languages is either unmitigated or emphasised. On the other hand, both British and Hong Kong critics try to avoid condemning the concert as a whole. Most criticisms in English and Chinese reviews are mitigated.

The cross-cultural quantitative similarities suggest two insights into this research.

First, although it has been a controversial topic whether or not Brown and Levinson's (1987) politeness theory is suitable for Chinese culture (Gu 1990; Mao 1994; Ji 2000), this very influential model seems to explain well the general trends observed in this comparative study. As Hong Kong has a close historical connection with the UK, the relation may impose an influence on the writing culture of Hong Kong Chinese concert reviews. Alternatively, as this study only compares their rhetorical features on the surface level, it is possible that Brown and Levinson's model is only applicable to both cultures within the limited confines of this study, and would encounter unforeseen shortfalls in a more thorough analysis.

Second, it is possible that in today's globalised world, the professional cultures of music criticism in British and Hong Kong newspapers can be very similar. The similarities may therefore be the result of professional culture having a stronger impact than cultural differences between English and Chinese.

To conclude, though general trends suggest more similarities than differences between English and Chinese reviews, further detailed investigations need to be conducted to look into the way specific strategies are realised. As part of this larger-scale ongoing project, in-depth semi-structured interviews are scheduled with music critics in order to obtain a better explanation of the quantitative results.

References

- Vijay K. Bhatia (2008).** Towards Critical Genre Analysis. In Vijay K. Bhatia, John Flowerdew & Rodney H. Jones (eds.) *Advances in Discourse Studies*. Routledge: London, UK, pp166-177.
- Marina Bondi (2009).** Historians at Work: Reporting Frameworks in English and Italian Book Review Articles. In Ken Hyland & Giuliana Diani (eds.), *Academic Evaluation: Review Genres in University Settings*. Palgrave MacMillan: Basingstoke, Hants, UK, pp179-96.
- Penelope Brown & Stephen C Levinson (1987).** *Politeness: Some universals in language usage*. Cambridge University Press: Cambridge, UK.
- Jonathan Culpeper, Derek Bousfield & Anne Wichmann (2003).** Impoliteness revisited: With special reference to dynamic and prosodic aspects. In *Journal of Pragmatics* 35, pp1545-79.
- Giuliana Diani (2009).** Reporting and evaluation in English book review articles: A cross-disciplinary study. In Ken Hyland & Giuliana Diani (eds.), *Academic*

- evaluation: Review genres in university settings*. Palgrave MacMillan: Basingstoke, Hants, UK, pp87-104.
- Andrew Duncan (1773)**. *Introduction Medical and Philosophical* Vol. 1, Part 1. John Murray: London, UK, pp6-9.
- Xianping Ge, Wanda Pratt & Padhraic Smyth (1999)**. Discovering Chinese words from unsegmented text (poster abstract). Paper presented to the *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*, 1999.
- Yueguo Gu (1990)**. Politeness phenomena in modern Chinese. In *Journal of Pragmatics* 14, pp237-57.
- Olivia Fong-wa Ha (2011)**. A diachronic study of music criticism: the case of record reviews. In Francoise Salager-Meyer & Beverley A. Lewin (eds.), *Crossed Words: Criticism in Scholarly Writing*. Peter Lang: Bern, Switzerland, pp339-64.
- Elizabeth Anne Hoger (1992)**. *Writing in the discipline of music: rhetorical parameters in writings about music criticism*. PhD Dissertation. Purdue University: West Lafayette, IN, USA.
- Janet Holmes (1988)**. Doubt and certainty in ESL textbooks. In *Applied linguistics* 9, pp21-44.
- Russell A Hunt & Douglas Vipond (1986)**. Evaluations in literary reading. In *Text-Interdisciplinary Journal for the Study of Discourse* 6, pp53-72.
- Ken Hyland (2000)**. *Disciplinary Discourses: Social Interactions in Academic Writing*. Longman: Harlow, UK.
- Ken Hyland (2004)**. Disciplinary interactions: Metadiscourse in L2 postgraduate writing. In *Journal of Second Language Writing* 13, pp133-151.
- Ken Hyland & Diani Giuliana (2009)**. *Academic evaluation: review genres in university settings*: Palgrave Macmillan: Basingstoke, Hants, UK.
- Ken Hyland & Polly Tse (2009)**. 'The leading journal in its field': evaluation in journal descriptions. In *Discourse Studies* 11, pp703-720.
- Fiona Hyland & Ken Hyland (2001)**. Sugaring the pill: Praise and criticism in written feedback. In *Journal of Second Language Writing* 10, pp185-212.
- Hiroko Itakura & Amy Tsui (2011)**. Evaluation in academic discourse: Managing criticism in Japanese and English book reviews. In *Journal of Pragmatics* 43, pp1366-1379.
- Hiroko Itakura (2012)**. Hedging praise in English and Japanese book reviews. In *Journal of Pragmatics* 44.
- Shaojun Ji (2000)**. Face and polite verbal behaviors in Chinese culture. In *Journal of Pragmatics* 32, pp1059-1062.
- Donna M Johnson (1992)**. Compliments and politeness in peer-review texts. In *Applied Linguistics* 13, pp51-71.
- Janet M Levy (1987)**. Covert and casual values in recent writings about music. In *Journal of Musicology*, pp3-27.
- Beverly A Lewin (2005)**. Contentiousness in Science: the Discourse of Critique In Two Sociology Journals. In *Text* 25, pp723-744.
- Jo Mackiewicz (2007)**. Compliments and criticisms in book reviews about business communication. In *Journal of Business and Technical Communication* 21, pp188-215.
- LuMing Robert Mao (1994)**. Beyond politeness theory: 'Face' revisited and renewed. In *Journal of Pragmatics* 21, pp451-486.

- Ana I Moreno & Lorena Suárez (2008).** A study of critical attitude across English and Spanish academic book reviews. In *Journal of English for Academic Purposes* 7, pp15-26.
- Greg Myers (1989).** The pragmatics of politeness in scientific articles. In *Applied Linguistics* 10, pp1-35.
- Patricia A. Rockwell (2006).** *Sarcasm and other mixed messages: The ambiguous ways people use language*, Edwin Mellen Press: Lewiston, NY, USA.
- Pamela Rogerson-Revell (2007).** Humour in business: A double-edged sword: A study of humour and style shifting in intercultural business meetings. In *Journal of Pragmatics* 39, pp4-28.
- Stanley Sadie (ed.) (1980).** *The New Grove Dictionary of Music and Musicians* (1st Edn.). Macmillan: London, UK.
- Françoise Salager-Meyer, María Ángeles Alcaraz Ariza & Maryelis Pabón Berbesí (2007).** Collegiality, critique and the construction of scientific argumentation in medical book reviews: a diachronic approach. In *Journal of pragmatics* 39, pp1758-1774.
- Robert D. Schick (1996).** *Classical Music Criticism*. Garland: New York, NY, USA.
- Percy A. Scholes (1970).** *The Oxford companion to music*. Oxford University Press: Oxford, UK.
- Philip Shaw (2004).** How do we recognise implicit evaluation in academic book reviews? In *Academic discourse: new insights into evaluation*. Peter Lang: Bern, Switzerland, pp121-40.
- Hilkka Stotesbury (2006).** Gaps and False Conclusions: Criticism in Research. In *Academic Discourse Across Disciplines* 42, pp123-48.
- William John Teahan, Yingying Wen, Rodger McNab & Ian H Witten (2000).** A compression-based algorithm for Chinese word segmentation. In *Computational Linguistics* 26, pp375-93.
- Oscar Thompson (1979).** *Practical musical criticism*. Da Capo Press: Boston, MA, USA.
- M. Lluïsa Gea Valor (2000).** *A pragmatic approach to politeness and modality*. Publicacions de la Universitat de Valencia: Valencia, Spain.

13

Do EFL Learners Discriminate Text-Level Main Ideas, Paragraph-Level Main Ideas, and Details During Reading? A Study With Self-Paced Reading

Yoshinobu Mori
University of Tsukuba
s1530038@u.tsukuba.ac.jp

Background

Successful reading comprehension requires the understanding of main ideas, or the information that is important to comprehend the gist of the text (e.g., Grabe & Stoller, 2011). Texts include various levels of information that differ in importance such as supporting details, main ideas summarizing a paragraph, and main ideas summarizing the whole text, with a superordinate text idea integrating the subordinate ideas. Such links among text ideas form a hierarchical structure (e.g., van Dijk & Kintsch, 1983). As a first step to establish mental representations in which the above text ideas are coherently connected to one another, it is necessary to identify the position each text idea holds in the hierarchical structure. Furthermore, to smoothly construct such coherent representations, readers need to discriminate the different levels of text information during their reading based on its relative importance.

Despite the importance of understanding the main idea, in second language (L2) reading, it may be difficult to prioritize more important information above the less important information, compared to first language (L1) reading (Miller & Keenan, 2011). Additionally, it can be seen that there is insufficient research on main idea comprehension during reading (e.g., Cirilo & Foss, 1980). Thus, it is worth investigating whether L2 readers can regulate their reading process based on the relative importance of text-level main ideas, paragraph-level main ideas, and the supporting details.

Past studies have examined main idea comprehension using post-reading tasks. Details are often excluded and the main ideas are included in summaries of L1 university students as well as L1 children (e.g., Brown &

Day, 1983; Williams et al., 1981) and L2 readers (e.g., Kim, 2001). These studies indicated that readers, including the readers of English as a foreign language (EFL), selected the main ideas to be included in the summaries. Additionally, employing an importance rating task, Ushiro et al. (2008) demonstrated that EFL readers rated main ideas as being more important than the supporting details. To directly tap the activation of text information in the mind, Mori (2015a) conducted a sentence recognition task after the EFL learners read each paragraph. The participants tended to recognize the main ideas more accurately and quickly than the supporting details, indicating that the EFL readers selectively activated main ideas more strongly than the supporting details in their minds. The above findings support the idea that L1 and L2 readers can discriminate between the main ideas and the details in post-reading tasks.

In contrast, insufficient studies have dealt with main idea comprehension during reading. Among the limited studies, Cirilo and Foss (1980) revealed that self-paced reading times for target sentences were longer when the sentences appeared as to be of a higher level information as compared to a lower level information in the hierarchical text structure. Moreover, Yeari et al. (2015), who employed eye-tracking measurements, elucidated the importance of text information in the reading process. Specifically, the first-pass reading times were longer for the main ideas than for the details. These studies reasoned that readers needed more time to integrate main ideas with their existing representations because main ideas typically introduce new topics at the beginning of a textual unit. On the other hand, readers needed less time for the supporting details since they are often written detailing the preceding context. However, it should be noted that the aforementioned studies only dealt with L1 undergraduates and further research on L2 is needed. Since L2 readers allot much cognitive resource to lower level processes (e.g., word recognition and syntactic parsing) unlike L1 readers, they may fail to allot sufficient cognitive resources towards higher level processes such as constructing globally coherent representations.

Among the various factors affecting main idea comprehension, one of the most important factors to consider is reading proficiency, especially in L2 contexts, since it is very limited in L2 reading. Several studies showed that good readers exceeded poor readers in main idea comprehension (e.g., Johns & Mayes, 1990; Winograd, 1984). From the perspective of text difficulty, Kim (2001) observed less successful main idea comprehension by Korean EFL readers in a text including difficult vocabulary. In contrast, other studies did not find the effect of reading proficiency on main idea comprehension

(e.g., Riley & Lee, 1996; Ushiro et al., 2009). Ushiro et al. (2009) used an easy text so that their participants would understand the contents without difficulty. Thus, reading proficiency is thought to only affect text comprehension, and not the main idea comprehension. Based on past research, the present study also used texts that were easy for EFL readers to comprehend in order to exclude the possible effect of difficulty in text decoding.

To summarize, past studies have yet to shed light on main idea comprehension in L2 reading whereas the L2 main idea comprehension has been explored in post-reading tasks (e.g., Kim, 2001; Mori, 2015a; Ushiro et al., 2008). Moreover, L2 readers are more likely to pay sufficient attention to main idea comprehension in easy texts. Therefore, the present study aimed to examine whether Japanese EFL readers could identify the varied importance of text-level main ideas, paragraph-level main ideas, and supporting details during reading, employing the self-paced reading measurement based on Cirilo and Foss (1980). The research question (RQ) was set as follows:

Do EFL learners pay more attention to text main ideas and less to details, compared to paragraph main ideas, during reading?

According to Cirilo and Foss (1980) and Yeari et al. (2015), it is predicted that the reading times are longer for text-level main ideas, paragraph-level main ideas, and supporting details in this order if EFL readers appropriately understand the varied importance of these text ideas. However, it should be also noted that self-paced reading times can reflect other cognitive processes apart from the targeted process (Haberlandt, 1994). Accordingly, this study conducted the importance rating task to confirm whether the ratings of the text information show a similar tendency to that of the reading times. That is, the importance ratings were used to confirm whether the participants regulated their reading times based on the importance of the text information.

Method

Participants

Fifty-three Japanese-speaking undergraduate and graduate students participated in the present experiment. All the participants had studied EFL for more than six years as a normal part of their formal education in Japan. Their majors were comprehensive human sciences, humanities and culture, humanities and social sciences, human sciences, informatics, life and environmental sciences, medicine and medical sciences, pure and applied

sciences, science and engineering, social and international studies, and systems and information engineering.

Materials

Six expository texts were adopted from Mori (2015b). These texts consisted of one text-level main idea, three paragraph-level main ideas, and other supporting details. The text main idea was written in the beginning of the texts, and the paragraph main ideas were written in the beginning of the paragraphs except in the first paragraphs: In the first paragraphs, the paragraph main ideas followed the text main ideas. A sample (“Energy” text) is as follows (The text main idea is boldfaced and the paragraph main ideas are underlined):

Electronic energy production from oil has major problems. The danger of producing electric energy is increasing. There is not as much oil left as before. It is necessary to drill even deeper to recover oil. It means greater danger to the workers. The development of nuclear power reflected the limited amount of oil.

The decreasing availability has made oil more expensive. It costs more money to produce oil. This situation puts the oil seller in a powerful position. Increasing costs of oil are the strongest motivation to seek alternative energy sources.

Our energy habits have damaged the environment. Burning oil sends out various particles into our atmosphere. Such particles include ozone and other gases. In the US, nearly half of the population lives in areas which have dirty air. Moreover, accidents with oil tankers did a lot of damage to ocean ecologies. One example is the tanker accident in Alaska.

Mori (2015b) revised the texts so that the EFL readers would not have difficulty in comprehension. Unfamiliar words to Japanese EFL university students were identified in a pilot study and paraphrased with high-frequency words. Additionally, all the texts had an almost equal and low readability level so that the EFL readers would not have difficulty in understanding them. Table 1 shows the features of the texts.

Text	Word	Flesch-Kincaid grade level	Sentence
Argentina	147	7.8	17
Distance	153	7.4	13
Energy	150	8.6	16
Environment	151	7.6	10
Support	155	8.2	17
Three Mile Island	148	8.5	14

Table 1: Features of the texts

Procedure

The participants were tested individually. First of all, the participants were given a general explanation of the experiment. They were then instructed to read each sentence onscreen for comprehension at their own pace by pressing a button on a response pad (Cedrus, model: RB-730). The author used SuperLab version 4.5 for windows to record the reading time for each sentence. The order of the texts was randomized. After reading all the texts, the participants rated the relative importance of each sentence in each text using a 5-point Likert scale ranging from 1 (not important) to 5 (important).

Data Analysis

The data for the two-line sentences on the screen were excluded from the following analyses. As most sentences of the “Environment” text were two lines on the screen, data from this text was excluded from the analyses. This is because the reading times for these sentences would reflect not only the reading comprehension but also saccade from the end of the first line to the beginning of the second line. The reading times greater than 2.5 of the standard deviation from each participant’s mean were deleted from the analysis. The data treatment deleted 0.26% of the data (8 of 3021 data points). The reading times were divided by the number of syllables. To examine whether the participants distinguished between the importance of text information while reading and the rating task, two one-way repeated measures analyses of variance (ANOVAs) were conducted for the importance ratings and reading times per syllable. Text information (Text-level main idea, Paragraph-level main idea, and Supporting detail) was a within-participant variable. We used an alpha level of .05 throughout the study.

Results

Table 2 and Figure 1 show the results of the importance ratings. The one-way repeated measures ANOVA yielded a significant effect of Text

information, $F(1.69, 88.11) = 88.36, p < .001, \eta^2 = .54$. Multiple comparison with the Bonferroni adjustment revealed that the importance ratings were significantly higher for the text main ideas than for the paragraph main ideas, $p = .002, M_{diff} = 0.35, 95\% \text{ CI } [0.12, 0.59]$, and significantly higher for the paragraph main ideas than for the supporting details, $p < .001, M_{diff} = 0.91, 95\% \text{ CI } [0.71, 1.11]$. This result confirmed that the participants regarded the text main ideas, paragraph main ideas, and supporting details more important, in this order. In other words, they were able to tell the differing importance of the text information in accordance with the hierarchical levels.

Text information	M	SD
Text main idea	4.08	.09
Paragraph main idea	3.73	.05
Detail	2.82	.06

Table 2: Descriptive statistics of the importance ratings

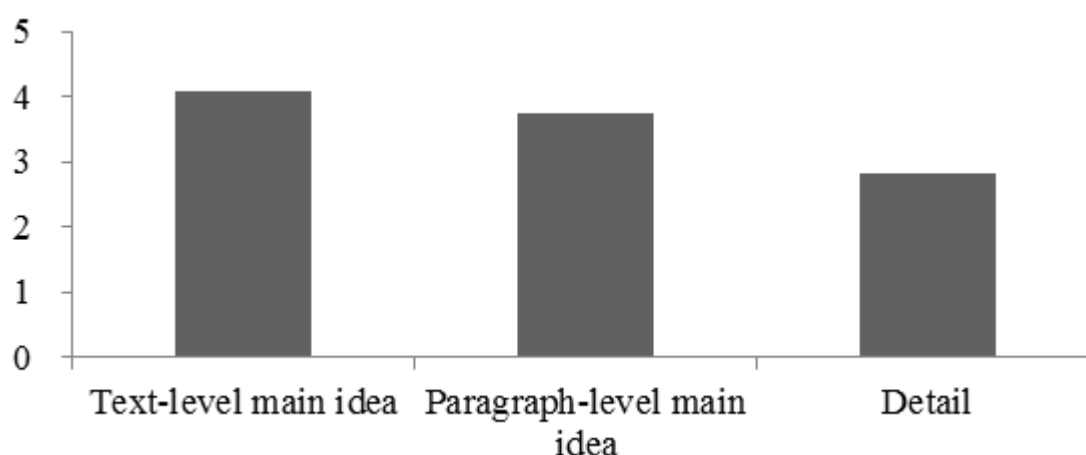


Figure 1: Mean importance ratings

Table 3 and Figure 2 show the results of the reading times per syllable. The one-way repeated measures ANOVA yielded a significant effect of Text information, $F(1.23, 63.97) = 34.54, p < .001, \eta^2 = .10$. Multiple comparison with the Bonferroni adjustment demonstrated that the reading times per syllable were significantly longer for the text-level main ideas than for the paragraph main ideas, $p < .001, M_{diff} = 104.09, 95\% \text{ CI } [57.71, 150.47]$, and significantly longer for the paragraph main ideas than for the supporting details, $p = .009, M_{diff} = 23.72, 95\% \text{ CI } [4.99, 42.46]$. This result indicated that the participants paid more attention to the text main ideas, paragraph main ideas, and the supporting details in accordance with their importance.

Text information	M	SD
Text main idea	544.75	222.20
Paragraph main idea	440.66	137.28
Detail	416.93	125.43

Table 3: Descriptive statistics of the reading times (millisecond) per syllable

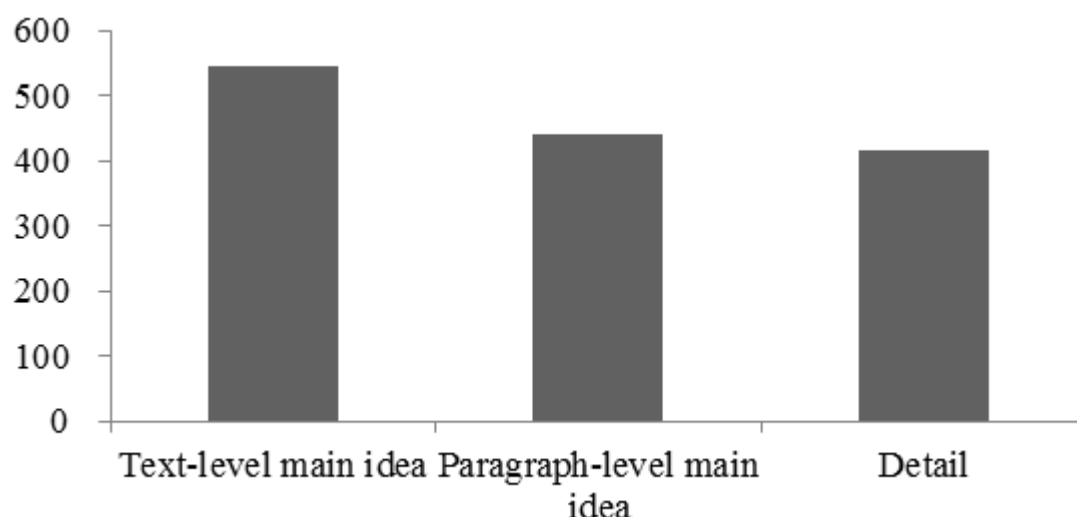


Figure 2: Mean reading times (millisecond) per syllable

Discussion

Statistical analyses for the importance ratings and reading times per syllable demonstrated that the Japanese EFL readers considered the text-level main ideas as more important, followed by the paragraph-level main ideas, and the supporting details and they spent longer times reading these text ideas in the same order. Addressing the RQ, the EFL readers distinguished the relative importance of these text ideas while reading when they do not have trouble in text comprehension (Ushiro et al., 2009).

While reading the first sentence of the texts, they may have identified these sentences as the text-level main ideas, relying on their location and meanings. That is, the main ideas often appear at the beginning of a textual unit, introducing and overviewing the unit (Cirilo & Foss, 1980; Yeari et al., 2015). The EFL readers in this study may have expected sentences to be connected with the incoming sentences, resulting in greater attention (i.e., slowing down their reading). On the other hand, when encountering the initial sentences of the paragraph, the participants probably identified the sentences as the paragraph-level main ideas merely from their meanings. The readers were likely to notice the new topics introduced in the sentences (Cirilo & Foss, 1980; Yeari et al., 2015). Thus, it took time to integrate the paragraph-level main ideas with their existing mental representations. At the

same time, the EFL learners may have paid attention to the paragraph-level main ideas to be linked with the following details. It should be noted that the readers were not able to perceive paragraph borders otherwise indicated as indentation since this study adopted the sentence-by-sentence reading method.

As for the text-level main ideas there is another possibility. The EFL readers might have paid attention to the text-level main ideas without considering their importance only because they were located at the beginning of the text. However, this possibility is less valid. The result of the importance ratings confirmed that the EFL readers regarded the text-level main ideas as significantly more important than the paragraph-level main ideas. Furthermore, if the participants had completely relied on locational information, they would have failed to distinguish the importance of the paragraph-level main ideas and the supporting details, which would have resulted in equal reading times for these text ideas. Therefore, although the EFL learners might have relied on the locational information of the main ideas, they were not completely dependent on it. Rather, they probably took into consideration the relative importance indicated by their meanings as well.

It is true that the locational information and meanings are difficult to separate in main idea comprehension because important information is usually placed at the beginning of the textual units. Thus, further studies are required to individually examine the effects of location and the meanings of text ideas on main idea comprehension during L2 reading.

The findings of this study have pedagogical implications. In order to allow EFL readers to engage in main idea comprehension while reading, simple and easy texts are appropriate. Using such texts, the learners can concentrate on text contents, and not be distracted by lower-level processes. This study suggests that EFL readers could regulate their reading process in accordance with the importance of text information in easy texts. Thus, teachers should sometimes let them read such simple and easy texts on their own without giving instructions. It may prompt self-regulated reading and lead to cultivating autonomous readers in the long term.

References

- Ann L. Brown & Jeanne D. Day (1983).** Macrorules for Summarizing Texts: The development of expertise. In *Journal of Verbal Learning and Verbal Behavior* 22, pp1-14.
- Randolph K. Cirilo & Donald J. Foss (1980).** Text Structure and Reading Time for Sentences. In *Journal of Verbal Learning and Verbal Behavior* 19, pp96-109.
- William Grabe & Fredricka L. Stoller (2011).** *Teaching and Researching Reading*. Pearson Education: Harlow, UK.
- Karl Haberlandt (1994).** Methods in Reading Research. In Morton A. Gernsbacher (ed), *Handbook of Psycholinguistics*. Academic Press: London, UK, pp1-32.
- Ann M. Johns & Patricia Mayes (1990).** An Analysis of Summary Protocols of University ESL Students. In *Applied Linguistics* 11, pp253-271.
- Sung-Ae Kim (2001).** Characteristics of EFL Readers' Summary Writing: A study with Korean university students. In *Foreign Language Annals* 34, pp569-581.
- Amanda C. Miller & Janice M. Keenan (2011).** Understanding the Centrality Deficit: Insight from foreign language learners. In *Memory & Cognition* 39, pp873-883.
- Yoshinobu Mori (2015a).** Do EFL Readers Use Macrorules in Their Mind Without the Summary Task? Evidence from activation of main ideas in the sentence recognition task. In *Kantokoshinetsu Association of Teachers of English Journal* 29, pp85-98.
- Yoshinobu Mori (2015b).** Relationship of Main Ideas with Subordinate and Superordinate Text Information in EFL Readers' Mental Representations. In *Annual Review of English Language Education in Japan* 26, pp349-364.
- Gail L. Riley & James F. Lee (1996).** A Comparison of Recall and Summary Protocols as Measures of Second Language Reading Comprehension. In *Language Testing* 13, pp173-187.
- Yuji Ushiro, Akari Kai, Chikako Nakagawa, Fuyumi Watanabe, Yuko Hoshino & Haruka Shimizu (2009).** Effects of Reading Perspective on EFL Learners' Summary Writing and Importance Rating. In *Annual Review of English Language Education in Japan* 20, pp11-20.
- Yuji Ushiro, Chikako Nakagawa, Akari Kai, Fuyumi Watanabe & Haruka Shimizu (2008).** Construction of a Macroproposition From Supporting Details: Investigation from Japanese EFL reader's summary and importance rating. In *The Japan Association of College English Teachers Journal* 47, pp111-125.
- Teun A. van Dijk & Walter Kintsch (1983).** *Strategies of Discourse Comprehension*. Academic Press: New York, NY, USA.
- Joanna P. Williams, Maravene B. Taylor & Sonia Ganger (1981).** Text Variations at the Level of the Individual Sentence and the Comprehension of Simple Expository Paragraphs. In *Journal of Educational Psychology* 73, pp851-865.
- Peter N. Winograd (1984).** Strategic Difficulties in Summarizing Texts. In *Reading Research Quarterly* 19, pp404-425.
- Menahem Yeari, Paul van den Broek & Marja Oudega (2015).** Processing and Memory of Central Versus Peripheral Information as a Function of Reading Goals: Evidence from eye-movements. In *Reading and Writing* 28, pp 1071-1097.

14 Language-specificity of forensic linguistics: the example of Hungarian

Eszter Tóth

University of Szeged, Hungary

esztertoth.04@gmail.com

Introduction

This paper is about forensic linguistics in Hungary, especially the analytical methods. It addresses the question of whether those methods are universal, so they can work in other languages, or whether they (or some of them) are unique, which means they are specific to languages. I would like to demonstrate some unique and universal examples, which can be useful to investigate.

Forensic linguistics is a really broad discipline whose subjects range from legal documents through trademarks up to author identification (Kontra, 2006). The term 'forensic linguistics' was first used by a Swedish linguist, Jan Svartvik in 1968 (Svartvik, 1968). He analysed the statement of Timothy John Evans, a man hanged in the 1950s for murder but who was later found to be innocent. That was in the late 60's but in Hungary, it was not until the 1970's and 1980's that linguists started to talk about the application of linguistics to criminal issues (Nagy, 1977). The only comprehensive work was written in 1980 by Ferenc Nagy (Nagy, 1980). That book presents several points of view on the topic, adding practical examples of the analytical methods. But it is now 35 years old and since then only brief studies have turned up. Some areas have priority over others because they are more researchable or there are more exact methods available for them. There are also problems for example with the terminology, the field's negative domestic appeal, and the misconception that linguists are graphologists.

Within forensic linguistics, I will deal with author identification. In the following, I would like to compare the international methods to the Hungarian ones. I will analyse Nagy's approach, and see which techniques are universal, and which are unique, dealing only with qualitative methods.

Idiolect as a central notion

On the first pages of his book Nagy characterises idiolect: “We are always talking about the individual styles of writers and poets, so one may assume that every person has their individual use of language even if they do not write to thousands of readers” (1980: 7). This is also a central notion in international methods. As Coulthard writes some decades later, “idiolect is a distinct and individual version of language. But it only becomes useful if an idiolect feature repeats itself, either within one text or across several texts by the same author” (2004: 432). Theoretically then, author identification is based on the concept of an idiolect as a distinct and individual version of language, but practically it is important that those distinctive patterns can be found consistently in the texts, because only consistent patterns can be detected. So we need to start from two important concepts: consistency and distinctiveness.

There are problems with analysing idiolects in the field of forensic linguistics, such as:

- a) how to measure (because of the shortness of the texts);
- b) how to determine distinctiveness;
- c) how to deal with the variability of natural languages, since individuals are not always consistent;
- d) how to extract the idiolect out of the texts.

Consistency and distinctiveness may be evidence that an idiolect exists, however mere observation and description of these features is not a theory of idiolect. Theories have to have explanatory power (Grant, 2010: 521).

Author identification: terms and methods in Hungary

As I mentioned above, I deal with author identification, analysing anonymous letters. Nagy calls it forensic text analysis, and it has two subcategories: the first of which is author identification (Nagy's term), and author characterization, (what Nagy calls 'limitation').

I will first look at author identification or one may call it document authentication or similarity detection. In this case texts are compared in two ways:

- a) Anonymous letters are compared with each other, and the question posed, whether it is possible that two or more anonymous letters have the same author.
- b) Anonymous letters are compared with known letters, and the question posed, whether it is possible that the anonymous letters and the known letters have the same author.

Secondly, Nagy discusses author characterization (sociolinguistic profiling) or, in his own terminology 'limitation' (of the possible authors). In that case, only the incriminating documents are available, we infer from them class characteristics like gender, age, education, social background or even profession. As Nagy points out: "Forensic text analysis is based upon the experience that the language which is common and uniform to everyone may show differences by social background, age groups, and quality and nature of education" (1980). The reason that these subcategories (author identification and characterization) are separated is the methodology; Nagy has two different analytical techniques. Here I will talk about just the first one, author identification.

According to Nagy, the most effective way to analyse texts is to do it with a complex analytical method. In this case, there are five aspects one may consider:

- 1) composition style
- 2) syntax
- 3) vocabulary
- 4) morphology
- 5) spelling

These aspects are not unknown since these are the methods, which are usually tested elsewhere, too. Analysis of the syntactic structure, the morphological inflections, the vocabulary and the spelling also appears in the international literature. Some papers (Kašić–Đorđević, 2010; Chaski, 2001) mention coherence and cohesion, which is 'composition style' in Nagy's terminology. Some major categories will be identical but with variations due to the different language systems. And two of them will only be mentioned in the Hungarian practice. There is also punctuation. Nagy does not mention it specially, but in the following section I will present a case where it was helpful.

Punctuation

In this case, I obtained four anonymous threatening letters to analyse. The question was whether they were from the same writer. The complex analytical method established that most likely one person had written all of the four letters, and thanks to punctuation, I could determine the order of arrival.

The first one says:

„Te bunkó! Menj innen a büdös kurva anyádba!”
“You jerk! Get out_of_here the stink whore your_mother_to!”

“You tool! Get out of here you mother fucker!” – there is only one exclamation mark.

The second says:

„Ne is próbálkozz! És takarodj te is innen!”
“No also try! And get you also out_of_here!”

“Do not dare to try! And get out of here!”

The difference between letters 1 and 2 is that letter one contains only some exclamation marks, but in letter 2 every sentence ends with it.

The third says:

„Nem kell a falunak a szeméttelp!!!”
“No need the town_to the rubbish-shoot!!!”

“This town doesn’t need the rubbish-shoot!!! – 3 exclamation marks.

The last one says:

„Tegyé le a szeméttelpről!!!!!! Felfogtad?????”
“Abandon the rubbish-shoot_from!!!!!! You_understand?????”

“Don’t you dare think about the rubbish-shoot!!!!!! Do you understand?????”
– 6 exclamation marks and 5 question marks.

The unknown author put more and more punctuation marks as he/she was getting more and more frustrated. The examples and the punctuation speak for themselves.

Composition style

The first of Nagy's aspects, composition style, may be new for those who read the literature in English, but it may be familiar from the pedagogic practice when teachers grade children's papers. They consider the coherence and cohesion of the text; that the wording is adequate and accurate for the topic etc. Nagy (1980: 64) says that a text with clear, consistent, on-topic wording and precise phrases apparently differs from a text with language frugality, inconsistency, scurrility and bad spelling.

Syntax

In this case, we analyse full stops, capital letters, disrupted structures or embedded structures (Eagleson, 1994), but in this paper the focus is word order. Compared to the basic English word order (SVO), Hungarian word order is variable. Hungarian is an agglutinating language with suffixes; these elements help to show the words' functions in the sentence. In English, word order defines the functions in the sentence. In Hungarian, thanks to the morphosyntactical elements, word order can convey other information, like topic and focus.

There is a syntactically assigned focus position in Hungarian sentences. It is important because if we analyse Hungarian texts, we will have to consider the order of the words, because it may be idiosyncratic. In the following case I analysed four letters, two anonymous letters and two known letters. It was my task to tell whether all the anonymous letters had the same author as the known ones. Using the complex analytical method, the word order contributed to the analysis. The anonymous letters had various syntactic structures but the known letters had not. The letters in question had sentences with Subject-Object-Verb order:

Ezen levelünket több médianak is megküldjük.
This our_letter-ACC several media_to also send_we.

We also send our letter to several media.

Nemleges válaszuk esetén a leírtakat végrehajtjuk.
Negative their_respond in_case_of the written_things-ACC accomplish_we.

If your answer is negative, we will execute what we have written above.

Subject-Verb-Object order:

Mi nem kímélünk senkit sem.
we no have_mercy_on_we nobody-ACC not_either.

We have no mercy on anybody.

Sokan féltik az életüket.
many fear the their_life-ACC.

Many fear for their lives.

Mi meghoztuk döntésünket.
we brought our_decision-ACC.

We made our decision.

Object-Subject-Verb order:

Az életét mindenki félti.
the his/her_life-ACC everyone fears.

Everyone fears for his/her life.

Verb-Subject order:

Nincs könyörület bennünk.
is_not mercy in_we.

We have no mercy.

By contrast, in the known letters, the same syntactical structure was discoverable. 90% of the sentences had the Subject-Verb-Object order:

A rendőrség a mai napig [...] vizsgáltatja a szőrös fenekemet!
the police the today until [...] have_examined the hairy my_butt-ACC!

The police have still examined my hairy butt [...]!

Ezek az emberek [...] besározták az egyenruhájukat!
these the men [...] tainted the their_uniforms-ACC!

These people have tainted their uniforms [...]!

...tisztára mossa saját bemocskolódott ruháját!
...clean wash own tainted clothes-ACC!

...to whitewash his own tainted clothes!

A külföldi média boldogan tájékoztatni fogja a fél világot!
the foreign media happily inform will the half world-ACC!

The foreign media will be happy to inform half the world!

Tetteik minősítik önöket!
your_actions label you-ACC!

Your actions define who you are!

The analysis revealed that the known writer could not have written the questionable letters. Among other important things, the order of the words helped to point to the truth.

The next aspect is morphology, but as an introduction to morphology I would like to present some morphosyntactical issues here concerning agreement.

Agreement

Hungarian requires that the verb and its subject agree in number and person. Lack of agreement or incorrect agreement may be evidence for distinctiveness, as in the following examples from anonymous threatening letters:

Követeléseink valamennyi büntetőügyre vonatkozik. [vonatkoznak]
our_demand-PL all to_criminal apply-SG [apply-PL]

Our demands cover all criminal cases.

Követeléseink (demands) is in plural so the verb *vonatkozik/ik* (apply) also needs to be plural *vonatkozik/nak*.

...függesszék fel valamennyi büntetőeljáráson hozott ítéletüket,
suspend all criminal_proceedings_at delivered sentence,
annak [azok] végrehajtását...
its [their] implementation...

Suspend all your sentences delivered at criminal proceedings, their implementation...

Because *valamennyi ítéletüket* (*all your sentences*) is plural the pronoun has to be in plural too.

Mindenki	célponttá	válik,	aki	követeléseinket	és	
everyone-SG	target	will_become,	who-SG	our_demands	and	
utasításainkat	nem	vagy	csak	részben	teljesítik	[teljesíti]
our_instructions	not	or	just	partly	fulfill-PL.	[fulfill-SG]

Everyone will become a target who does not or just partly fulfills our demands and instructions.

Here the verb *teljesíti/k* (*fulfil*) is in plural but it is incorrect, because it agrees with the pronoun *mindenki* (*everyone*), which is technically singular.

Könyörtelen	[Könyörtelenek]	és	hidegvérű ek	vagyunk.
merciless	[merciless-PL]	and	cold-blooded-PL	are_we.

We are merciless and cold-blooded.

The predicate agrees in number with the subject and if it is copulative, both parts agree in number with the subject. The subject here is *mi* (*we*) so the adjectives need to be in plural.

Morphology – Vowel harmony

Another point of interest in Hungarian is vowel harmony. Hungarian has back, front and neutral vowels. Neutral vowels are phonetically front vowels but may combine with back vowels, too. Vowel harmony regulates that words with front harmonic vowels get front vowel suffixes, like *ütő|ről* (*from the club*); words with back harmonic vowels get back vowel suffixes, like *kar|ba* (*into the arm*); and in words where the last syllable has a neutral vowel and the previous has back there is variability, like in *fotel|be|ba* (*in the armchair*). Vowel harmony is worth mentioning because this variability may be an idiolect feature. For example to choose the *-ban* (*in*) suffix to the root *fotel* (*armchair*) marks a certain age group: the elder people.

There are some examples in a case where someone made a fake Facebook profile and harassed a girl. The girl had an assumption about the harasser, that he was someone whom she had known, and she had had e-mails from

this person before. Looking at the e-mails, and doing a complex analysis, among other features, the difference in the vowel harmony made me think that he was not the harasser.

Examples from the Facebook chat:

alkesz|nek (to an alcoholic)

bohém|nek (to Bohemian)

fater|rel (with father)

Examples from the known e-mails:

farmer|ban (in jeans)

póker|ban (in poker)

Ágnes|sal (with Agnes)

There are two other features for analysis: vocabulary and spelling. However in those, there is no difference between the international and the Hungarian methods. So I will focus on the following aspect which is absolutely unique in Hungarian investigations. It is very important in Nagy's complex analytical method but does not appear to be in the English literature. It is reference.

Reference

Reference is a relationship between the language and the world. What Nagy (1980) means by reference in forensic linguistics is how the authors refer to the world by using their individual version of the language. We can illustrate this with the four anonymous threatening letters (mentioned above). In this case literal translation is not as important as the meaning, but I have highlighted the cardinal phrases.

The first aspect is that the word 'media' appears in the texts, both in the anonymous letters and in the known letters. The anonymous letters were from 2008 and 2009, and the known letters were from 2009.

2008

Ezen levelünket több **médiának** is megküldjük. (We also send our letter to several **media**.)

Válaszukat valamennyi **médiában** közölhetik velünk. (Your answer can be announced in any **media**.)

2009

A **médiának** és a rendészeti szerveknek köszönhetően... (Thanks to the **media** and the authorities...)

Mindenki aki azt a levelet megkapja köteles értesíteni [...] a **médiákat**... (Everyone who gets this letter has to inform [...] the **media**...)

Known letters

...mivel a magyar **média** fél önöktől... (...as the hungarian **media** fear you...)

A külföldi **média** boldogan tájékoztatni fogja a fél világot! (The foreign **media** will be happy to inform half the world!)

The second aspect is the target of the threat itself which differs in the questionable and the known letters. The first one was threatening the judges, prosecutors, officers, lawyers, defendants, witnesses etc., but the known author was threatening the police.

The next aspect is the consequences: what will happen if they do not obey?

2008

Akik nem engedelmeskednek azok, **célponttá** válnak... (Everyone becomes a **target** who doesn't obey...)

2009

...aki nem tarja be az utasításinkat [...], azok számunkra nem lesz több, mint egy egyszerű levadászandó **célpont**. (...those who won't follow our instructions will be nothing but an easy-hunted **target** for us.)

Consequences like these did not appear in the letters from the known author. So it is a difference too, like the threat.

The next issue is their connections: Who are they in contact with?

2008

Sok **radikális csoporttal** vagyunk kapcsolatban... (We are in contact with many **radical groups**...)

2009

Radikális, szélsőséges csoportokkal vagyunk kapcsolatban...
(We are in contact with **radical, exaggerated groups**...)

There is also a difference, because the known author has not mentioned anything about his connections either with radical groups or with anybody. And the last aspect to consider is the verbalization of the demand, which also differs in the questionable and the known letters. While in the questionable letters, they are talking about suspense and postponement, the known author “just” demands that the police do what he wants.

2008

...függesszék fel... (...suspend...)

...halasszák el... (...postpone...)

...követeljük, hogy... (...we demand, that...)

2009

Követelésünk a következő... (Our demand is the following...)

...függesszék fel... (2x) (...suspend...)

Nem hozhatnak ítéletet, nem menthetnek fel... (You may not adjudicate, may not acquit...)

Known letters

...utoljára szólítom fel önöket... (...this is my last call upon you...)

...felszólítom... (...I call upon...)

As you can see, the anonymous letters could all be ascribed to the same author, because the complex analytical method established their common authorship by the incorrect agreement, the same syntactical structures, the wording and the spelling mistakes. Finally reference also contributed to the conclusion.

Summary

This paper presented an overview of author identification in Hungary. Hungarian analytical methods were compared to international ones. The question raised at the beginning of this paper was whether the methods used

in Hungary were universal or unique. The answer is that most of the categories which one may use to analyse during an investigation are universal, but within these categories the methods can be unique to a language because of its structure. In Hungarian, those may be methods relating to the word order, agreement, vowel harmony, case endings and so on. The two aspects which are unique in forensic linguistics in Hungary are composition style and reference.

References

- Carole E. Chaski, (2001).** Empirical evaluations of language-based author identification techniques. In *Forensic Linguistics* 8:1, pp1-65.
- Malcolm Coulthard (2004).** Author identification, idiolect and linguistic uniqueness. In *Applied Linguistics* 25:4, pp431-447.
- Robert Eagleson (1994).** Forensic analysis of personal written texts: a case study. In J. Gibbons (ed.), *Language and the Law*. Longman: London, UK.
- Tim Grant (2010).** Text messaging forensics. Txt 4n6: Idiolect free authorship analysis? In Malcolm Coulthard. & Alison Johnson (eds.), *The Routledge Handbook of Forensic Linguistics*. Routledge: London, UK, pp508-522.
- Zorca Kašić & Jelena, P. Đorđević (2010).** Forensic linguistics in criminal investigation. In G. B. Milošević (ed.), *NBP Journal of criminalistics and law*. Beograd: Academy of Criminalistic and Police Studies: Belgrade, Serbia, pp111-126.
- Miklós Kontra (2006).** Nyelv és jog. In F. Kiefer (ed.), *Magyar nyelv*. Akadémiai Kiadó: Budapest, Serbia, pp1018-1038.
- Ferenc Nagy (1977).** Nyelvészeti összehasonlító vizsgálat a személyazonosításban. In Csáki et al. (eds.), *A nyelvész szerepe a kriminalisztikában*. Petőfi Kiadó: Kecskemét, Serbia, pp37-46.
- Ferenc Nagy (1980).** *Kriminalisztikai szövegnyelvészet*. Akadémiai Kiadó: Budapest, Serbia.
- Jan Svartvik (1968).** *The Evans Statements – A case for forensic linguistics*. Statens Humanistiska Forskningsrad: Göteborg, Sweden.

15

Vocabulary learning grounded in an ESP community: Design and effect of a basic medical ESP course

Hiroko Suzuki^a & Miho Fujieda^b

^a*Tokai University*

^b*Kyoto College of Medical Science*

Introduction

The perception of English as a foreign language (EFL) is currently yielding to a new understanding of English as an international language (EIL) in which students need opportunities both to interact with academic content and to draw on their linguistic knowledge and skills. Doing so will, in turn, allow them to fulfill the social needs they require to participate in today's globalized society. To this end, MEXT¹⁶ has determined that Japanese junior high school students, who at the age of 15 begin developing the four skills in English, should acquire about 1200 basic words through the use of MEXT-approved textbooks. This linguistic knowledge is expected to provide the foundation for senior high school communication activities around a variety of topics such as social or scientific issues. Students are expected to acquire an additional 1800 words from senior high school textbooks, empowering students to begin pre-ESP education at the tertiary level. However, a major challenge for students in tertiary pre-ESP courses, when faced with reading texts featuring topics relevant to their majors, is their lack of vocabulary. A survey of about 2500 non-English-major college students (Ono et al., 2005) found that 33% had only a junior high school level command of English vocabulary, while another 34% had a first year of senior high school level of command.

The college students in this study specialize in radiological technology (RT). A survey (Fujieda, 2013) conducted with graduates of the program indicates that their recognition of English as a necessary communication tool was strengthened, allowing them to cope with changing times and emerging

¹⁶ The Ministry of Education, Sports, Science and Technology in Japan

professional trends. Survey respondents also felt that English medical vocabulary teaching at the tertiary level should be emphasized.

According to the needs analysis done as part of the survey, a basic medical ESP course was designed and implemented for students majoring in RT. Students were expected to learn basic terms and concepts effectively and draw on them to acquire productive skills. As a key element in the foundation of the departmental ESP curriculum and materials development, preliminary word lists from three specialized corpora were created. Although preliminary (Fujieda, Suzuki & Koyama, 2013), the introductory level list, while expected to introduce keywords necessary for reading authentic texts about health literacy and medicine, did not attempt to cover the keywords in the texts used in the Basic Medical English course (BME) targeted in this study. The underlying assumption was that, although students need to be exposed to more words in the introductory-level word lists in order to read authentic texts for health literacy, they can only acquire a certain amount of vocabulary, since the one-semester BME course only meets 15 times.

Thus, this study investigates whether the texts used in the current BME course promote student vocabulary and reading skills sufficient for reading the authentic health literacy texts selected for the final BME project.

Background

The BME Course in this study is offered at a small four-year private vocational college exclusively for future radiological technologists, or RTs (elsewhere called x-ray technicians or radiographers). One month before graduation, the students take a national license exam, consisting of 14 subjects. These are paper-based multiple-choice tests including both theoretical (e.g., basic medicine, radiation physics, and medical imaging engineering) and practical (e.g., radiation safety management, medical imaging informatics, and radiographic technology) content. Successful completion of the ministry-specified educational program, which is equivalent to graduation from college, is a prerequisite for taking the exam and obtaining a license.

The result of the English placement test upon their admission to the college shows a wide range of English proficiency levels, from false beginner to high intermediate. Initial class surveys have revealed that student attitudes toward English also vary; some of them do not hesitate to say openly, “I hate English.” Since English is not part of the national license exam, the position of English in the curriculum tends to be somewhat peripheral, although the

college's mission statement emphasizes the social need for "internationally-minded" healthcare professionals.

BME materials

Considering the current proficiency level of the target students and consulting with content teachers and drawing on the concept of a constellation of genres (Swales, 2004), the course engaged students in a series of authentic texts. The instructor found reading materials on the Internet and decided to use authentic texts written for the purpose of health education directed at the general public in the U.K. and the U.S., with many articles coming from the websites of the BBC's Science and Nature Human Body and Mind, and Medline Plus Interactive Health Tutorials by the U.S. National Library of Medicine and the National Institutes of Health.

Figure 1 gives an overview of the kinds of English learners of medical ESP at the target college specialized in RT are exposed to and where. The order in which the texts were sequenced took account of both students' English proficiency and their content knowledge development in L1 as advanced by health education articles for the general public.

After learning senior high school English using MEXT-approved textbooks aiming to help students acquire 3,000 vocabulary words by high school graduation, the students targeted in this study learned fundamental scientific and medical terms relevant to human health through Science English Course in their first semester and BME in the second semester. These are represented as Phase (1) in Figure 1, in which students begin the process of becoming RTs in a medical discourse community. Phase (2), which takes place in the students' second year, introduces English vocabulary items that were chosen by considering fundamental L1 subjects relevant to RT. Thus, BME is a sort of pre-ESP taking place in the latter half of Phase (1), where the students' focus is to learn fundamental English terms to describe human body structures and relevant diseases, with the goal of reading texts in Phase (1) and at the beginning level of Phase (2).

Third-year students at the target college experience working as student RTs in a hospital, which stimulates in them a stronger consciousness to become professional RTs in the future. In Phase (3), students read introductory English RT textbooks in the RT subjects. Phase (4) requires gradually stronger English proficiency, providing a transition to life as graduate students or professional RTs and enabling them to read academic reports, research papers and descriptions or instructions of the devices relevant to RT

and to communicate within their professional community. As such, the English level of the materials at Phase (4) corresponds to texts from periodicals such as Radiology or the American Journal of Roentgenology.

As a key element of the departmental ESP curriculum and materials development, preliminary word lists from three specialized corpora were created (Fujieda, Suzuki & Koyama, 2013). The purpose was to illustrate what words the students needed exposure to before reading RT related textbooks or articles in RT technical journals. With their overall college career in mind, three specialized corpora were compiled from sources targeting audiences with different levels of specialized knowledge: patient education materials at Phase (1), introductory university textbooks at Phase (2), and research articles at Phase (3) and (4) in Figure 1.

The outcome of compiling these three corpora from sources targeting audiences with different levels of specialized knowledge, as illustrated in Figure 1, has enabled the authors to simulate a developmental nature of the learner vocabulary, since the words in the list are supposed to help the students to learn through reading activities with texts at these levels. In this sense, the lists serve as a guideline for designing a coherent ESP program in which course planners can create teaching and learning materials that support learners' professional development (Nation & Webb, 2011). One of the three word lists created from the analytical results of the three corpora in this study are used in analyzing the difficulty level of the reading texts used in BME: the word list for patient education materials (Level 1).

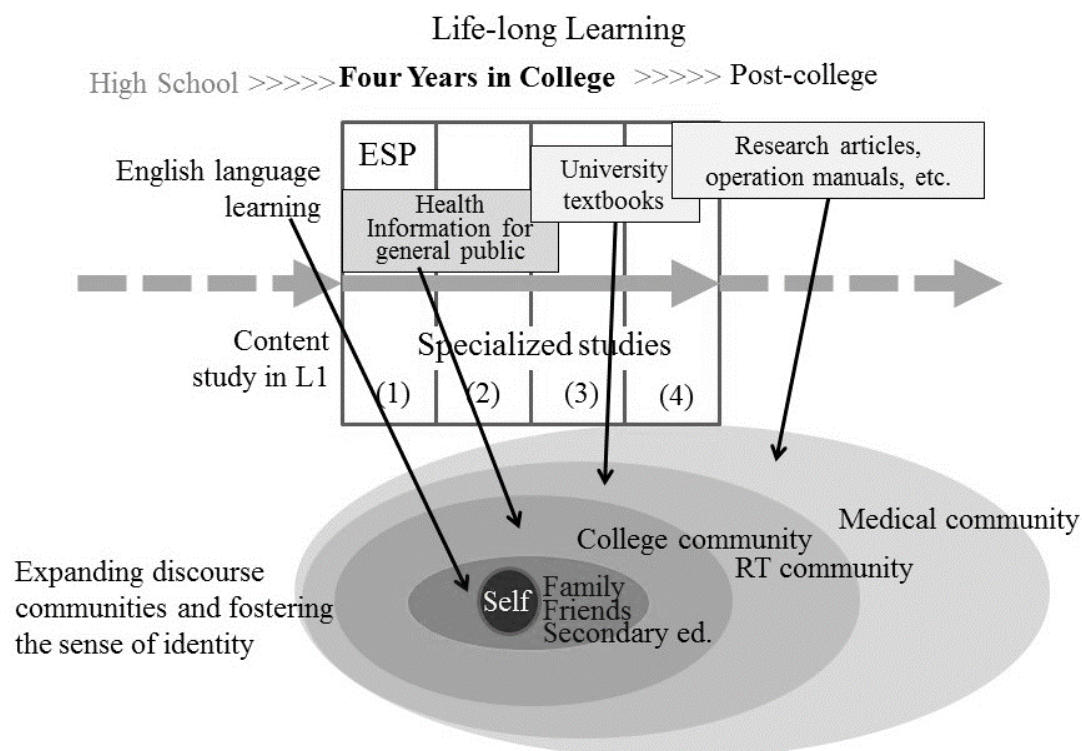


Figure 1: Illustrating a developmental process of learner vocabulary toward becoming a radiological technologist in hospitals and medical institutions

Course Design

In order to figure out the needs of English for RTs, which are similar to those of other medical professionals, a systematic needs analysis was done (Fujieda & Suzuki, 2010; Fujieda, 2013). The results revealed that basic needs such as recognizing and using medical terminology in daily medical interactions, reading doctors' instructions in English, and reading screen prompts are skills that all RTs have to master. The results were referred to in selecting the BME materials at pre-ESP level and in verifying the practical effectiveness of their English and contents.

Table 1 illustrates the contents and the materials used in the BME courses. The four items, weeks of the semester, phases of the course, contents and text names of the materials used throughout the semester are shown chronologically. First, in the preparation phase, nouns of basic body parts are introduced, followed by interactive online exercises and illustration labeling exercises. These materials help the students organize terms with the aid of the relevant information from the reading material "04 Body structure."

Week 1: Preparation Basic body parts (head, upper body & lower body)
01 body parts
02 internal organs
03 interactive body
04 body structure ®, 05 vocabulary
Week 2-10: Foundation Thematic reading
Reading 1 “Lungs” (respiratory system)
06 lungs ®, 07 vocabulary, 08 comprehension quiz
09 respiratory system-illustration diagram
10 lungs-illustration diagram, pair-work
11 respiratory system ®, 12 respiratory system-puzzle
Reading 2 “Heart” (circulatory system)
13 heart ®, 14 vocabulary, 15 comprehension quiz
Mid-term Examination 16 writing about a body part
17 heart-review, pair work
18 circulatory system ® (optional), 19 puzzle
Reading 3 “Bones” (skeletal system)
20 bones (1) ®, 21 vocabulary, 22 comprehension quiz
23 bones (2) ®
24 human skeleton-worksheet
25 heart attack-PowerPoint slides used for the model presentation
Final Exam 26 writing about a disease chosen for your project
Week 11-15: Application Individual projects
Reading 4 Self-selected “Common Health Problems”
Ex. 27 pneumonia ®, 28 vocabulary
29 lung cancer ®, 30 vocabulary

Table 1: Overview of Contents and Materials in the BME Course

® = *reading texts*

The foundation phase provides three reading materials about physical description and function of three main human organs. These are more closely related to the students’ future profession (i.e. lungs, the heart, and bones). Before reading the English passages, students review what they have learned in their L1 subject courses by posting any anatomical information they know about the organs in Japanese on the class website as homework, followed by class discussion.

After these core reading activities, at the application phase, a mini-project called “Common Health Problems” starts. Each student chooses a health problem related to the three main organs they’ve learned about, reads about the problem, and prepares a PowerPoint presentation. Students are encouraged to select the topic that they are most interested in. Their task is to explain the problem as a healthcare professional to somebody in their care. This activity extends the previous readings, allowing students more

autonomy. Students who have chosen the same topic are invited to help each other.

Vocabulary in BME materials

The previous study determined the Level 1 word coverage of the BME reading texts (Suzuki & Fujieda, 2012). It revealed that the eight reading texts (Table 1, ⑧) featured approximately 10% of the Level 1 vocabulary, 14% of the introductory activity vocabulary (Table 1, 04), and 15% of the final project vocabulary (Table 1, ex. 29). Table 3 shows how many Level 1 vocabulary items from previous texts recur in new texts. The figures in the one-timers are cumulative, demonstrating how many of the words in the texts student will not be exposed to again. Table 3 demonstrates what students were faced with in each reading text, revealing that recurring words were fewer than new words, which were, in turn, far fewer than one-timers. The results indicate that the texts made it relatively difficult for students to incidentally learn the target vocabulary through experiencing authentic reading.

	04. Body structure	06. Lungs	13. Heart	20. Bones	29. Lung cancer
One-timer	13	41	60	92	116
Recurring	4	7	8	7	14
New	21	9	15	14	38

Table 3: Level 1 words in the reading texts

However, since the BME materials were chosen based on the course design to situate the target language and to systematically develop activities considering the students' L1 knowledge and communication skills in English, the materials were designed to reintroduce contexts and feature highly relevant contents (Figure 1). The one-timers are meaningful when we consider how embedded literacy (Martin 2013) impacts a target reading text. The concept of embedded literacy is a complex structure of knowledge related to the topic as well as word groups to form meaning of knowledge, which supports current reading comprehension. Language is mediated so a reader can recognize communicative events as instances of particular genres, leading a discourse community (Swales, 1990) to assign particular terms or labels to these particular perspectives on the notion of genre, based on the concepts of prototype, intertextuality and inference (Paltridge, 1995).

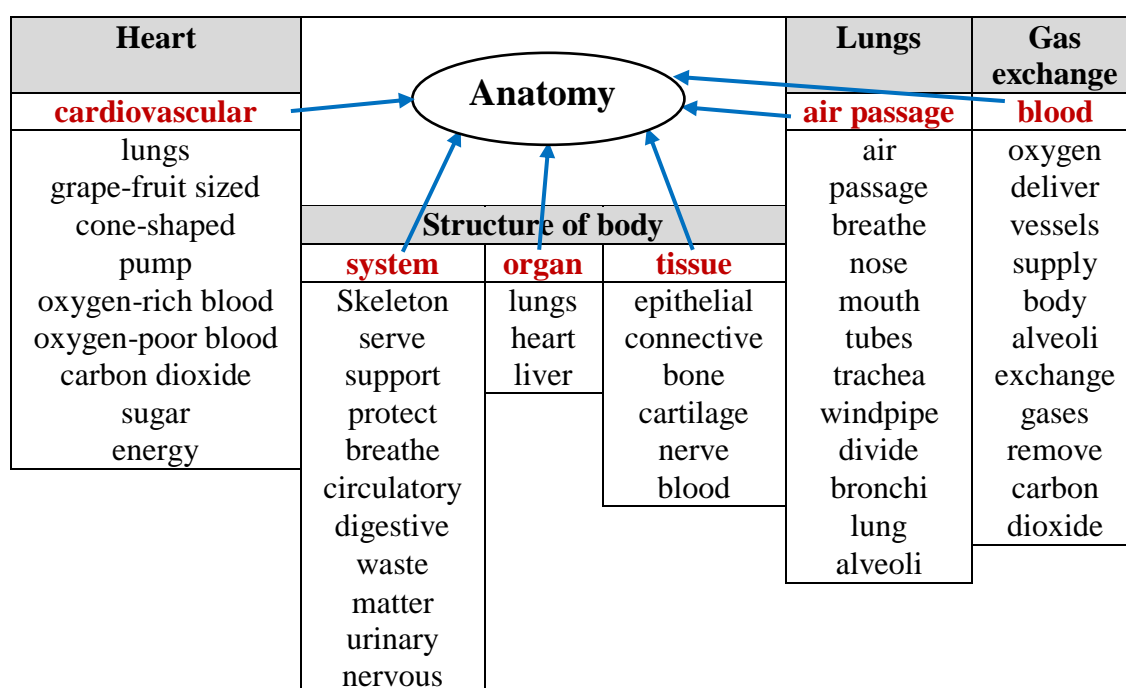


Figure 2: Keyword “Anatomy” and relevant concepts and word groups

Thus, text 29, “Lung cancer,” for example, identifies seven key concepts: anatomy, abnormal growth of cells, causes, two types of cancer, symptoms, diagnosis and treatment. Each of them forms a word group; for example, the word group for “anatomy” included lungs, blood, oxygen, breathe, release, carbon dioxide, trachea, bronchial tubes, alveoli, small blood vessels, secrete, mucus, trap dirt, and cilia. When the students read the text “Lung cancer” in the final project phase, their text comprehension was supported by knowledge of concepts relevant to the contents of “Lung cancer,” which they had learned through materials in the previous phases (Table 1) with the word groups to explain the concepts. The word “Anatomy” in Figure 2, for example, is a concept connected to “Lung cancer.” Figure 2 shows word groups listed below the relevant key concept; highlighted words in the word groups are those featured in the “Lung cancer” text.

As demonstrated by the results of the vocabulary analyses of the BME materials, the course design combines theory and practice, drawing on the concept of situated learning (Lave & Wenger, 1991). In this model, the enrichment and deepening of student’s understanding of the target vocabulary acquisition is an incidental outcome of experiential learning, embodiment, and generative processing.

Analyses: Students' output in their target vocabulary

Purposes

The BME course design was aimed at students in an optimally situated learning community. So in order to evaluate how the coursework actually affected student learning of the target medical terms, the amount of target vocabulary in the texts written in the final writing test (Table 1, 26) as their output was analyzed. The analyses of BME's online pre-reading vocabulary quizzes and post-reading assignments and reflection on contents also attempts to clarify the relationships between students' perceptions of vocabulary learning, their learning strategies, and their depth of comprehension of target medical terms.

Results

Table 4 overviews data on BME student English and writing protocols from the final test. The placement test results indicate that the average English proficiency, especially regarding vocabulary and grammar, was around the first-year senior high level when they entered college. Assuming that students acquired the vocabulary required by MEXT's Course of Study implemented in 2013, the students' vocabulary size may be approximately 1,500 words.

	Mean	SD	N
Placement (100)	49.6	13.0	89
V-quiz (50)	36.4	10.5	91
Writing (10)	7.9	1.3	90
Total token	201.4	61.4	90
Type/token	0.6	0.1	90
Level 1* type	15.1	5.8	88

Table 4: BME student data

** Level 1 is a corpus-based specialized wordlist for radiological technology students. It consists of the 306 most frequently used word families in a patient education materials corpus.*

“V-quiz” in Table 4 means 5 quizzes on the vocabulary they were learning throughout the course; the five scores are summed up here. “Writing” shows the average score of the writing test targeted in this study, with 10 points as the maximum possible. The average of 7.9 points is relatively high, and may be attributable to the fact that they had created PP slides to organize their ideas with the keywords beforehand and used them as reference in their writing. “Total token” is the length of the students' written texts. They wrote

about 200 words in average, which is relatively high, considering the starting point of their English proficiency level and sparse experience of writing English before entering this college. “Type/token,” the ratio of number of vocabulary items to total number of words in a text, is a way of indicating the semantic density of a text, as well as the learner vocabulary learned through the course. “Level 1 type” shows that their texts included 15% of Level 1 words, on average.

	Place- ment	V-quiz	Writing	Total token	Type/ token	Level 1
Placement	1	.448**	.358**	.158	.004	.034
V-quiz		1	.577**	.284**	-.109	.141
Writing			1	.536**	-.390**	.342**
Total token				1	-.609**	.633**
Type/token					1	-.403**
Level 1						1

Table 5: The correlational relationships between six kinds of student data

**** Correlation is significant at the .01 level (2-tailed).**

Table 5 demonstrates the correlational relationships between the six items in Table 4. Writing vs. V-quiz and Writing vs. Total token have significant correlations at the .01 level. This indicates that the instructor, who scored the writing texts as an impressionistic evaluation, seems to have been affected by the students’ accuracy of vocabulary and text length. On the other hand, Type/token ratio and Text length showed negative correlations. These figures indicate that the BME students, regardless of their text lengths, tended to use various Level 1 terms.

Text length	Total token	Type	L1-token	L1-type	No. of students
Group L	255	132	40	19	39
Group S	160	94	23	12	49
	type/ token (%)	L1type/ token (%)			
Group L	52	14			
Group S	59	14			

Table 6: text length and number of L1-words in the texts between Groups L and S

In order to explore the meaning of negative correlations between “Total token” and “Type/token” ratio, two groups are created to compare the averages of these items. As shown in Table 4, the average total token was 201 words, so Group L (L stands for long) includes the texts of more than or equal to 201 word and Group S (S stands for short) are those of fewer than

201 word texts. Four categories: the average words of total token, type, Level 1- (L1 in Table 6) Token and L1-type between Groups L and S were compared. The total tokens between the two groups were significantly different, but the types, L1-tokens and L1-types gradually reduced their differences. Moreover, the differences between type/token ratios of all and Level 1 words are not meaningful. The line plots in Figure 3 illustrate their significance.

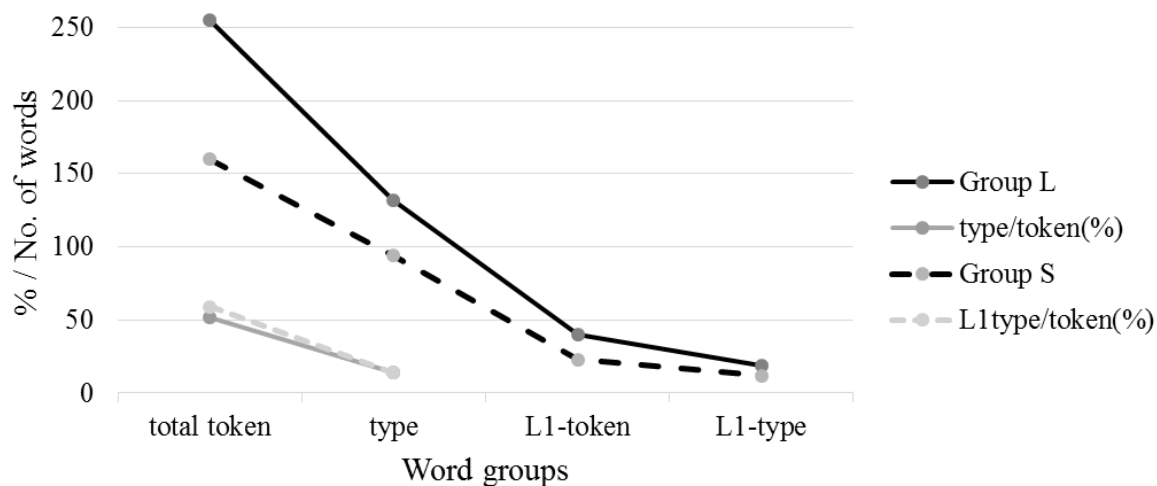


Figure 3: Comparing Group L and S texts according to text length and type/token ratio

The data indicate that the BME students understood or learned the target vocabulary deeply enough to produce it in their writing, although the text length and the amount of information that Group L explained in their writing were almost twice those of Group S.

Group	V-quiz	Writing	total token	Level 1	No. of students
Heart-L	40.3	8.5	235	13	8
Lung-L	42.5	8.7	248	21	14
Heart-S	30.0	6.8	186	11	9
Lung-S	23.4	6.7	189	18	9
	spelling	mechanics	structure		
Heart-L	3.1	8.6	4.4		
Lung-L	2.4	8.6	1.7		
Heart-S	2.9	11.1	6.5		
Lung-S	2.9	14.4	3.2		

Table 7: Differences in accuracy of the texts between Groups L and S

In order to identify the differences between Groups L and S more clearly, the next analysis sought differences in language accuracy between the two

groups' texts. Since the vocabulary, especially Level 1 terms, used in the reading materials is different for each topic, this analysis focused on the two articles: "Lung cancer" and "Heart attack" (Table 1, Reading 4), which were the texts chosen by the most students. Each text was analyzed for spelling mistakes, mechanical revision points, such as, singular or plural forms, articles, verb forms, missing be-verbs or general verbs, and structural revision points which require reforming a whole sentence. Heart-L is the group who read the "Heart attack" text and got the highest Group L scores on their writing, and Lung-L is the group getting the highest scores in Group L after reading the "Lunch cancer" text. Heart-S and Lung-S are the groups with lower Group S scores on their writing. The borderline between the high and the low scores in writing was determined at 8 points because of the average point, 7.9. The upper half of Table 7 shows the student data of the vocabulary quizzes and about their writing and the lower half compares three types of grammatical mistakes. Interestingly, all groups had few spelling mistakes, probably as an outcome of the five vocabulary quizzes throughout the course. However, both lower "Heart-S" and "Lung-S" groups in Table 7 had many more mechanical mistakes and structural problems than the other groups.

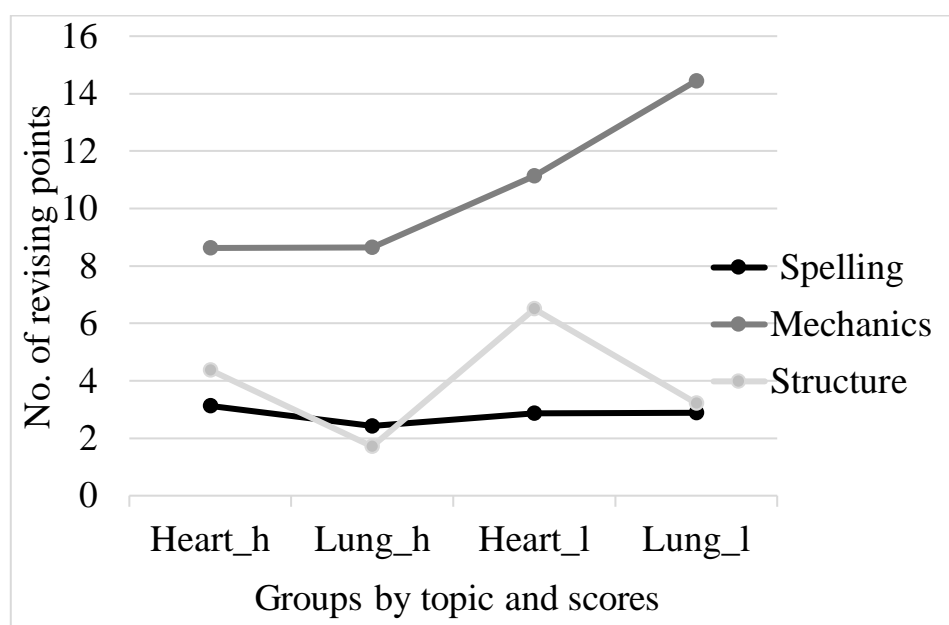


Figure 4: Grammatical accuracy in Group L & S' texts

Figure 4 demonstrates that for all groups, regardless of their test scores and text length, the target ESP vocabulary, including Level 1 terms used in their writing, were satisfactory in terms of accuracy and number of words. However, the texts in Group S had so many grammatical mistakes that readers could not grasp the meaning clearly.

Student text analysis

The following two examples demonstrate differences in the texts' readability.

Student A in Group L

Topic: Lung cancer
Writing score: 9/10 points
Token total: 390 words
Type/token: 0.47
Level 1 used: 32 words

Student A's comment on the project: I learned that many people die from lung cancer. Especially I was satisfied with an opportunity to deepen my understanding about specialized knowledge in medicine, for instance, there are several kinds of cancers and what are their causes. So I enjoyed studying in BME.

Student A created eight slides. The following two slides, forming diagrams with key phrases, pictures and symbols, explain two types of cancer and their causes.

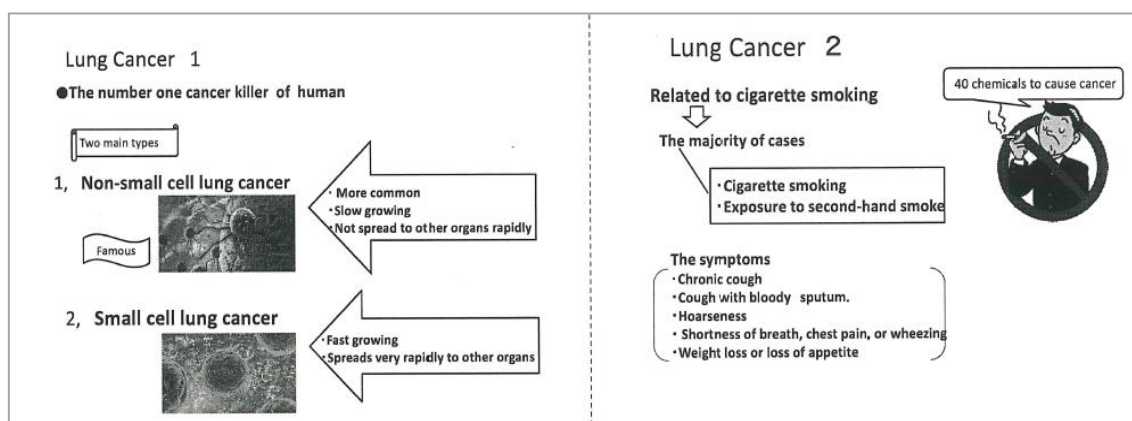


Figure 5: Student A's Slides

Based on them, Student A wrote:

Lung cancer is the number one cancer killer of men and women. There are two main types of lung cancer: non-small cell and small cell. Non-small cell lung cancer is more common, slow growing, and does not spread to other organs rapidly. Small cell lung cancer is not as common as non-small cell. But it is fast growing, and spreads very rapidly to other organs. Cigarette smoking or exposure to second-hand smoke causes the majority of lung cancer cases. 40 chemicals of cigarette cause cancer.

The underlined sentences demonstrate that Student A formed sentences by using the phrases in the slides and appropriately adding verbs, articles and connecting markers.

Student B in Group S

Topic: Heart attack

Writing score: 7/10 points

Token total: 248 words

Type/token: 0.52

Level 1 used: 14 words

Student B's comment on the project: (I chose this topic because) my father has a heart disease and has collapsed a couple of times. (I wanted to emphasize) the causes and how to prevent heart attacks (through the presentation). I was benefited by acquiring new knowledge about heart attack in this project.

Student B's following two slides illustrate similarities between heart attack and angina and their different treatments.

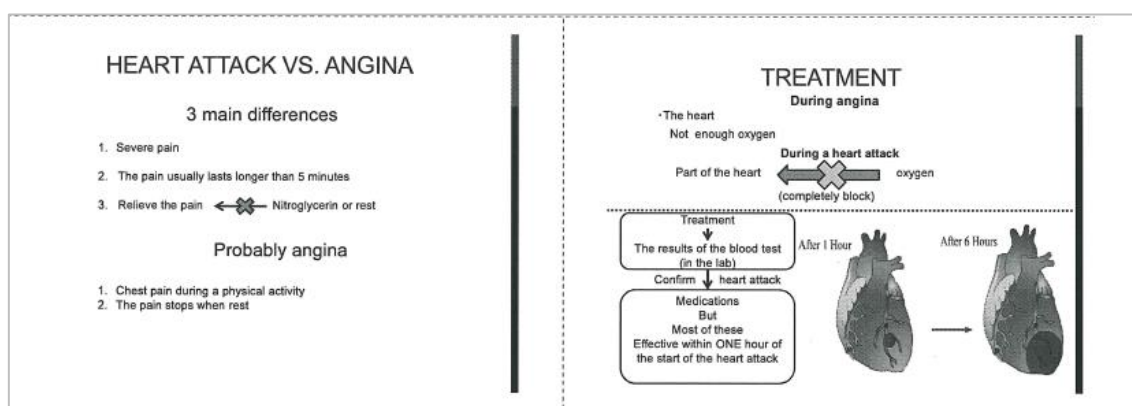


Figure 6: Student B's Slides

Based on these slides, student B wrote:

Heart attack is similar to angina. Its 3 main differences are severe pain. the pain usually lasts longer than 5 minutes, nitroglycerin or rest don't relieve the pain but angina are signs that chest pain during a physical activity, and the pain stops when rest. First treatment, you confirm heart attack by the blood test. When heart attack, you are medicationed by doctor, but most of these, it (is) effective within ONE hour (from) the start of the heart attack.

About a half of the 80 words in student B's writing were unclear due to grammatical mistakes, which are highlighted, and structural problems, which

are double-underlined. Student B's writing has two significant drawbacks. One is capitalization. Most of the students in Group S did not seem to have constraints in terms of when to use capital letters. It may be due to their lack of procedural knowledge about how a sentence unit is formed. The other point is that the subject shifts from second to third and back. Students may have difficulty in finding the appropriate subject, especially if the equivalent Japanese sentence does not require a subject. The subjective shift from second to third person in a text also indicates that the students' writer perspectives have not emerged; if they are not conscious of having a reader, they may eventually lose sight of their objective.

Concluding remarks

The vocabulary analyses of the reading texts demonstrate that the BME coursework could be situated for the students by integrating what they were learning in L1 specialized subjects. Their responses on the course evaluation survey revealed that they were able to 1) gradually see connections between their L1 content area studies and what they have learned in BME, and 2) shift their perceptions of medical vocabulary learning from item memorization to content learning in context. However, some students seemed to lack grammatical knowledge and awareness and had difficulty in forming appropriate sentence structures. On the other hand, they used more of the target vocabulary than expected and their spelling had few mistakes. This may have been because the vocabulary quizzes that they had in every section were effective, but the little explicit grammar instruction did not give them clues how to be aware of grammatical mistakes in their own texts.

These results imply that the students paid more attention to expressing their meanings with the target vocabulary in fully contextualized situations. However, this does not necessarily mean that the students needed more grammar instruction, since they would not have been able to easily apply their learned grammatical knowledge in their spontaneous writing process.

This study has clarified that a way to improve the course design, specifically, the necessity of having incidental grammar learning through writing in the pre-BME course to give students opportunity to learn how to aware of grammatical problems in their writing. So, in each activity in BME, they can deepen their understanding of the target vocabulary by expressing themselves more precisely without too much context-dependent understanding.

References

- Miho Fujieda (2013).** *A survey of English use by practicing radiological technologists*. Paper presented at the sixteenth JASMEE Academic Meeting, Chiba, Japan.
- Miho Fujieda & Hiroko Suzuki (2010).** *Corpus-informed vocabulary learning in ESP: A sociocultural view*. Paper presented at APEC-RELC International Seminar Conference, Singapore. Abstracts, p. 24.
- Miho Fujieda, Hiroko Suzuki & Yukie Koyama (2013).** *Creating corpus-informed word lists for a college radiology ESP program*. Paper presented at Vocab@Victoria, Victoria University of Wellington, New Zealand.
- Jean Lave & Etienne Wenger (1991).** *Situated learning: Legitimate peripheral participation*. Cambridge University Press: Cambridge, UK.
- James Martin (2013).** Embedded literacy: Knowledge as meaning. In *Linguistics and Education* 24, pp23-37.
- Hiroshi Ono, Muraki, Hidearu Muraki, Norio Hayashi, Naoki Sugimori, Hironari Nozaki, Toshiharu Nishimori Machiko Baba, Keiko Tanaka, Takeshi Kuniyoshi and Shien Sakai (2005).** Nihon no daigakusei no kio gakuryoku kouzou to rimediaru kyouiku [A development of placement test and e-learning system for Japanese university students]. In *NIME Kenkyu Hokoku [NIME Research Report]* 6, pp1-6.
- Hiroko Suzuki & Miho Fujieda (2012).** Iryokei ESP project gakushu no sekkei to sono kouka [Designing Medical ESP project-based course design]. In *Educational Development, Tokai University* 5, pp17-58.
- John Swales (2004).** *Research genres: Exploration and applications*. Cambridge University Press: Cambridge, UK.

16

An investigation of brain synchrony between students and their teacher during a secondary school EFL lesson

Reiko Yamamoto^a, Takayuki Nozawa^b,
Hyeonjeong Jeong^b, Shigeyuki Ikeda^b,
Kohei Sakaki^b, Ryuta Kawashima^b,
Craig Smith^c & Yasushige Ishikawa^c

^a *Osaka International University*

^b *Institute of Development, Aging and Cancer, Tohoku University*

^c *Kyoto University of Foreign Studies*

r-yamamo@oiu.jp

takayuki.nozawa.b5@tohoku.ac.jp

hyeonjeong.jeong.b8@tohoku.ac.jp

shigeyuki.ikeda.e2@tohoku.ac.jp

kohei.sakaki.q1@dc.tohoku.ac.jp

ryuta.kawashima.a6@tohoku.ac.jp

c_smith@kufs.ac.jp

y_ishika@kufs.ac.jp

Introduction

Synchrony from an Applied Linguistics perspective

There is evidence that synchrony is an innate and fundamentally necessary part of the human ability to engage in social interaction with other people. Tomasello (1999) claims that just after being born, an infant starts proto-conversation with an interlocutor, a fosterer. This proto-conversation has a distinct turn-taking structure: a response to an approach by another person, and vice versa. For example, a new-born baby mimics adults' body movements, especially the movements of the mouth and the head. A six-month-old infant begins imitative learning with gaze-following behavior in synchronization with adults' attention and other behavior. An infant uses direct gestures such as pointing with a finger to show something to adults. These gestures indicate that the infant tries to create synchrony with an adult directed to a target to which the infant is paying attention.

According to Condon & Sander (1974), the human infant moves in precise and sustained synchronous organizations of change of movement with the articulated structure of adult speech. The movement of a body part – head, eyes, mouth, elbow, wrist, finger, thumb, eyes, and hips – can be examined by varying the number of scanned frames. Micro transcriptions reveal a synchronous organization of behavior which engages the speaker and the listener. Synchrony of organized correspondences between adult speech and infant body movement suggests a bond between human beings which is based on expressions of participation within shared organizational forms.

Kendon (1990) reported that the way in which individuals synchronize with one another can vary. The phenomenon of movement mirroring appears to occur only between a speaker and the person directly addressed. Other participants move concurrently with the movement of the speaker and listener, as an axis of interaction between them is set up. Their movements are either of quite different form from those of the direct addressee, or else they have a different timing. The coordination of the listener's movements with the speaker's behavior is brought about by the listener's response to the stream of speech.

Marsh et al. (2009) investigated how synchronization of behavior and emergent social units of perception and action are reflected in feelings of connection to other people. Bodily synchrony and physical cooperation contribute to sociality; and, social synchrony and cooperative social action contribute to the emergence of a social unit. These studies show that human beings, from infancy, initiate their own learning by actively creating, and by responding to, social interaction.

Synchrony from a Neuroscience perspective

The following four studies suggest that synchrony plays an essential role in social interaction. Hasson et al. (2012) shows that brain-to-brain coupling constrains and shapes the actions of each individual in a social network, leading to complex joint behaviors that could not have emerged in isolation. These joint behaviors reveal the existence of forces that operate in interaction and shape social worlds.

In Stephens et al. (2010), functional magnetic resonance image (fMRI) was adopted to record the brain activity of both speakers and listeners during natural verbal communication. The speaker's spatiotemporal brain activity was used to model the listener's brain activity. It was found that the speaker's

activity is spatially and temporally coupled with the listener's activity; however, this brain coupling vanishes when participants fail to communicate. Moreover, although generally the listener's brain activity mirrors the speaker's with a delay, areas of the brain that display predictive anticipatory responses were found. The extent of neural coupling was reflected in quantitative measures of story comprehension; and, the greater the anticipatory speaker-listener coupling was, the greater the understanding. The observed alignment of production and comprehension-based processes is a mechanism by which brains convey information.

Schippers et al. (2010) explored whether brain synchronization occurs during continuous streams of actions. The brain activity of both gesturer and guesser in a game of Charades was measured. It was found that the guesser's brain activity in regions of the brain which are involved in mental processing echoed the gesturer's brain activity.

Insights for educational planning are available from what has been learned about the ways human beings create and sustain inter-personal communication. In a study of the means of learning, Kent (2013) argues that the most effective teaching occurs when teachers and students are in synchrony with one another because teaching is a natural cognitive activity that requires human interaction. Kent (2013) claims that the dynamic and complex process of synchronization can inspire students to participate actively in their own learning. Thus, teaching conditions may be optimal when synchrony exists in interaction. Kent (2013) believes that a learning environment where synchronization occurs may even supersede the constraints of skill levels and accommodate a growing, changing brain.

The ultimate aim of future steps of the study reported in this paper is to investigate whether English language learning is facilitated when brain synchrony occurs between students and their teacher.

Definition

In this study, synchrony is defined as measurable inter-personal brain synchronization between students and their teacher. The inter-personal brain synchronization was measured by ultra-small near-infrared spectroscopy (NIRS) devices worn by the participants (see Figure 1). The Institute of Development, Aging and Cancer, Tohoku University, the Hitachi Co., Ltd., and the Japan Science and Technology Agency created an ultra-small NIRS device by using Large-Scale Integration (LSI), a process of integrating thousands of transistors on a single silicon semiconductor microchip

(Tohoku University, Hitachi Co., Ltd., & Japan Science and Technology Agency, 2011).



Figure 1: NIRS device

Inter-personal brain synchronization was confirmed by the measuring of the number of body movements counted by the method presented in Condon & Sander (1974).

Research questions

This study sought to answer the following research questions:

- 1) Would synchrony between the teacher and the students occur during the warm-up activities; and also, during the language learning tasks of an EFL lesson?
- 2) Would synchrony that occurred during the warm-up activities evoke synchrony in the learning tasks?
- 3) Would the level of synchrony during the warm-up activities be similar to the level of synchrony which occurred during the language learning tasks?
- 4) Would a higher level of synchrony occur when students were actively engaged, as measured by body movements, in the lesson?

In order to answer the above four research questions, an experiment was designed (see Experiment design section), and synchrony was measured by ultra-small NIRS devices worn by the participants. The synchrony was also confirmed by measuring the number of body movements counted by the method described by Condon & Sander (1974).

Experiment design

Two groups (i.e., Class A and Class B) of nine randomly-selected secondary school Japanese female students and their Japanese female English-as-a-foreign language (EFL) teacher took part in the study. The same teacher taught both Class A and Class B. It was assumed that the delivery and reception of the same lesson materials may vary even with classes that are expected to be similar because of human relationship factors and variations in individual emotional and physical states.

Two lesson components were created: 15 minutes of warm-up activities which were intended to establish a positive and enjoyable team-work atmosphere, and 35 minutes of language learning tasks which followed the course plan of study and practice. To avoid an order effect of the warm-up activities on the learning task, the warm-up activities were positioned before and after the language learning tasks.

A meeting with the teacher beforehand was held to discuss how the warm-up activities and the language learning tasks would be conducted. The warm-up activities and the language learning tasks were performed in English by the students. The teacher used English and Japanese for the lesson instructions and other social communication.

It was decided that the warm-up activities would consist of the students singing songs together with the use of pictures as teaching aids. It was expected that this type of warm-up activity would evoke synchrony. The songs were ‘Five Little Monkeys’ the first time and ‘Ten Fat Sausages’ the second time the teacher taught Class A and Class B. The language learning tasks had three components: answering the teacher’s questions, translating the teacher’s Japanese into English, and self-assessment of a writing assignment using a check list. Among those three components, the third component, student self-assessment of the completion of a writing assignment using a check list, was excluded from the analysis of the synchrony, since it was anticipated that student-teacher interaction would not occur during that component. Table 1 shows an overview of the design of experiment.

	Class A	Class B
First time	15-minute warm-up activities	35-minute language learning tasks
	35-minute language learning tasks	15-minute warm-up activities
Second time	35-minute language learning tasks	15-minute warm-up activities
	15-minute warm-up activities	35-minute language learning tasks

Table 1: Overview of the design of the experiment

Validation of the study

Participants

Two groups (i.e., Class A and Class B), each of nine different randomly-selected secondary school Japanese female students and the same female Japanese EFL teacher, took part in the study.

Methods

Inter-personal brain synchronization was measured by ultra-small NIRS devices worn by the participants. Wavelet transform coherence (WTC) analysis (Grinsted et al., 2004) was applied to measure the brain synchronization at different time scales. WTC values for each of all possible pairs were time-averaged at each period (wavelet time scale). Differences in their averages between (i) the warm-up activities versus the language learning tasks, (ii) the language learning tasks after versus before the warming-up activities, and (iii) the warm-up activities of the more versus less engaged classes, were statistically tested using paired t-tests. A false-discovery rate (FDR) adjustment was applied for the multiple testing for periods.

Inter-personal brain synchronization was confirmed by measuring of the number of body movements as counted by the method described by Condon & Sander (1974). All four lessons were recorded with three digital video cameras from three different angles: from the right front, the left front, and the back of the classroom. The movement of the following seven body parts – head, eyes, mouth, elbow, wrist, finger, and eyes – of all of the students who wore the NIRS devices in the two groups was observed and the number of body movements was counted by an expert observer.

Results

The WTC analysis, as is shown in Figure 2, revealed that in both Class A and Class B, inter-personal brain synchronization during the warm-up activities was significantly greater than that during the language learning tasks (FDR-adjusted $q < .05$).

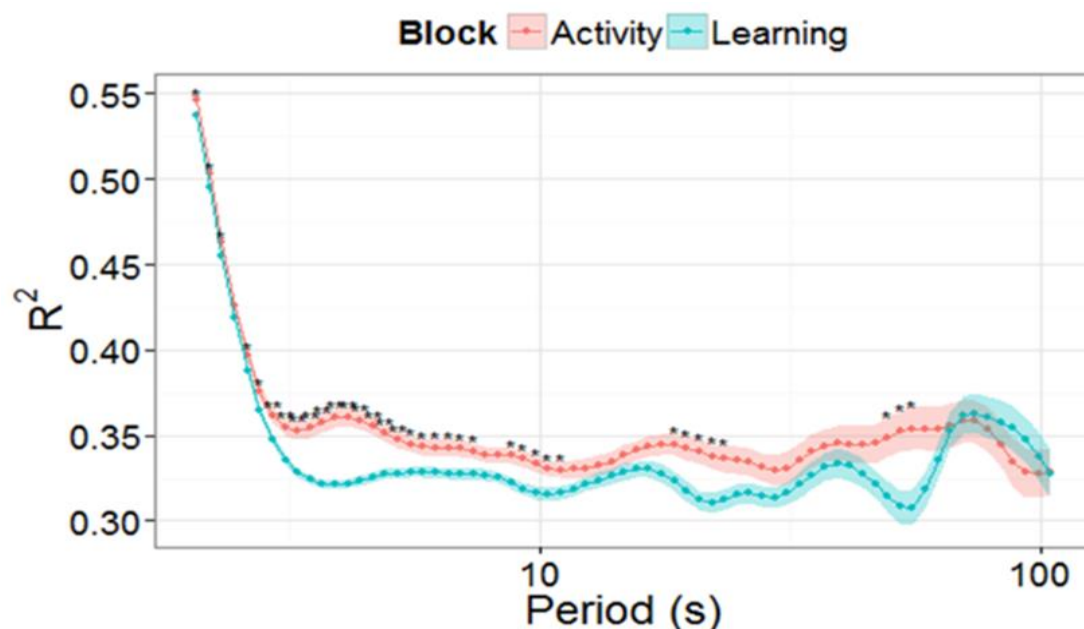


Figure. 2: Difference in inter-personal brain synchronization between the warm-up activities and the language learning tasks
 Note: * = FDR-adjusted $q < .05$

As is shown in Figure 3, no significant difference in inter-personal brain synchronization was found between the brain synchronization which occurred during the learning tasks whether the warm-up activities were conducted before or after the learning tasks.

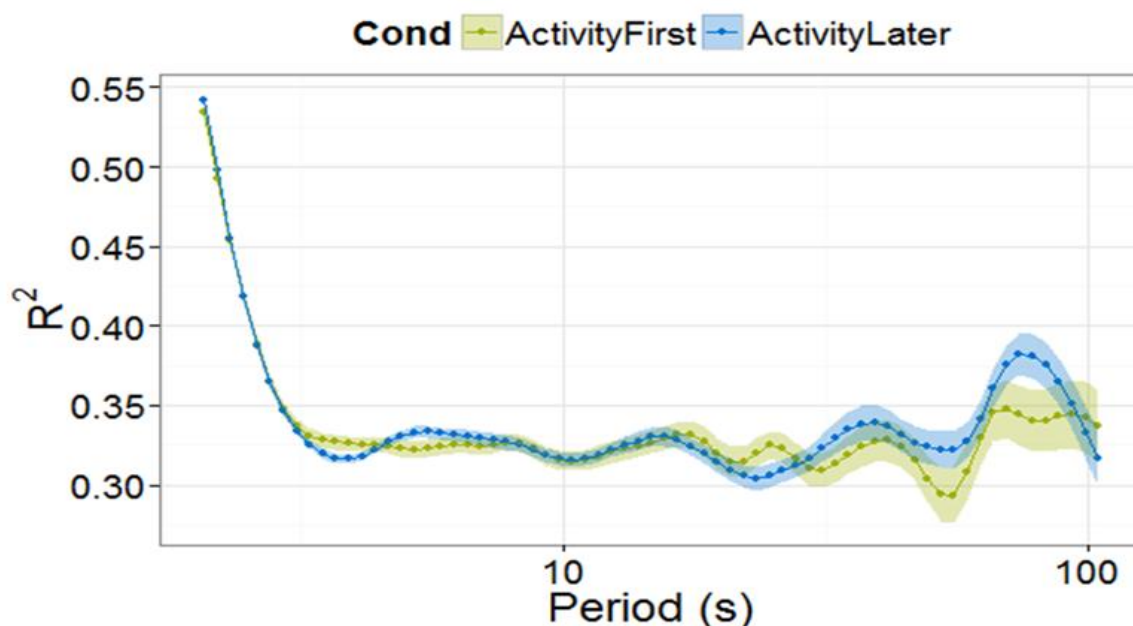


Figure 3: Difference in inter-personal brain synchronization during the language learning tasks when the warm-up activities were before (*ActivityFirst*) and after (*ActivityLater*) the learning tasks

As is shown in Figure 4, higher brain synchronization occurred in Class B, in which students were more actively engaged in the warm-up activities in the lesson than in Class A (FDR-adjusted $q < .05$).

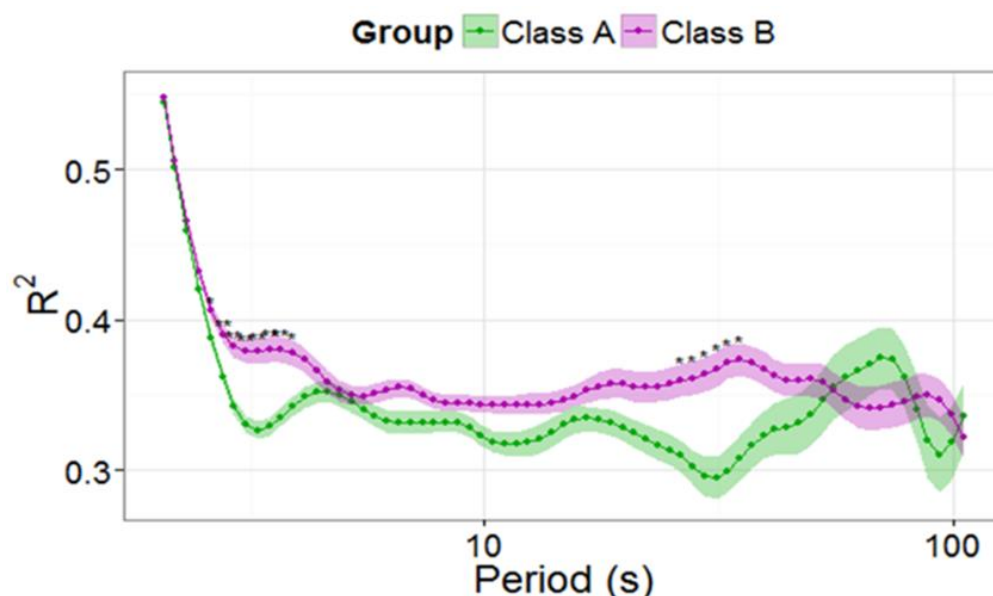


Figure 4: Difference in inter-personal brain synchronization during the warm-up activities between the two classes
*Note: * = FDR-adjusted $q < .05$*

The number of body movements during the warm-up activities, in both Class A and Class B, was greater than that during the language learning tasks. In Class B, a larger number of body movements was observed. Fewer body movements were observed during the language learning tasks in both Class A and in Class B (see Figures 5 and 6).

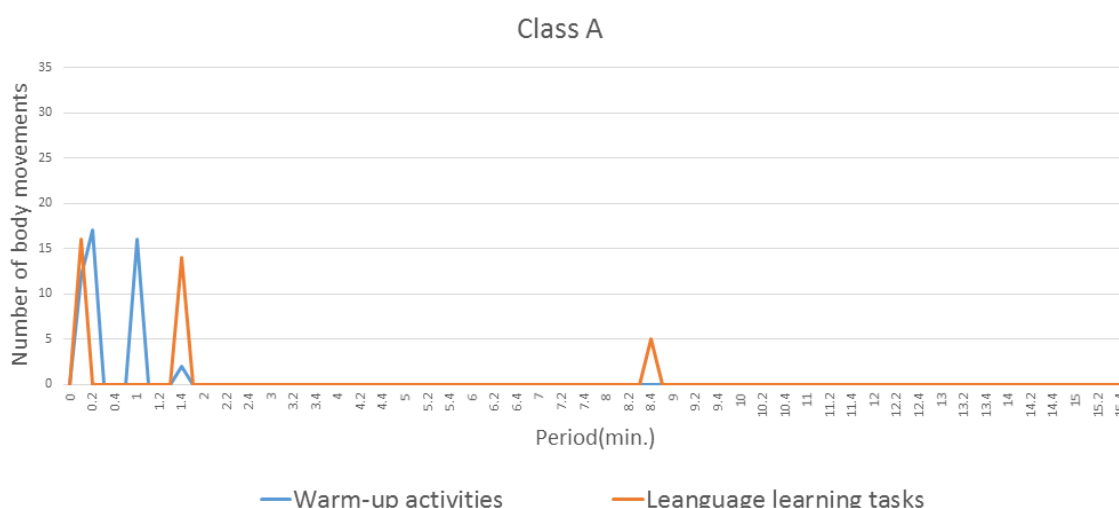


Figure 5: Difference in the number of body movements between warm-up activities and language learning tasks in Class A

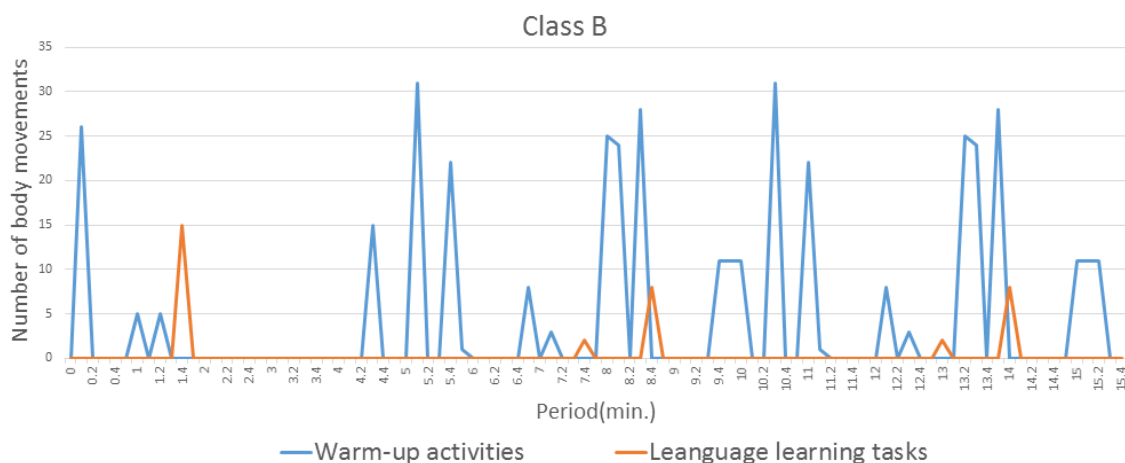


Figure 6: Difference in the number of body movements between warm-up activities and language learning tasks in Class B

Summary of the findings

Research question No.1: Would synchrony between the teacher and the students occur during the warm-up activities; and also, during the language learning tasks of an EFL lesson?

Synchrony was present during the warm-up activities and the language tasks with both classes.

Research question No.2: Would synchrony that occurred during the warm-up activities evoke synchrony in the learning tasks?

According to Figure 3, a comparison of the brain synchronization during the language learning tasks before the warm-up activities with that after the warm-up activities reveals no significant difference. Thus, the synchrony of the warm-up activities before the language learning task did not influence the level of synchrony during the language learning tasks.

Research question No.3: Would the level of synchrony during the warm-up activities be similar to the level of synchrony which occurred during the language learning tasks?

Figure 2 shows that synchrony during the warm-up activities was greater than that during the language learning tasks.

Research question No.4: Would a higher level of synchrony occur when students were actively engaged, as measured by body movements, in the lesson?

According to Figure 4, a higher degree of brain synchronization occurred in Class B, in which more student body movements were observed in the warm-up activities in the lesson compared to that of Class A. The number of body movements is shown in Figures 5 and 6.

Discussion

The inter-personal brain synchronization during the warm-up activities was significantly greater than that during the language learning tasks in both Class A and Class B. There was higher inter-personal brain synchronization in Class B during the warm-up activities in which more active student engagement was observed. These findings were confirmed by the number of body movements. This indicates that synchrony between the students and their teacher was apparently influenced by the students' engagement in the warm-up activities, and that the students' engagement was possibly generated and facilitated by the teacher-student verbal and non-verbal communication characteristics of the warm-up activities.

It is interesting to note that even with a random selection of students and the same teacher using the same lesson plan, groups of students may respond in the unique ways which teachers refer to as 'class chemistry'.

The experiment was designed to eliminate any possible order effect of the warm-up activities on the learning task by placing the warm-up activities in pre- and post-learning activity positions for both classes. This was done in the expectation that the warm-up activities would evoke synchrony in the learning tasks and thus, it was thought necessary to have comparable data. However, it was found that although the warm-up activities raised and sustained synchrony levels between students and the teacher during the course of the warm-up activities, when the warm-up activities were in the pre-learning task position the high degree of synchrony did not carry over to the learning tasks. In fact, the level of synchrony was similar during the learning task segment of the lesson whether the warm-up activities were conducted before or after the learning tasks.

The rationale for including warm-up activities in language lessons centers on intuitive views that warm-up activities have a positive influence on the efficacy of language teaching. This belief has been based on mainly anecdotal evidence. This study offers a replicable evidence-based method of assessing the relationship between teachers and their students which could help researchers investigate whether an integration of warm-up activities and language learning tasks may be beneficial for class groups of students.

Teachers' intuition that the warm-up activities have an inter-personal bonding impact was demonstrated in this study; however, in terms of synchrony, the learning tasks were not connected to the warm-up activities. Identification of the characteristics of warm-up activities that bring students and teachers together would help in the planning of lessons which engage students in learning.

Limitations

The number of the body movements was counted to confirm inter-personal brain synchronization. However, due to an oversight in procedure, inter-rater reliability was not measured. It should have been done, and inter-rater reliability will be measured in future studies. Nevertheless, the recordings of all the lessons were monitored carefully and the students' movements of the seven body parts were counted with great attention by the expert observer. In addition, to ensure the accuracy of the count of the number of the students' body movements, a second expert in the field was recruited to watch the recordings.

Conclusion

Research studies could be designed to explore whether synchrony promotes EFL learning and if so, what aspects of English language learning are influenced and to what degree for which types of learner. Optimal conditions for learner-teacher communication brain synchrony would be a useful subject of study.

This paper reported a study of a native Japanese-speaking teacher's lesson. Worthwhile future studies could make comparisons with lessons taught by English native speakers, and also Japanese- and English-native speakers' team-taught lessons. This area of research holds promise for the improvement of educationally constructive relationships between teachers and students.

Acknowledgements

This study was supported by a Grant-in-Aid for Scientific Research (#26370645) from the Japan Society for the Promotion of Science, a Grant-in-Aid for Scientific Research (#15H01771) from the MEXT, Japan, and by the Cooperative Research Project Initiative of Institute of Development, Aging and Cancer, Tohoku University, Japan (#21 (2014), #44 (2015)). The data presented, the statements made, and the views expressed are solely the responsibility of the authors.

References

- William S. Condon & Louis W. Sander (1974).** Synchrony demonstrated between movements of the neonate and adult speech. In *Child Development* 45, pp456-462.
- A. Grinsted, J. C. Moore & S. Jevrejeva (2004).** Application of the cross wavelet transform and wavelet coherence to geophysical time series. In *Nonlinear Processes in Geophysics* 11, pp561-566.
- Uri Hasson, Asif A. Ghazanfar, Bruno Galantucci, Simon Garrod & Christian Keysers (2012).** Brain-to-brain coupling: A mechanism for creating and sharing a social world. In *Trends in Cognitive Sciences* 16, pp114-121.
- Adam Kendon (1990).** *Conducting Interaction: Patterns of behavior in focused encounters*. Cambridge University Press: Cambridge, UK.
- Alexis Kent (2013).** Synchronization as a classroom dynamic: A practitioner's perspective. In *Mind, Brain, and Education* 7, pp13-18.
- Kerry L. Marsh, Michael J. Richardson & R. C. Schmidt (2009).** Social connection through joint action and interpersonal coordination. In *Topics in Cognitive Science* 1, pp320-339.
- Marleen B. Schippers, Alard Roebroek, Remco Renken, Luca Nanetti & Christian Keysers (2010).** Mapping the information flow from one brain to another during gestural communication. In *Proceedings of the National Academy of Sciences of the United States of America* 107:20, pp9388-9393.
- Greg J. Stephens, Lauren J. Silbert & Uri Hasson (2010).** Speaker-listener neural coupling underlies successful communication, In *Proceedings of the National Academy of Sciences of the United States of America* 107:32, pp14425-14430.
- Tohoku University, Hitachi Co., Ltd. & Japan Science and Technology Agency (2011).** *Senyo LSI no kaihatsu de kogata keiryoka wo jitugen shita chokogata NIRS no shisakuki wo kaihatsu [Developing an ultra-small NIRS prototype utilizing an LSI.]*. Retrieved from <http://www.hitachi.co.jp/New/cnews/month/2011/09/0914.pdf>
- Michael Tomasello (1999).** *The Cultural Origins of Human Cognition*. Harvard University Press: Cambridge, USA.

17

'An international journal publishing high quality, original research': self-evaluative categories in journal descriptions

Davide Simone Giannoni
University of Bergamo
giannoni@unibg.it

Introduction

Disciplinary cultures have been likened to intellectual territories, whose inhabitants share “codes of conduct, sets of values and distinctive intellectual tasks” (Becher 1981: 109). Such factors determine not only the way research is conducted but also how it is published and received within each discipline. Against this backdrop, academic journals play a key role in the dissemination of new scholarship. More specifically, they allow disciplinary communities: (i) to develop legitimate research spaces capable of attracting like-minded scholars with converging interests; (ii) to determine the nature and limits of new research in terms of coverage, methodology and epistemology; (iii) to act as gatekeepers that regulate and police access to such spaces.

While the shift from print to digital has simplified the publication process and made it relatively inexpensive, it has complicated the search for visibility and recognition. Only well-established, reputable journals are likely to publish content that will be widely noticed and eventually cited elsewhere; bibliometric data is often cited as evidence of a journal's relative standing and impact. Thus editors and publishers wield considerable power but are also accountable to readers and subscribers. At the same time, their success requires an ability to attract good submissions. Journals descriptions (JDs) are a common example of self-presentation written with this purpose in mind.

For Tse and Hyland (2010:1880) “surprisingly little has been said about the genres concerned with the distribution, rather than the production, of knowledge. These ‘carrier genres’, however, play an indispensable role in the mechanisms of delivering and promoting knowledge”. Other examples of print-based and/or digital genres belonging to the same category are back-cover blurbs (Basturkmen 2009), publishers' homepages (Gea-Valor 2006)

and book descriptions (Giannoni 2009).

Despite their brevity and apparent simplicity, JDs contain “more than simply a statement of aims and scope, as lexical and rhetorical choices both reflect and define what is valued by the target academic community” (Tse & Hyland 2010: 1888). Accordingly they can provide interesting insights into the ethos and priorities that shape a journal’s development and choice of contents. Their audience is necessarily diverse, as JDs are relevant “to novice researchers deciding on a venue for their next paper; for researchers seeking to narrow their search for personally relevant material; to librarians and course leaders assembling bundles of subscriptions for their universities” (Hyland & Tse 2009: 718).

Variation across disciplines

Despite their pervasive presence in publishers’ websites, JDs have been largely neglected by authors researching academic discourse. The evidence in hand, however, suggests that disciplinary variation is not very great. In a study of JDs in Applied Linguistics, Sociology, Biology and Mechanical Engineering, Hyland & Tse (2009: 710) found “considerable cross-disciplinary uniformity”. Similarities have been observed also between JDs in Applied Linguistics and Medicine, with both favouring “an applied and generally multidisciplinary orientation” harnessing “a range of concerns and audiences” (Giannoni 2008: 214-215). Convergence across disciplines may be explained by commercial considerations but also by a general trend to standardisation noticed in other digital genres. Moreover, it is worth considering that editorial claims about a journal’s coverage may not entirely match its content: Kidd’s (2002) small-scale study of APA journals, for example, shows a mismatch between invited submissions and actual content.

Hyland and Tse’s work (2009) draws on a sample of 200 JDs from 10 academic publishers and only considers titles published in print as well as online. Wordlists were searched for “items carrying evaluative meanings” (p. 707) – mainly qualifiers but also a few nouns and an adverbial. The items were then used to identify five functional categories (Positioning, Standing, Audience, Quality Assurance, Publishing Practices) linked to “tendencies of lexicogrammatical and rhetorical choices in this genre” (p. 707). Unfortunately, the most frequent evaluative items in each discipline are listed but not quantified. The authors conclude that, “the expression of value almost always draws on positive items [...] emphasizing, among other things, reach, novelty, ranking, importance, and scholarliness” (p. 710).

Evaluative categories

By grouping evaluative lexis into semantic categories, an analyst can map out the most prominent values signalled in a given type of text. Hyland and Tse (2009) refer to these categories as 'attributes' but other authors have adopted different labels to indicate the same thing: parameters of value (Thetela 1997), parameters of evaluation (Thompson & Hunston 2000), semantic categories (Pérez-Llantada Auría 2008), evaluative dimensions (Bednarek 2009), values (Breeze 2011). Rather than opt for abstract categorisations based on top-down taxonomies, this strand of studies views 'values' essentially as evaluative parameters built into the lexicon. Although they do not exclude moral principles (expressing ethos), most of the values signalled explicitly in language are of a factual, utilitarian nature. Thus academic discourse signals such attributes as relevance, size and assessment in speech (Swales & Burke 2003), relevance, size, novelty and goodness in research articles (Giannoni 2012) and relevance in lectures (Deroey & Taverniers 2012a, 2012b).

Taking the lead from Hyland and Tse's (2009) pioneering work and building on the lexical approach to the study of values, the present chapter seeks therefore to answer three research questions:

- What are the most frequent evaluative items used in JDs?
- What evaluative categories (i.e. values) do these items encode?
- What do the results indicate about different disciplines?

Material

To assemble an adequate number of texts from reliable sources, a range of international journals were required. For this purpose twenty top-ranking titles, listed by impact factor, were selected for each of the four disciplines investigated by Hyland & Tse (2009), using the Science Edition of Journal Citation Reports (ISI 2013) for Biology and Mechanical Engineering, and the Social Sciences Edition for Linguistics¹⁷ and Sociology. After checking each journal's full title and publisher details, its homepage was consulted to locate the JD. In the case of Linguistics, the first 21 entries had to be included, as one journal (*Language*) appeared to lack a general online description.

Most of the 80 journal homepages considered are hosted by publishers (see Appendix), who tend to adopt a similar template for all of their titles. Where available as an option, the full description rather than its short version was accessed. When deciding what material to incorporate into the JD corpus,

¹⁷ Journal Citation Reports do not provide a separate category for Applied Linguistics.

text detailing the scope, aims and content of each journal was included but not subscription/submission details, benefits to authors or other peripheral content. Topic-lists were also omitted to avoid skewing results and were tagged as ‘aaa’ in the corpus. Below is a quantitative breakdown, inclusive of data in square brackets from Hyland and Tse’s (2009) corpus and SSTTR (Standardised Type Token Ratio), generated using a well-known concordancer (Scott 2007).

	LING	SOC	BIO	ME	All
Tokens per discipline	2,544	2,995	2,850	2,736	11,163
Tokens per text	127.20	149.75	142.50	136.80	139.54
	[232.50]	[239.50]	[146.50]	[214.40]	[208.25]
STTR	40.85	39.53	43.55	40.20	40.73

Table 1: Quantitative breakdown of JD corpus

The maximum divergence in length across the corpus is relatively small, with Sociology JDs 18% longer than those in Linguistics. The STTR was also similar. Greater differences are noticeable between these texts and those in Hyland and Tse (2009), which were significantly longer in all disciplines with the exception of Biology. It is interesting to note that approximately one third of JDs had no title. Most of them were headed ‘Aims’ and ‘Scope’, or ‘About this journal’ and ‘About [name of journal]’. Other options were ‘Description’, ‘Mission Statement’ and ‘Overview’.

Method

A wordlist was generated for each discipline using WordSmith Tools (Scott 2007). At first only the first 20 items in each list were considered but they included hardly any evaluative lexis and consisted almost exclusively of grammar words (*and, of, the*) and disciplinary terms (*language, social, biology*). These are listed in Table 2 below, with the number of occurrences given before each item.

LING	SOC		BIO	ME
157 and	187 the		153 and	163 of
136 the	167 of		138 the	159 the
130 of	157 and		133 of	150 and
84 language	80 in		78 to	63 in
67 in	73 to		69 in	60 to
66 to	52 research		50 a	43 are
47 journal	48 a		43 journal	40 journal
43 research	42 social		39 biology	37 a
40 a	41 is		37 is	37 is
30 articles	38 journal		36 articles	29 papers
30 is	37 on		36 research	26 on
30 on	28 sociological		32 for	26 with
22 theoretical	28 sociology		31 biological	24 for
21 as	27 for		30 are	24 research
21 or	23 that		25 reviews	22 aaa
19 for	21 an		24 as	20 as
19 issues	19 articles		23 on	20 energy
18 an	19 as		22 an	20 engineering
18 are	19 from		21 that	19 thermal
16 speech	19 review		18 rhythms	17 or

Table 2: Top 20 lexical items in JD corpus

It was therefore necessary to broaden the coverage, opting instead for any word with 10+ occurrences in the general wordlist. This produced a total of 181 types, 40 of which were identified as candidates, being either qualifiers (*broad, important, empirical*) or other words associated with a journal's qualities (*including, well, theory*).

To complete the picture, keywords for each discipline (with the whole corpus used as reference corpus) were generated and those with a minimum keyness of 2 and at least 10 occurrences singled out. However, the items of any interest (because evaluative) also occurred in the general wordlist, so no additions were necessary.

Finally, the concordance lines of the 40 candidates were manually inspected so as to narrow down the count to occurrences that: (i) encode an evaluative parameter, and (ii) target the journal and/or its subject, content, authors and readership. The procedure confirmed a total of 33 types, distributed in 185 relevant tokens). These were then grouped into seven semantically related categories whose constituents, arranged in decreasing order, are summarised in the following section.

Results

After normalising the counts to occurrences per 10,000 words, the picture that emerges is that shown in Table 3. Overall the degree of value marking is not dissimilar across disciplines, ranging from 408.8 in LING to 464.2 in ME (+14%). Orientation is by far the most prominent variable, with ten times more markers than Primacy (167.2 vs. 16.5). It is followed by Relevance and Inclusiveness. The other values (Novelty, Worth, Globalism) are less frequent, all at around 31-34 words each. Conversely, in academic ‘About us’ pages it is Globalism and Worth that top the list (cf. Giannoni 2014); Primacy comes last at 16.5, while in ‘About us’ pages it ranks third.

	LING	SOC	BIO	ME	Mean
Orientation	180.8	163.6	108.8	215.6	167.2
Relevance	51.1	71.1	115.8	106.0	86.0
Inclusiveness	70.7	100.2	73.7	51.2	73.9
Novelty	35.4	43.4	35.1	25.6	34.9
Globalism	35.4	40.1	31.6	18.3	31.3
Worth	31.4	23.4	35.1	43.9	33.4
Primacy	3.9	23.4	35.1	3.7	16.5
Total	408.8	464.1	435.1	464.2	443.0

Table 3: Distribution of values (occurrences/10,000 words)

Turning to the degree of differentiation between domains, it is significant that Orientation ranks first in all fields except Biology, whereas Relevance is second in Biology and Mechanical Engineering, and Inclusiveness comes second in Linguistics and Sociology. On the other hand, Globalism is weak in Linguistics and Mechanical Engineering. Finally, Primacy is particularly low in Linguistics and Mechanical Engineering.

The same data, given below (Figure 1) in column chart form, graphically illustrates the greater prominence given to Orientation, Relevance and Inclusiveness across disciplines. Compared to these, all the other values have a relatively low profile.

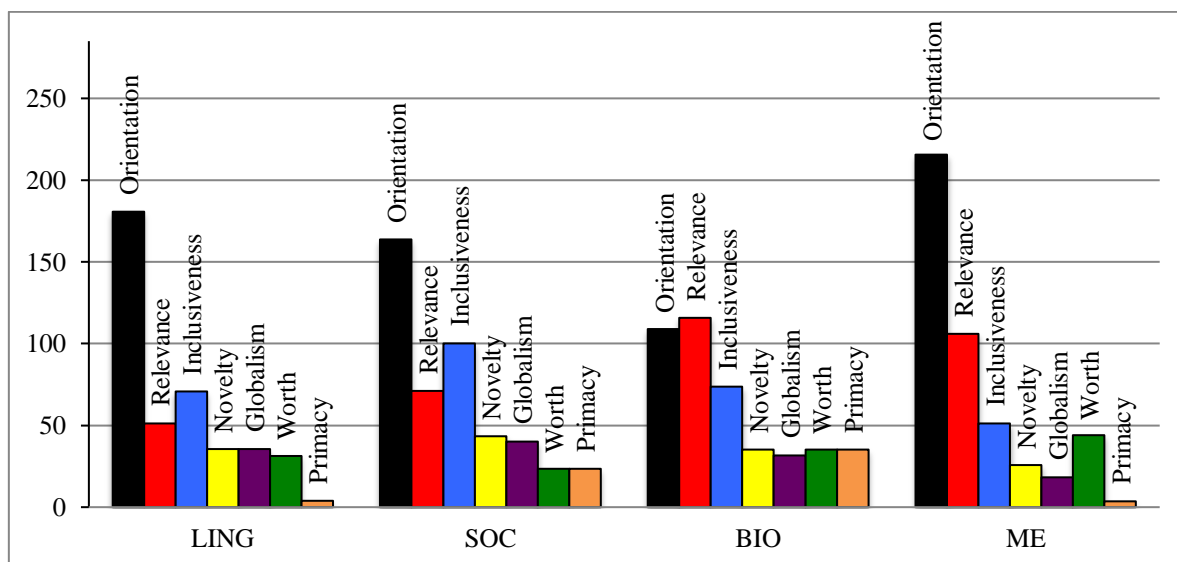


Figure 1: Values present in JD corpus

These findings suggest that Linguistics and Mechanical Engineering are less competitive fields, placing little emphasis on Primacy, compared to Sociology and Biology. Instead, the former are much more likely to emphasise Orientation— arguably because they are more detailed when defining their scope and purpose. Sociology is the field most concerned with negotiating disciplinarity (Inclusiveness) and Mechanical Engineering the one most concerned with Worth.

As in Hyland and Tse (2009), these results can also be interpreted by comparing soft (Linguistics, Sociology) vs. hard (Biology, Mechanical Engineering) disciplines. Table 4 below shows how this pans out for each of the seven values considered. The total figure is very similar but Inclusiveness and Globalism are over-represented in soft JDs, while Relevance and Worth are over-represented in hard JDs.

	Soft	Hard	Mean
Orientation	172.2	162.2	167.2
Relevance	61.1	110.9	86.0
Inclusiveness	85.4	62.4	73.9
Novelty	39.4	30.3	34.9
Globalism	37.7	24.9	31.3
Worth	27.4	39.5	33.4
Primacy	13.6	19.4	16.5
Total	436.4	449.6	443.0

Table 4: Soft vs. hard fields (occurrences/10,000 words)

Orientation

The first group of items consists of qualifiers and nouns that define a journal's orientation, generally in terms of greater or lesser interest in theoretical/speculative as opposed to experimental/empirical research. The results (Table 5) indicate a clear prevalence of the former in the soft disciplines and of the latter in the hard ones. Here and in all subsequent groupings, the concordance lines of each item were inspected to exclude non-relevant occurrences from the count; specific criteria for exclusion are mentioned below.

	LING	SOC	BIO	ME	All
<i>theoretical</i>	22	15	3	6	46
<i>experimental</i>	3	-	9	14	26
<i>scientific</i>	6	1	7	10	24
<i>empirical</i>	7	15	-	1	23
<i>theory</i>	3	12	2	2	19
<i>applied</i>	1	-	3	9	13
<i>applications</i>	1	-	1	10	12
<i>critical</i>	2	6	2	2	12
<i>basic</i>	1	-	4	5	10
Total	46	49	31	59	185

Table 5: Items marking Orientation (occurrences)

The following examples illustrate how each item is used in the corpus:

- *All research papers must clearly explain the theoretical background*
- *It publishes experimental and theoretical papers*
- *It promotes scholarly and scientific discussion of issues*
- *The journal is open to empirical reports and review articles*
- *JMF features original research and theory*
- *Articles on either basic or applied work are suitable.*
- *Industry-relevant papers showing interesting practical applications*
- *It offers critical survey articles of recent research*
- *Articles in the journal pertain to basic and applied chronobiology*

Non relevant occurrences were those incorporating the item in the name of the journal (e.g. *Journal of Experimental Biology*) or of a discipline (*researchers in applied linguistics*) because such uses are not strictly speaking evaluative.

Relevance

Items in the Relevance category draw attention to the significance of

something that is deemed to be worth noting or considering. Hyland and Tse (2009) refer to this quality as 'Importance', which they assign to three adjectives (*important, relevant, significant*). Altogether, Relevance is more prominent in the hard sciences, which may mean that submissions will be treated more selectively or may simply be a self-promotional strategy stressing the journal's quality.

	LING	SOC	BIO	ME	All
<i>interest</i>	2	6	12	8	28
<i>important</i>	2	5	7	2	16
<i>significant</i>	1	2	7	5	15
<i>relevant</i>	3	1	3	4	11
<i>established</i>	2	3	1	5	11
<i>essential</i>	2	3	1	4	10
<i>special</i>	1	1	2	1	5
Total	13	21	33	29	86

Table 6: Items marking Relevance (occurrences)

The examples below illustrate how such items are used in the corpus:

- *Emphasis is on exceptional quality and general interest*
- *Some of the important topics discussed include*
- *Opinion papers and reviews of significant timely issues*
- *The work presented is relevant to each of these disciplines*
- *Journal of Consumer Culture is an established journal*
- *Applied Thermal Engineering provides essential reference material*
- *Q&As on topics of special or topical interest*

Non-relevant cases include metadiscoursal uses (*it is important that papers are presented*) and items occurring within lexicalised collocations (*occasional special sections*).

Inclusiveness

The third group comprises words that define a journal or its content in terms of inclusiveness, i.e. of breadth and scope. This is an essential aspect of editorial policy, which risks being too broad on the one hand and too narrow on the other. The results in Table 7 point to a clear prevalence of Inclusiveness in Sociology, which is a particularly rich and diverse field of enquiry, while it is under-represented in Mechanical Engineering.

	LING	SOC	BIO	ME	All
<i>interdisciplinary</i>	6	9	3	1	19
<i>broad</i>	2	6	10	-	18
<i>both</i>	3	7	1	5	16
<i>wide</i>	4	4	3	3	14
<i>range*</i>	2	2	2	1	7
<i>any</i>	-	1	-	4	5
<i>only</i>	1	1	2	-	4
Total	18	30	21	14	83

Table 7: Items marking Inclusiveness (occurrences)

**when not co-occurring with broad or wide (otherwise total = 19)*

Examples of each item are given below:

- *Social Networks is an interdisciplinary and international quarterly*
- *Work and Occupations provides you with a broad perspective*
- *It focuses on both analytical and experimental research*
- *This interdisciplinary approach spans a wide range of interests*
- *Biological Reviews covers the entire range of the biological science*
- *work on any aspect of the biological foundations of language*
- *Poetics publishes not only advanced research reports but also*

Occurrences not deemed relevant included the use of ‘range’ as a verb (*papers may range from*), of ‘only’ as an intensifier (*Only under compelling circumstances will an Editor*) and of ‘any’ collocating with an external referent (*this journal is a valuable addition to any library*).

Novelty

Novelty is another of the attributes mentioned in Hyland and Tse (2009) for such items as *new*, *current*, *original*, *innovative*, *recent*. Only two of these had 10+ occurrences in the JD corpus (see Table 8 below), with a prevalence of *new*. The overall figure was not too dissimilar across disciplines.

	LING	SOC	BIO	ME	All
<i>new</i>	6	12	7	5	30
<i>recent</i>	3	1	3	2	9
Total	9	13	10	7	39

Table 8: Items marking Novelty (occurrences)

Examples of relevant occurrences:

- *AJS prizes research that offers new ways of understanding the social*
- *The journal provides new information and theoretical approaches*
- *It offers critical survey articles of recent research on specific topics*

Non-relevant instances mainly consisted of uses that were not self-evaluative (e.g. *In recent years, the number of published manuscripts*).

Globalism

Globalism markers highlight a journal's international standing, in terms of contributors, readership and prestige. It was encoded by only two items and is a quality less marked in Mechanical Engineering.

	LING	SOC	BIO	ME	All
<i>international</i>	9	11	6	2	28
<i>world</i>	-	1	3	3	7
Total	9	12	9	5	35

Table 9: Items marking Globalism (occurrences)

Examples of the items in this group include:

- *Bilingualism: Language and Cognition is an international peer-reviewed journal*
- *The FASEB Journal is the world's most cited biology journal*

Fully lexicalised forms were excluded from the count (e.g. *international relations specialists; research into language with relevance to real-world problems*).

Worth

General references to value are grouped under this label, which comprises three different parts of speech. They are all positive and point to attained or pursued worth.

	LING	SOC	BIO	ME	All
<i>high</i>	3	1	7	5	16
<i>advance</i>	2	3	2	4	11
<i>quality*</i>	1	3	1	3	8
<i>well</i>	2	-	-	-	2
Total	8	7	10	12	37

Table 10: Items marking Worth (occurrences)

**when not co-occurring with high (otherwise = 20)*

Examples of these markers are given below:

- *It maintains a high level of readability and scholarship*
- *Only papers which represent a genuine advance in the state of the science*
- *The journal publishes the highest quality empirical and theoretical*

research

- *Group-studies on well defined samples*

Non-relevant occurrences were, again, fully lexicalised forms (*well-being of relevant organisms*) and/or those evaluating an external referent (*under high confining pressures and at high temperatures*).

Primacy

The last group is similar to Hyland and Tse's (2009) Ranking, assumed to include such markers as *main, major, best, excellent, highest, leading, most*. Only the last two of these occurred 10+ times in the present JD corpus. Being ahead of others appears to be prized in Biology and Sociology but almost ignored in Linguistics and Mechanical Engineering.

	LING	SOC	BIO	ME	All
<i>leading</i>	-	6	4	1	11
<i>most</i>	1	1	6	-	8
Total	1	7	10	1	19

Table 11: Items marking Primacy (occurrences)

Examples of these items include:

- *Astrobiology is the leading peer-reviewed international journal*
- *This authoritative journal disseminates the most current findings*

Excluded from the count were instances where 'most' is not an adverb but a quantifier, as in *most of the reviews will be of a specialist nature*.

Conclusions

These results are indicative of the way editors perceive their journal, assert its status and claim a position in the field. The four disciplines considered here deploy the same repertoire of values, with Orientation emerging as the dominant evaluative theme. However, interesting differences were observed across the corpus. Sociology and Biology are more competitive (+Primacy), with less emphasis on the journal's stance (–Orientation). Linguistics and Mechanical Engineering are less competitive (–Primacy) with more emphasis on the journal's stance (+Orientation). These differences cut across the hard/soft field division adopted in Hyland and Tse (2009). Turning to individual domains, Sociology is the one most concerned with defining Inclusiveness; Mechanical Engineering is the least concerned with Inclusiveness (being a strong, well-established field) but the one most concerned with Worth.

The type of analysis attempted here has various advantages and some caveats. First of all it ensures that non-relevant occurrences are omitted by manually investigating all concordance lines. Secondly, it does not attempt to map all the lexis associated with each category but only the most frequent, explicit evaluative items. Admittedly, the grouping and labelling of categories is based on the analyst's judgement and may entail a margin of error. Nevertheless, it provides a cross-section of the most frequent value markers associated with JDs and allows comparisons across the corpus.

This is, to my knowledge, one of the rare follow-ups to Hyland and Tse's work on JDs – a genre that deserves closer investigation, although its familiarity appears to have bred neglect. It would be interesting to extend the same kind of analysis to other evaluative genres (e.g. book reviews, review articles, blurbs, introductions) in these disciplines to uncover similarities with JDs. At the end of the day, however, we are left with a question that cannot be answered by textual analysis alone: whether and to what extent the values embedded in JDs are an expression of disciplinary culture or merely part of the rhetorical toolbox of editorial promotionalism.

References

- Helen Basturkmen (2009).** Back Cover Blurbs: Puff Pieces and Windows on Cultural Values. In Ken Hyland & Giuliana Diani (eds), *Academic Evaluation: Review Genres in University Settings*. Palgrave: Basingstoke, UK, pp68-86.
- Tony Becher (1981).** Towards a Definition of Disciplinary Cultures. In *Studies in Higher Education*, 6:2, 1981, pp109-122.
- Monika Bednarek (2009).** Dimensions of Evaluation: Cognitive and Linguistic Perspectives. In *Pragmatics & Cognition*, 17:1, 2009, pp146-175.
- Ruth Breeze (2011).** Disciplinary Values in Legal Discourse: A Corpus Study. In *Ibérica* 21, pp93-116.
- Katrien L.B. Deroey & Miriam Taverniers (2012a).** "Just Remember This": Lexicogrammatical Relevance Markers in Lectures. In *English for Specific Purposes*, 31/4, 2012, pp221-233.
- Katrien L.B. Deroey & Miriam Taverniers (2012b).** "Ignore That 'cause It's Totally Irrelevant": Marking Lesser Relevance in Lectures. In *Journal of Pragmatics*, 44:14, 2012, pp2085-2099.
- Maria-Lluisa Gea-Valor (2006).** Is an Image Worth a Thousand Words? Net Representations of the Book Industry. In *Revista de Lingüística y Lenguas Aplicadas*, 1:1, 2006, pp37-48.
- Davide Simone Giannoni (2008).** Popularizing Features in English Journal Editorials. In *English for Specific Purposes*, 27:2, 2008, pp212-232.
- Davide Simone Giannoni (2009).** Negotiating Research Values across Review Genres: A Case Study in Applied Linguistics. In Ken Hyland & Giuliana Diani (eds), *Academic Evaluation: Review Genres in University Settings*. Palgrave: Basingstoke, UK, pp17-33.

- Davide Simone Giannoni (2012).** From “Our Methods Apply Equally Well” to “The Model Does a Very Poor Job”: A Corpus-Based Study of Academic Value-Marking. In *Proceedings of the Corpus Linguistics Conference 2011*. University of Birmingham: Birmingham, UK. Available at <http://www.birmingham.ac.uk/research/activity/corpus/publications/conference-archives/2011-birmingham.aspx>.
- Davide Simone Giannoni (2014).** “One of the UK’s Top Universities”: Self-Reported Value Claims in Academic About Us Webpages. Paper presented at the AELFE Conference ‘Academic Communication in a Digital Age’, Stockholm University, Sweden, 12-14 June 2014.
- Ken Hyland & Polly Tse (2009).** “The Leading Journal in Its Field”: Evaluation in Journal Descriptions. In *Discourse Studies*, 11:6, 2009, pp703-720.
- Sean A. Kidd (2002).** The Role of Qualitative Research in Psychology Journals. In *Psychological Methods*, 7:1, 2009, pp126-138.
- Maria Carmen Pérez-Llantada Auría (2008).** Stance and Academic Promotionalism: A Cross-Disciplinary Comparison in the Soft Sciences. In *Atlantis*, 30:1, 2008, pp129-145.
- Mike Scott (2007).** *WordSmith Tools 5.0*. Oxford University Press: Oxford, UK.
- John M. Swales & Amy Burke (2003).** “It’s Really Fascinating Work”: Differences in Evaluative Adjectives across Academic Registers. In Pepi Leistyna & Charles F. Meyer (eds), *Corpus Analysis. Language Structure and Language Use*. Rodopi: Amsterdam, Netherlands, pp1-18.
- Puleng Thetela (1997).** Evaluated Entities and Parameters of Value in Academic Research Articles. In *English for Specific Purposes*, 16:2, 1997, pp101-118.
- Geoff Thompson & Susan Hunston (2000).** Evaluation: An Introduction. In Susan Hunston & Geoff Thompson (eds), *Evaluation in Text. Authorial Stance and the Construction of Discourse*. Oxford University Press: Oxford, UK, pp1-27.
- Polly Tse & Ken Hyland (2010).** Claiming a Territory: Relative Clauses in Journal Descriptions. In *Journal of Pragmatics*, 42:7, 2010, pp1880-1889.

Appendix: Journals covered by corpus

Linguistics

- Brain and Language (Elsevier)
- Journal of Memory and Language (Elsevier)
- Research on Language and Social Interaction (Taylor & Francis)
- Studies in Second Language Acquisition (Cambridge University Press)
- Language and Cognitive Processes (Taylor & Francis)
- Language Learning & Technology (Michigan State University / University of Hawai'i)
- Journal of Speech, Language, and Hearing Research (American Speech-Language-Hearing Association)
- Bilingualism: Language and Cognition (Cambridge University Press)
- Journal of Second Language Writing (Elsevier)
- Applied Linguistics (Oxford University Press)
- Language Teaching (Cambridge University Press)
- Journal of Sociolinguistics (Wiley)
- American Journal of Speech-Language Pathology (American Speech-Language-Hearing Association)
- Journal of Neurolinguistics (Elsevier)
- Applied Psycholinguistics (Cambridge University Press)
- Interaction Studies (John Benjamins)
- Mind & Language (Wiley)
- Journal of Communication Disorders (Elsevier)
- Journal of Child Language (Cambridge University Press)
- Computational Linguistics (MIT Press)

Sociology

- American Sociological Review (SAGE)
- American Journal of Sociology (University of Chicago Press)
- Annual Review of Sociology (Annual Reviews)
- Annals of Tourism Research (Elsevier)
- Sociological Theory (SAGE)
- Population and Development Review (Wiley)
- Sociological Methods and Research (SAGE)
- Sociology of Education (SAGE)
- Social Networks (Elsevier)
- Sociology of Health & Illness (Wiley)
- European Sociological Review (Oxford University Press)

- Journal of Consumer Culture (SAGE)
- Journal of Marriage and Family (Wiley)
- Work and Occupations (SAGE)
- Socio-Economic Review (Oxford University Press)
- The Sociological Quarterly (Wiley)
- Sociology of Religion (Oxford University Press)
- Poetics (Elsevier)
- Social Science Research (Elsevier)
- International Political Sociology (Wiley)

Biology

- PLOS Biology (PLOS)
- Biological Reviews (Wiley)
- Physics of Life Reviews (Elsevier)
- eLife (eLife Sciences Publications)
- BMC Biology (BioMed Central)
- Philosophical Transactions of the Royal Society B: Biological Sciences (Royal Society)
- FASEB Journal (FASEB)
- Bioscience (American Institute of Biological Sciences)
- Proceedings of the Royal Society B: Biological Sciences (Royal Society)
- Quarterly Review of Biology (University of Chicago Press)
- Bioessays (Wiley)
- Biology Direct (BioMed Central)
- Bioelectrochemistry (Elsevier)
- Geobiology (Wiley)
- Biology Letters (Royal Society)
- Journal of Biological Rhythms (SAGE)
- Interface Focus (Royal Society)
- Journal of Experimental Biology (The Company of Biologists)
- Chronobiology International (Informa)
- Astrobiology (Mary Ann Liebert Inc)

Mechanical Engineering

- Progress in Energy and Combustion Science (Elsevier)
- International Journal of Plasticity (Elsevier)
- Journal of Vibration and Control (SAGE)
- Proceedings of the Combustion Institute (Elsevier)
- Combustion and Flame (Elsevier)

- IEEE/ASME Transactions on Mechatronics (IEEE/ASME)
- Aerosol Science and Technology (Taylor & Francis)
- International Journal of Machine Tools & Manufacture (Elsevier)
- Journal of Aerosol Science (Elsevier)
- Applied Thermal Engineering (Elsevier)
- International Journal of Thermal Sciences (Elsevier)
- Wind Energy (Wiley)
- International Journal of Heat and Mass Transfer (Elsevier)
- Mechanical Systems and Signal Processing (Elsevier)
- Nonlinear Dynamics (Springer)
- Journal of Fluids and Structures (Elsevier)
- Tribology Letters (Springer)
- Tribology International (Elsevier)
- Experimental Thermal and Fluid Science (Elsevier)
- International Journal of Mechanical Sciences (Elsevier)

18 **A comparison study of teacher supported peer feedback and teacher feedback for EFL writing**

Dr Huahui Zhao
Manchester Metropolitan University
H.Zhao@mmu.ac.uk

Abstract

Despite increasing interest in using peer feedback in the teacher dominated EFL writing instruction in China, few studies have provided prolonged evidence on teacher supported peer feedback there. It is important to examine the different nature and effectiveness of peer and teacher feedback as such findings will provide valuable implications for how to effectively integrate peer assessment into entrenched teacher assessment.

The current cross writing tasks case study explored the different nature and effectiveness of peer and teacher feedback in an EFL writing class at a university in China where peer assessment was continuously supported by different teacher intervention strategies. Eighteen second-year English majors along with their writing tutor participated in this study for over sixteen weeks across eight writing tasks. Three data sets were collected, including first draft wherein peer and teacher feedback contained, revised drafts wherein peer and teacher feedback was incorporated, and students' perception data of peer and teacher assessment.

Results showed that learners provided significantly smaller amount of and less varied feedback than their tutor and they used significantly less peer than teacher feedback in their revised drafts. However, the students valued peer assessment in terms of its complementary role to teacher assessment and its distinguishable role from teacher assessment. This paper concluded with possible recommendations for future research on peer assessment for EFL writing.

Introduction

In the past two decades, peer feedback has been increasingly used in ESL/EFL (English as a Second/Foreign Language) writing instruction. This has stimulated growing concerns about the use of peer feedback for ESL/EFL writing. However, a search and reading on empirical studies on

peer feedback¹⁸ so far reveal the following knowledge gaps in existing research agenda on peer assessment:

1. Few studies are conducted in China where probably the largest number of English learners are instructed in English in the world;
2. Studies on written peer feedback decreased dramatically in the past decade, even though written feedback is still a principal means of assessing writing;
3. Almost no longitudinal evidence has been provided regarding the impact of peer feedback on writing, albeit peer feedback has been suggested to vary from writing tasks and peer commentators (Cheng, 2004; Connor & Asenavage, 1994; Hedgcock & Lefkowitz, 1992; McGroarty & Zhu, 1997);
4. The relationship between the nature and the effectiveness of peer feedback has been hardly examined; however, research on teacher feedback has suggested their tight relationships (Conrad & Goldstein, 1999; Ferris & Hedgcock, 1998).

The current study aimed to fill in the aforementioned knowledge gaps via a) setting the research context in an EFL writing classroom in a university in China, b) exploring the amount, nature and impact of peer written feedback in comparison with that of teacher written feedback, c) collecting data across nine writing tasks with different peer commentators, and d) examining the relationship between the nature and use of peer feedback. It is expected that this study could shed new insight on the role of peer feedback in facilitating Chinese university learners improving their writing quality based on content analysis of feedback, writing drafts and learners' perception data.

Literature review

Investigations into feedback for EFL writing has been stimulated by the critical role of feedback in writing, as argued by Hyland and Hyland (2006a):

Feedback is a key component of teaching second language writing, with process, social constructivist, and academic literacy approaches all employing it as a central part of their instructional repertoires. (p.15)

¹⁸ The literature search started from journal papers in and after 1990, including *Journal of Second Language Writing*, *Assessing Writing*, *ELT Journal*, *TESOL Quarterly*, *Language Testing*, *Language Teaching*, *International Journal of Applied Linguistics*, *Applied Linguistics*, *System*, *Journal of English for Academic Purpose*, *Language Teaching Research*, *Language Learning*, *Annual Review of Applied Linguistics* and *Review of Educational Research*. In addition, key earlier studies and key theoretical references reported in the reviewed studies were also searched.

The argument indicates two important theoretical frameworks supporting peer assessment.

The process and product theory regards writing as a cognitive process wherein a supportive environment should be established to allow students to be acknowledged as writers and encouraged to take risks and engage in creating meaning decisions for writing (Lockhart & Ng, 1995). This is highlighted in the process approach writing theory. Flower (1989) suggests peer response as a case in point to illustrate how writing and context interact. Hayes (1996) thereafter modified the Flower and Hayes 1981 model by adding the social environment into the task context, suggesting writing as a social construct involving the interaction with audiences and collaborators. The role of audiences and collaborators highlights the value of peer assessment for EFL writing which creates the chance for writers to interact with peer audiences to improve the quality of their writing from the audience's perspective.

Another important theoretical root for peer assessment is the socio-cultural theory. This theory suggests that cognitive development is a socially and culturally mediated process via external sources including more knowledgeable peers (e.g. Aljaafreh & Lantolf, 1994; Donato, 1994; Villamil & Guerrero, 2006; Vygotsky, 1978). Through the notion of the Zone of Proximal Development (i.e. the distance between the actual developmental level determined by independent problem solving and the higher level of potential development determined through problem solving in collaboration with more capable peers or seniors, Vygotsky, 1978:86), Villamil and Guerrero (1996) articulate that:

..his [Vygotsky's] concept of 'zone of proximal development' which recognises the importance of peer assistance in the solutions of tasks and, consequently, in learning, seems particularly applicable to the kind of collaborative instructional activity that occurs during peer revisions. (p.54)

Socio-cultural theory lends strong support to integrating peer assessment into writing instruction to develop learners' writing ability. The two theoretical roots have been adopted as the main frameworks for current studies in peer assessment.

Examination of the quantity of feedback has revealed that students provide less feedback than writing teachers. For instance, Yang et al. (2006) reported that the writing teacher made more feedback instances (N=235) than the 12 Chinese EFL peer reviewers (N=225). However, teacher feedback was found to not cover all the language areas that were addressed

by peer feedback. Hu's (2005) study observed that the students made many valid suggestions which he, as the writing teacher, had not provided while commenting on students' writing. Accordingly, a complementary role of peer to teacher feedback in helping students with their revisions was suggested.

As for the nature of peer and teacher feedback, their focus and type were found to be very different. Caulk (1994) asserted that the peer reviewers focused more on content whereas she, as the writing teacher, paid more attention to form and clarity given their importance for academic writing. In line with this, Yang et al (2006) noted that peer feedback was more meaning-oriented whereas teacher feedback was balanced on both surface (e.g. grammar, vocabulary and sentence structure) and meaning-oriented (e.g. content) changes. As far as the type of feedback is concerned, Caulk (1994) observed that teacher feedback was more general, without supplying revision strategies, whereas peer feedback was more targeted, providing specific revision strategies.

Finally, teacher feedback was observed to be incorporated more frequently in students' next drafts than peer feedback. Paulus (1999) claimed that the 12 undergraduate international students in his pre-freshmen composition class were far more likely to incorporate teacher feedback (87%) than peer feedback (51%) in their revisions. As with Paulus' participants, the 12 Chinese college EFL learners in Yang et al.'s 2006 study also appeared to use a larger amount of teacher feedback (90%) than peer feedback (76%) in their subsequent drafts.

The existing studies reveal that learners confirm the value of peer assessment in particular in terms of its benefits in developing content (Mangelsdorf, 1992) and increasing feedback resource (Jacobs, Curtis, & Huang, 1998), and the chance of reading peer writers' assignments via providing peer feedback (Hu, 2005; Tsui & Ng, 2000). On the other hand, studies also reveal learners' concerns over the quality of peer feedback (e.g. Connor & Asenavage, 1994; Paulus, 1999).

To mitigate learners' concerns, training in peer assessment has been suggested. Studies in training (Berg, 1999; Patri, 2002; Rollinson, 2005; e.g. Stanley, 1992; Tuzi, 2004) justify it as engineering the efficiency of peer assessment. Hu (2005) showed that training made students more positive towards peer review, provided far superior comments both in quantity and quality to those made by untrained students (p.334 - 335). This value has been corroborated by the findings in Min's (2005) project with 18 EFL sophomore students in her composition class in Taiwan. The

data before and after teacher guidance on peer assessment revealed that the provision of guidance resulted in significantly more comments on global issues (e.g. idea development and organisation), more specific and focused ways in elaborating writers' ideas, and increasing confidence for less-advanced reviewers in regarding themselves as competent readers.

The existing studies have provided important guidelines for the current study in a number of ways. Firstly, following the existing research line, the current study will use teacher feedback as the comparison baseline to suggest the nature and effectiveness of peer feedback. Secondly, learners' perceptions were collected to supplement feedback and assignment data. Last but not least, training was provided to support peer assessment throughout the research period.

Current study

The following four hypotheses were formed based on the existing research evidence:

- **Hypothesis 1:** The writing teacher will provide more feedback than the peer reviewers on the same assignments.
- **Hypothesis 2:** The focus and the type of peer and teacher feedback are different.
- **Hypothesis 3:** Teacher feedback will be used more frequently than peer feedback in students' revisions.
- **Hypothesis 4:** Learners appreciate peer assessment as a form of writing assessment.

Methods

Participants

Eighteen second-year English majors with their writing teacher participated in this study over 16 weeks. These students obtained high English examination marks in the university matriculation examination (on average 120/150) and were assessed as intermediate English learners by their writing teacher who came from United States and had taught EFL writing for about two years in China. Before this, he was a professor in a college in United States for over 30 years and could, thus, be considered to be an experienced instructor. Full consent of undertaking this study was achieved from participants via written consent forms.

Instructional context

The prevailing English writing pedagogy in the research context was suggested to be almost completely teacher-dominated and examination-

oriented by the students and the writing teacher in the pre-assessment survey (i.e. before the use of peer feedback in their writing class) which asked about learners' previous English writing learning experiences (e.g. whether learners participated in peer assessment in their previous English writing classes; what kind of feedback learners received from their teachers; and how they would describe their previous English writing learning experience). Consequently, none of the students claimed to have had previous experiences of peer assessment. The influence of learners' previous teacher- and examination-driven learning experiences on the use of peer assessment will be discussed in the discussion section.

Training in peer assessment

Training in peer assessment was provided considering the participating students' lack of experience in peer assessment.

The teacher demonstrated how he provided feedback on writing prior to and throughout the four-month period of this study. The teacher read through student writing samples, stopped at the problematic areas, and demonstrated how he provided feedback on those areas. Grammar and wording were the two common aspects throughout all training sessions given that accuracy is the main criteria for examination essays.

Students' ongoing support needs were addressed based on the researcher's (i.e. the author's) ongoing analysis of peer feedback. The small amount of indirect peer feedback (i.e. peer feedback without providing revision solutions, as opposed to direct feedback which provided revision solutions) in the first three writing tasks led to this support during the remainder of the study in response to those problematic areas where students were unable to provide revision suggestions. Thus, the teacher encouraged students to highlight these areas with question marks and/or provide indications of the problems (for example 'needs rewording').

Data collection

Students were randomly paired by assigning a number from 1-9 and the two students receiving the same number were grouped in one pair, thus having different collaborators across tasks. Eight sequential steps were followed to integrate peer assessment into existing teacher assessment, as shown in Figure 1.

The data collected included: i) students' first drafts (see Appendix A) which were commented on by peers and the teacher, ii) students' revised drafts wherein peer and teacher feedback was incorporated (see the three

bold steps in Figure 1), and iii) students' perceptions of peer assessment via interviews. Eleven students were invited to participate in the interview sessions. Two interview questions were asked: (1) Which assessment mode do you prefer: peer assessment, teacher assessment or both? and (2) Can you explain your preference?

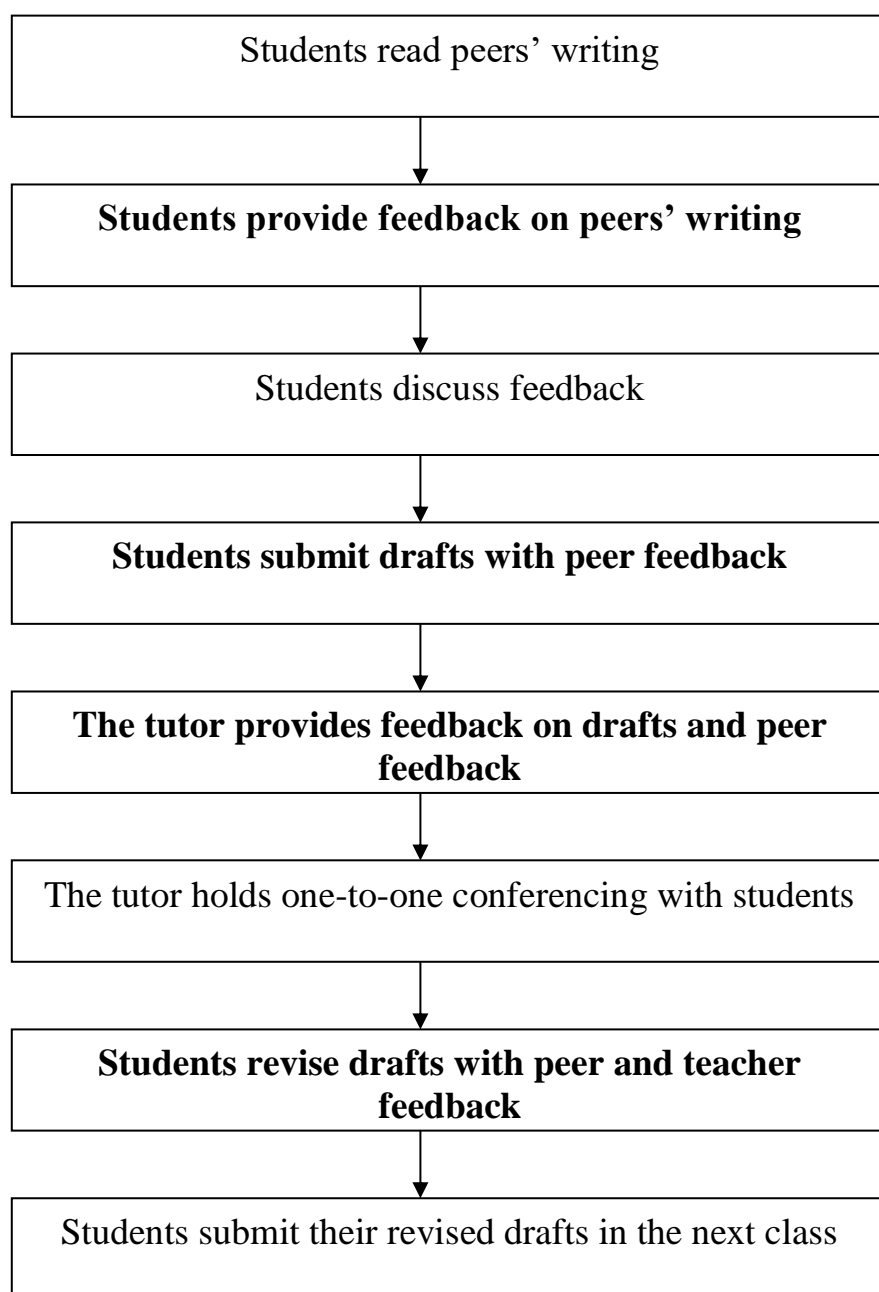


Figure 1: Flowchart of the integration of peer feedback

Data construction of feedback data

Table 1 provides an overview of the construction of feedback data, with a summary of the dates when the assignments were submitted, the number of assignments that were collected at each date, the gender of the peer reviewers, and the genre of assignments.

Date	Quantity	Gender: by peer assessors	Genre
28/02/06	10	5: male 5: female	10: argument
07/03/06	14	7: male 7: female	14: letter
14/03/06	11	5: male 6: female	11: poetry
11/04/06	9	5: male 4: female	2: fiction 6: argument 1: research paper
25/04/06	6	4: male 2: female	2: fiction 3: argument 1: research paper
08/05/06	7	3: male 4: female	5: fiction 2: research paper
16/05/06	10	5: male 5: female	1: fiction 8: argument 1: research paper
30/05/06	9	4: male 5: female	1 poetry 8: argument
IN TOTAL	76	38: male 38: female	41: argument 14: letter 12: poetry 5: fiction 4: research paper

Table 1: Data construction of peer feedback

Table 1 shows that the number of assignments collected at each date varied, depending on the quantity of assignments submitted and the continuity of drafts (i.e. assignments with second drafts widely different from first drafts were excluded from the dataset). The gender of the peer commentators of the 76 assignments was rather balanced. Genres of student writing covered a wide range, including fiction, letters, poetry, arguments and research papers as a result of self-selected writing topics. Among them, arguments appeared to be the most popular.

Data analysis

An inductive approach was adopted to undertake content analysis of feedback and students' revisions with feedback in response to the first three hypotheses.

One text-specific corrective feedback was selected as the unit of analysis. Text-specific corrective feedback refers to feedback provided in the text on a specific language area in students' writing, aiming at helping writers to improve the appropriateness of this area.

Text-specific corrective feedback was analysed in terms of its focus, type and use. The focus of feedback was categorised into grammar, wording, sentence structure, organisation, style and content, the most focal areas of teacher feedback identified by Hedgcock and Lefkowitz (1994). The type of feedback was classified into direct feedback and indirect feedback to differentiate the degree of explicitness of feedback which is considered to have a large influence on students' use of feedback (e.g. Ferris, 2006; Ferris & Hedgcock, 1998; Hyland & Hyland, 2006b). The use of feedback was examined based on the incorporation of peer and teacher feedback in subsequent drafts. Appendix B provides an overview of the coding categories; definitions and examples are given.

The interview data was analysed thematically, namely: learners' preference for the assessment mode and their justification for their preference.

Results

Hypothesis 1: The writing teacher will provide more feedback than the peer reviewers on the same assignments.

Total amount of peer and teacher feedback

A total of 620 feedback instances ($SD=7.37$) were provided by the peer evaluators on the 76 assignments, compared to 1,552 by the teacher on the same assignments ($SD=12.46$). The teacher, thus, provided over twice the quantity of feedback than the peers. Figure 2 demonstrates that the teacher generated a larger amount of feedback on the eight sets of assignments across genres.

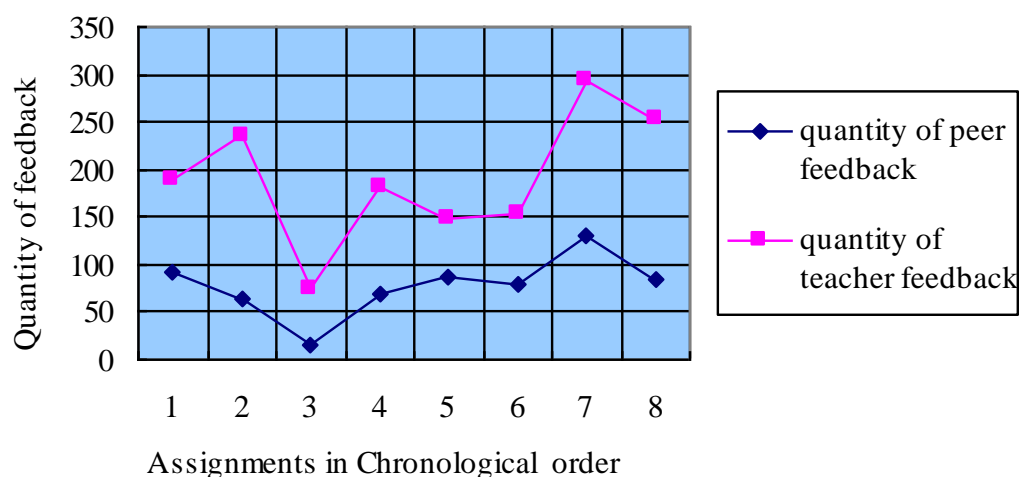


Figure 2: Total quantity of feedback

The finding was further investigated by the employment of paired-sample t-tests on the amount of peer and teacher feedback. The test results suggest that the difference in the amount of peer and teacher feedback was statistically significant ($t=10.634$, $df = 75$, $p < .05$). The negative value of the difference in the mean suggests that the teacher made a mean of 12.26 more instances than the peer reviewers on the 76 assignments.

Amount of peer and teacher feedback on each focus

Similar findings were obtained when the amount of peer and teacher feedback was examined in terms of their focus and type via paired-samples t-tests. With the exception of feedback on organisation, significant differences were observed in feedback on all the other six foci: feedback with grammar ($t=8.749$, $df=75$, $p < .05$), wording ($t=5.709$, $df=75$, $p < .05$), mechanics ($t=5.445$, $df=75$, $p < .05$), sentence structure ($t=3.987$, $df=75$, $p < .05$), style ($t=3.670$, $df=75$, $p < .05$) and content ($t=5.321$, $df=75$, $p < .05$). The negative t values indicate that the writing teacher made more feedback than the peers on these six foci.

Table 2 provided the specific amount feedback on each focus which revealed the writing teacher provided over 65% of feedback instances on all focuses including organisation.

Nature of feedback		FG	FW	FM	FSS	FO	FS	FC
PF	N	76	76	76	76	76	76	76
	Sum	226	239	48	25	13	7	62
	SD	2.935	4.150	1.005	.575	.444	.372	1.251
	% of Total Sum	26.3%	34.9%	23.0%	28.7%	33.3%	17.1%	24.5%
TF	N	76	76	76	76	76	76	76
	Sum	632	446	161	62	26	34	191
	SD	6.626	4.212	2.508	1.080	.703	.790	2.910
	% of Total Sum	73.7%	65.1%	77.0%	71.3%	66.7%	82.9%	75.5%

Table 2: Case summaries: focus of peer and teacher feedback

FG=Feedback on Grammar; FW=Feedback on Wording; FM=Feedback on Mechanics; FSS=Feedback on Sentence Structure; FO=Feedback on Organization; FS=Feedback on Style; FC=Feedback on Content

Amount of direct and indirect peer and teacher feedback

Paired sample t-tests showed the teacher provided significantly more direct ($t=9.571$, $df=75$, $p<.05$) and indirect feedback ($t=4.198$, $df=75$, $p<.05$) than the peer reviewers. To be specific, students provided 452 instances of direct feedback ($SD=5.433$) whilst the teacher provided 1,259 instances of direct feedback ($SD=11.575$). In other words, in contradiction with Caulk's (1994) finding, the teacher, rather than the students, offered more revision solutions to students' writing. This is graphically demonstrated in Figure 3.

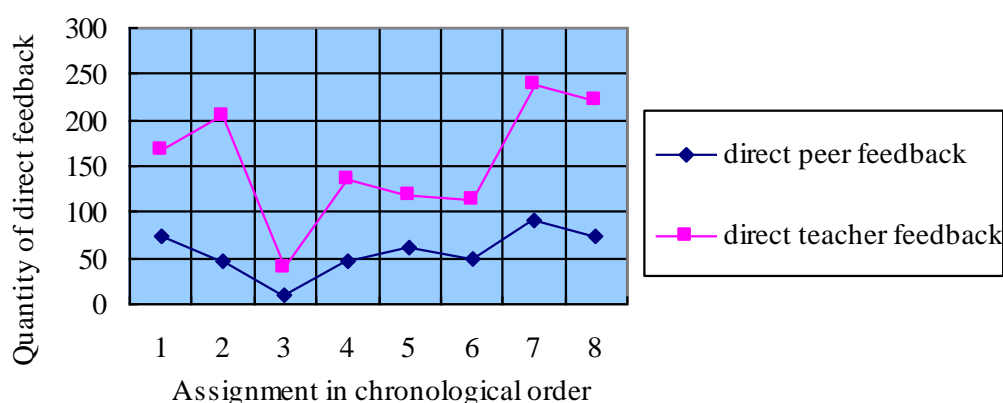


Figure 3: Quantity of direct feedback

It can be observed from Figure 3 that the teacher generated more direct feedback than the peer reviewers on all the eight assignment datasets. Figure 4 shows that all the collected assignment datasets contained more indirect teacher than peer feedback. To be specific, the students provided 156 instances of indirect feedback ($SD=2.55$) whilst the teacher provided 256 instances of indirect feedback ($SD=3.34$).

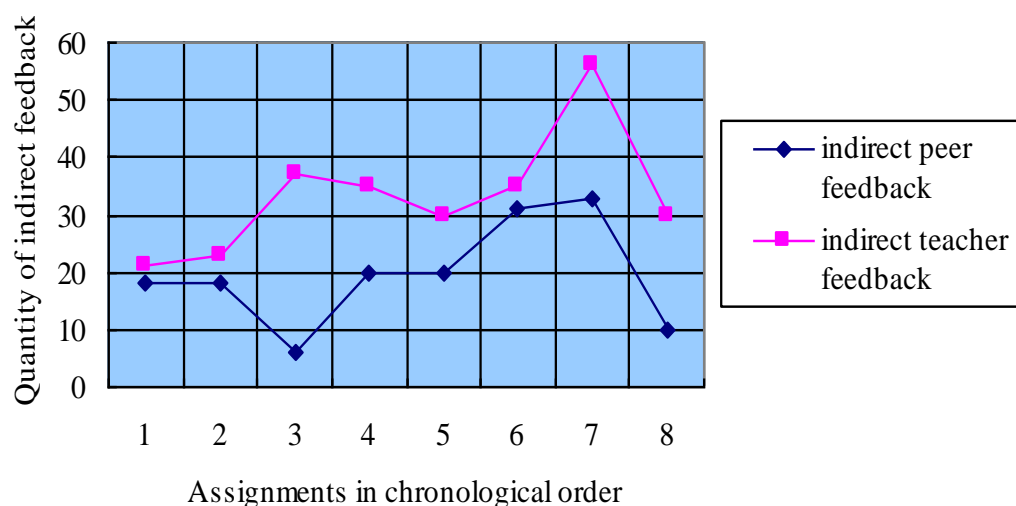


Figure 4: Quantity of indirect feedback

Based on the foregoing, we can see that hypothesis 1 has been confirmed. That is to say, the writing tutors provided significantly more feedback on the 76 assignments in terms of the total amount, the amount of feedback on each focus and of each type.

However, teacher feedback was observed not to cover all the language areas which the peer collaborators addressed, congruent with the finding in Hu's study of Chinese ESL college learners (Hu, 2005), 49% of peer feedback points were not addressed by the writing teacher in this study. This percentage could be larger because the teacher provided comments on peer feedback, and the 49% peer feedback instances only referred to those which were not commented by the teacher. On the other hand, the teacher generated a total of 82% of feedback instances on students' writing which were not provided by the peer commentators. In this sense, peer and teacher feedback could play a complementary role in assisting students in revising their drafts.

Hypothesis 2: The focus and the type of peer and teacher feedback are different.

Figures 5 and 6 show broadly similar focus of peer and teacher feedback, with grammar and wording as the most frequently commented language aspects and the style and organisation as the least frequently commented aspects. On the other hand, the students provided 13 more instances of feedback on wording than grammar whilst the teacher provided 186 more instances of feedback on grammar than wording. The students provided six more instances of feedback on organisation whilst the teacher provided eight more instances of feedback on style than organisation.

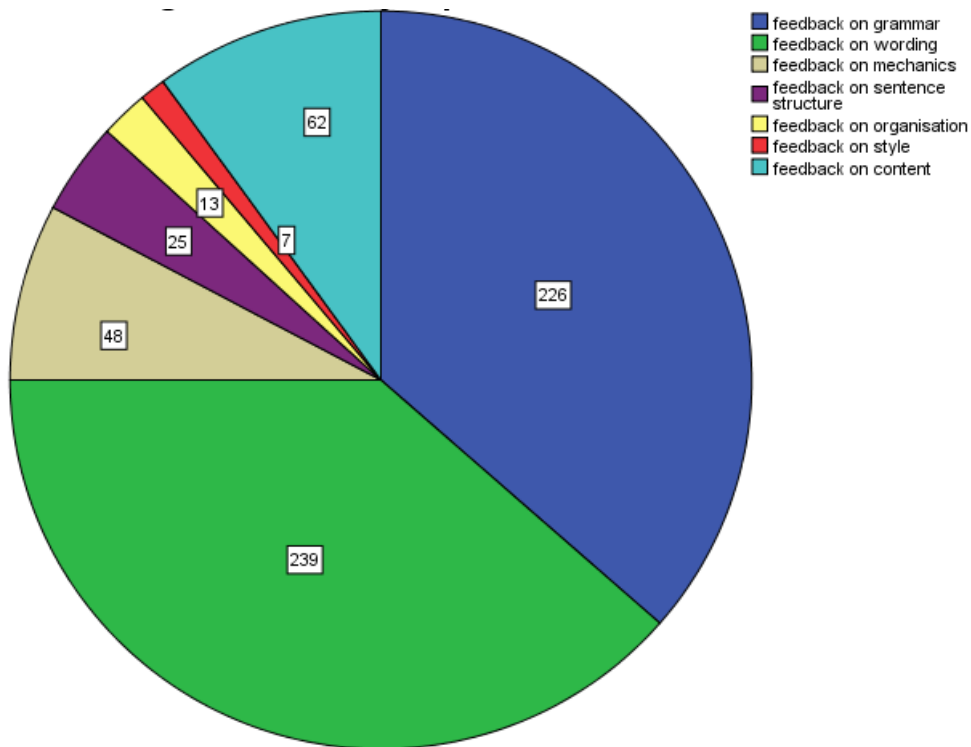


Figure 5: Quantity of peer feedback on each focus

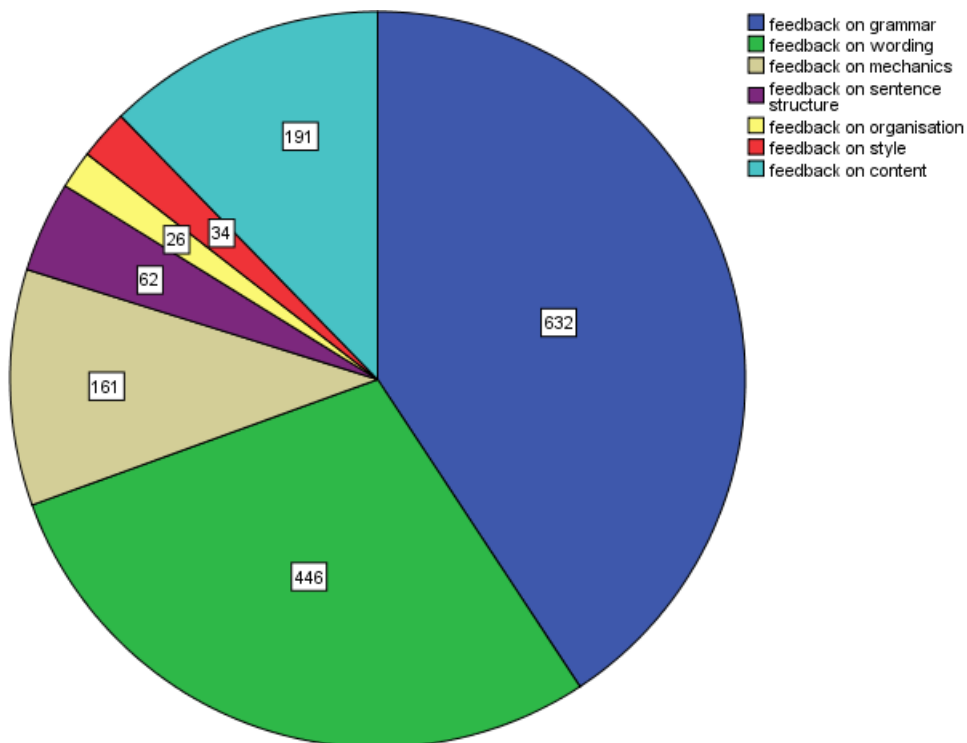


Figure 6: Quantity of teacher feedback on each focus

Figures 7 and 8 show the similarity of peer and teacher feedback type.

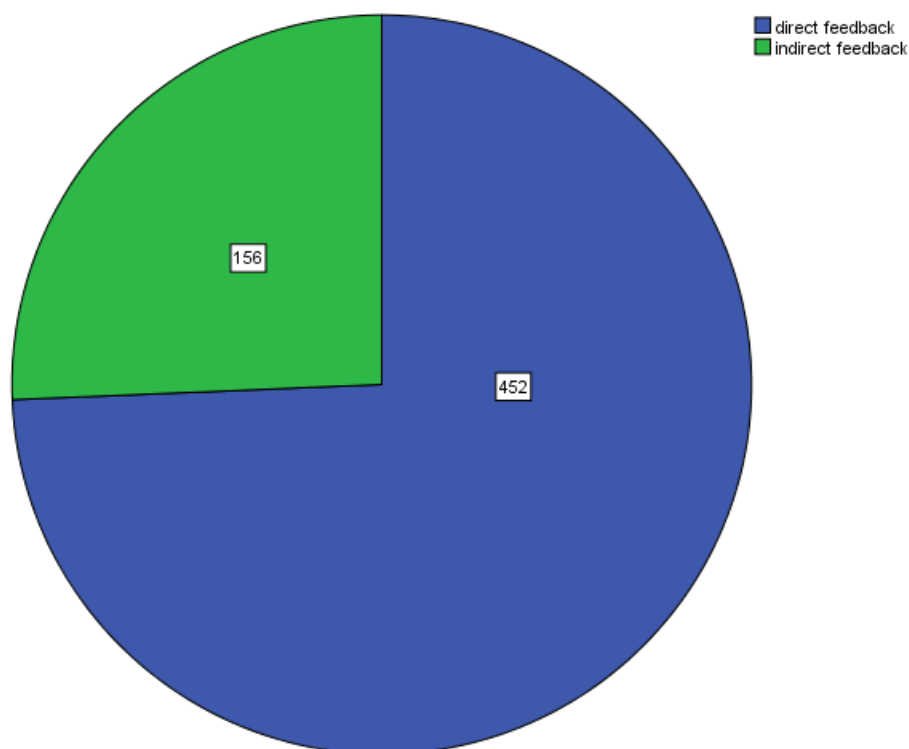


Figure 7: Quantity of direct and indirect peer feedback

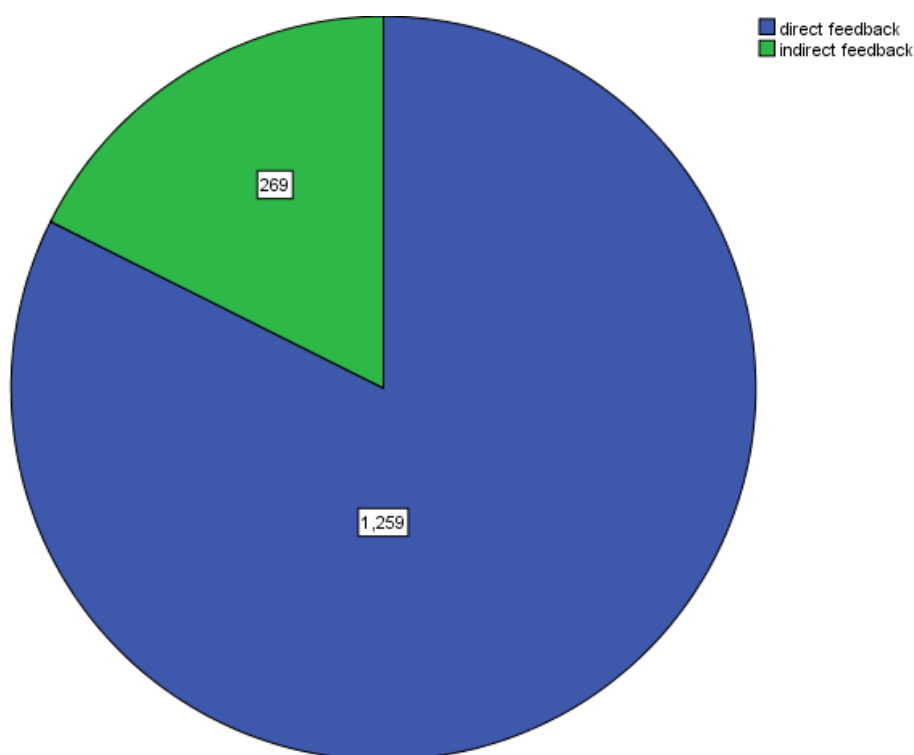


Figure 8: Quantity of direct and indirect teacher feedback

The vast majority of teacher and peer feedback was direct feedback. In this sense, Hypothesis 2 has been rejected. That is to say, the type of peer and teacher feedback was similar in this study.

Hypothesis 3: Students will use teacher feedback more frequently than peer feedback in their revisions.

Paired-sample t-tests were employed to explore this assumption. The test results show that students fully ($t=11.522$, $df=75$, $p<.05$) or partially used ($t=2.543$, $df=75$, $p<.05$) significantly more teacher feedback than peer feedback in their revised drafts.

Figure 9 graphically demonstrates the difference in the quantity of fully used peer and teacher feedback in each assignment dataset. It can be seen from Figure 9 that the students incorporated more teacher feedback than peer feedback instances in their assignments. This finding has, thus, confirmed the significant difference achieved via paired-samples t tests.

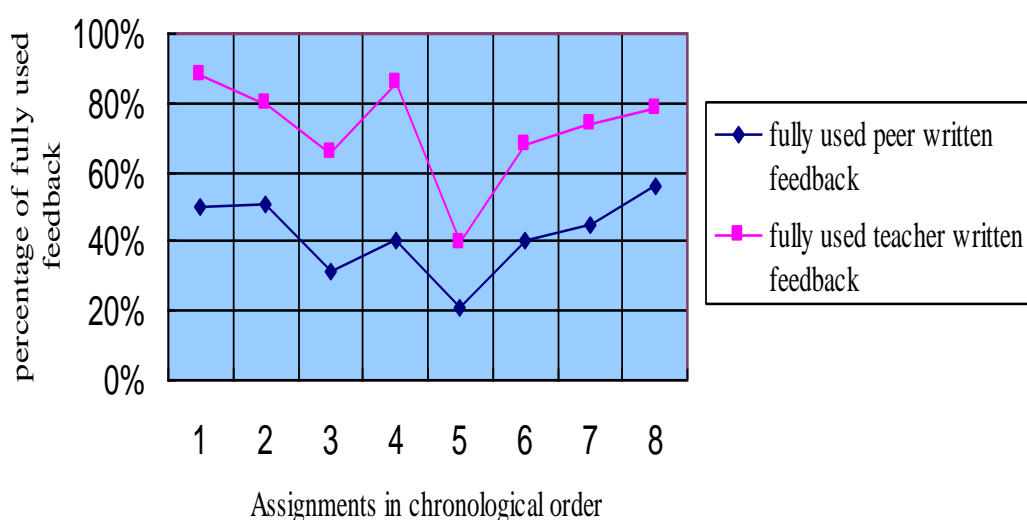


Figure 9: Fully used written feedback

To be specific, we can observe in Table 3 that 81.1% of the total feedback points which were fully used in students' revised drafts were made by the teacher, against 18.9% by the peer readers. Similarly, 62.7% of the comments that were partially used in students' revisions were generated by the teacher, against 37.3% by the peer commentators. We can thereby see that teacher feedback in this research context exerted greater impact on students' revisions.

Nature of feedback		Fully used	Partially used
PF	N	76	76
	Sum	267	53
	Std. Deviation	3.772	1.286
	% of Total Sum	18.9%	37.3%
TF	N	76	76
	Sum	1147	89
	Std. Deviation	10.345	1.829
	% of Total Sum	81.1%	62.7%

Table 3: Use of peer and teacher feedback

We can see that Hypothesis 3 has been supported by the current study. That is to say, teacher feedback affected students' revisions more significantly than peer feedback in this study.

Hypothesis 4: Learners appreciate peer assessment as a form of writing assessment.

All the eleven interviewees expressed their willingness to have peer assessment as a form of assessment to supplement teacher assessment with regard to the unique values of peer assessment.

Firstly, peer assessment created opportunities to read and learn from peers' drafts, as Shu said:

I feel it will help me because I may *learn some words* from them. If I meet words I don't know, I will look them up, finding out whether it is right or not. If it is wrong, I will *pay attention to it next time when I write*. In addition, I can *learn some styles and use them in my own writing later*. When I provide peer feedback, I am learning.

(Shu/prei/P49)

Similar views were expressed by Mo:

I can learn from their writing by reading it, another resource for learning. Two students will write in two different styles. In addition, I can learn some new words, sentence structures and organisation from theirs.

Mo/posti/P31)

In addition, students claimed that reading peers' work also encouraged them to be more reflective about their own writing, as articulated by Zhang:

Their mistakes can also reflect some of mine. I can also see my shortcomings via reading their errors. (Zhang/posti/P5)

Similarly, Ping also perceived peer feedback as a resource to reflect on her own writing:

For example, in poems, I wrote some words she was not familiar, so she commented that it was beautiful but I used some too flowery words. So I have to think about my purpose of writing a poem, to make it understood or to make it beautiful. So I have to weight it up and integrate it into final versions. (Ping/posti/P41)

Peer assessment, it is therefore suggested, might play a metacognitive role in prompting learners to reflect further on their own writing. The value of reading peers' writing aligns with the 16 students in Leki's (1990) study which highlighted the value of reading peers' writing.

Secondly, peers understood each other's writing better than the tutor given their shared educational and cultural background. Shu justified:

Peers, because sometime we think differently from the tutor but the same as peers, sometime they can understand what I am writing about but the tutor cannot.

(Shu/prei/P42)

Similarly, Shen talked about 'the smaller gap' between students than between students and the tutor, in terms of English proficiency, life experiences and attitudes towards social affairs. This, from his point of view, contributed to students' better understanding of others than did the tutor:

I sometime cannot understand peers' work due to different logics. However, we have a wider gap with the tutor than peers owing to different life experiences. So it's even harder to communicate with the tutor than with peers. We students have similar levels of English proficiency and similar views of certain social issues. Similar life experiences could also help us to understand each other's writing.

(Shen/posti/P23)

Thirdly, peers generated more dialogue than the tutor on grounds of the students' more equal social status. Xin explained this as follows:

I think it is good to discuss with peers because we can point out each other's problems and explain them in details. Further, the atmosphere between peers is more relaxing than that with the tutor, so we can discuss the problems more deeply. When I discussed with Art, I would not discuss my

writing so deeply. For example, I didn't discuss the organisation of the whole essay...I mean peers are in a more equal social status and thus the atmosphere between us was more relaxing.

(Xin/posti/P21)

A similar view was expressed by Zhang:

Peer feedback is discussed more interactively. It is not only feedback. It is feed forward and back between us. Therefore, we have a clear idea of which instance should be used and which will not be used.

(Zhang/posti/P27)

Last not least, peer assessment complemented teacher assessment due to its different focuses. Yao, for instance, assumed that teacher feedback should focus more on global issues whereas peers might comment more on surface areas:

Sometime teachers should focus more on organisation. But the peers, they should be more able to point out grammatical mistakes.

(Yao/pri/P36)

Shen also suggested the different focus of peer and teacher feedback:

Teachers provided deeper and more global feedback than our peers. For another thing, the tutor can comment on coherence between sentences but peers could not...but peer feedback is more helpful in ideas.

(Shen/posti/P19)

Although findings in the focus of peer and teacher feedback above suggested broadly similar focuses of peer and teacher feedback on the 79 assignments as a whole, feedback on individual assignments might receive different focused peer and teacher feedback. Further, students also commented, albeit less frequently, that: (1) peers could spot problems that student writers themselves could not observe and (2) peer assessment expanded readers from the writing tutor only. All these values about peer assessment have also been suggested by other studies looking at peer assessment (Hu, 2005; Mendonca & Johnson, 1994).

Based on students' perception data, Hypothesis 4 is confirmed.

Summaries and Discussions

Through the examinations of the four hypotheses, the following key findings were obtained in this study. The teacher provided significantly more feedback instances than the peer reviewers on the 76 assignments. Peer feedback played a complementary but not substitutable role to teacher

feedback in terms of the amount of peer feedback which was not addressed by the teacher and the amount of teacher feedback which was not made by the peer evaluators. The focus and the type of peer and teacher feedback were broadly similar: Wording and grammar as the two foci which were most frequently commented and direct feedback emerged as the dominant type. Teacher feedback exercised a greater effect on students' revisions than peer feedback, evidenced by its significantly more frequent use in students' revisions. Students valued peer assessment in terms of its distinguishable roles from teacher assessment.

The findings align with the findings that the students provided and used less peer feedback than teacher feedback. This could also be explained via learners' previous long-term teacher- and examination-oriented learning experiences of English writing. That is to say, the traditional dominant role of teachers in learners' learning experiences would made learners rely on teacher feedback over peer feedback.

On the other hand, the results go against the existing evidence of different focused peer and teacher feedback. The broadly similar focus and type of peer and teacher feedback contradicts findings which corroborate Leki's complaint about peer feedback provided by ESL learners:

They are likely to respond to surface concerns of grammar, mechanics, spelling, and vocabulary, taking refuge in the security of details of presentation rather than grappling with more difficult questions of meaning.
(as cited in Villamil & Guerrero, 2006, p. 493)

The similar focus could be explained by a teacher's modelling to students in Ferris' (2003):

both L1 and L2 researchers have suggested that peer feedback, for better or for worse, can be influenced by the priorities modelled by the teacher in giving feedback and in structuring the class in general and peer response session in particular. (p.131)

As far as this study is concerned, the broadly similar focus of peer and teacher feedback might be explained by the teacher's demonstration of how he provided feedback on student writing.

The current study confirms the existing evidence on the complementary role of peer to teacher feedback in terms of peer feedback on problematic areas which were not commented by the writing tutor and the unique roles of peer assessment. The findings should encourage teachers in the Chinese or similar teacher-dominated instruction contexts to embark on the use of

peer assessment despite learners' previous experience of peer assessment. Provided with appropriate training, learners even without prior experience of peer assessment are able to provide corrective feedback to assist peer learners' revisions of writing assignments.

Implications for future research

Although the sample size of this study is rather small, the current study generated relatively longitudinal evidence on the use of peer feedback across nine writing tasks of different genres and the peer evaluators; therefore, the findings on peer feedback in this study could be viewed as comprehensive by comparison with those studies based on one writing task. The influence of writing tasks and peers' individual factors (e.g. their English language proficiency and gender) was beyond the scope of this paper but worthy of further investigation in future work. In addition, it has been suggested in the empirical literature that students did not necessarily understand the feedback which was incorporated in their revisions (Goldstein, 2006; Hyland, 2003). Such feedback does not contribute to the development of learners' English writing proficiency. In this sense, the use of feedback should not be taken as the sole criterion for suggesting the effectiveness of feedback for writing. In future work, it is essential to investigate whether or not students understand the feedback they used in their revisions.

Conclusions

The findings in this study show that when peer and teacher feedback was implemented at the same stage of English writing, the latter appeared to be more influential than the former on students' revisions. However, peer feedback was found to complement teacher feedback by increasing feedback input, varying feedback resources, and creating learning resources for students. It is expected that the findings in this study will encourage writing tutors in the Chinese or other similar educational contexts to integrate peer assessment into their writing instruction. It is also expected that the current study could provide a baseline for further studies on peer assessment in the Chinese or other similar English education contexts.

References

- Ali Aljaafreh & James P. Lantolf (1994).** Negative feedback as regulation and second language learning in the zone of proximal development. In *The Modern Language Journal*, 78:4, pp465-483.
- E. Catherine Berg (1999).** The effects of trained peer response on ESL students' revision types and writing quality. In *Journal of Second Language Writing*, 8:3, pp215-241.

- Yuh-Show Cheng (2004).** A measure of second language writing anxiety: scale development and preliminary validation. In *Journal of Second Language Writing*, 13:4, pp313-335.
- Ulla Connor & Karen Asenavage (1994).** Peer response groups in ESL writing classes: how much impact on revision. In *Journal of Second Language Writing*, 3:3, pp256-276.
- Susan M. Conrad & Lynne M. Goldstein (1999).** ESL student revision after teacher-written comments: text, contexts, and individuals. In *Journal of Second Language Writing*, 8:2, pp147-179.
- Richard Donato (1994).** Collective scaffolding in second language learning. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research*. Ablex: Norwood, NJ, USA.
- Dana Ferris (2003).** Responding to writing. In B. Kroll (Ed.), *Exploring the dynamics of second language writing*. Cambridge University Press: Cambridge, UK.
- Dana Ferris (2006).** Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: contexts and issues*. Cambridge University Press: Cambridge, UK.
- Dana Ferris & John S. Hedgcock (1998).** *Teaching ESL composition: purpose, process and practice* (Vol. 9). Lawrence Erlbaum: Mahwah, N.J, USA.
- Linda Flower (1989).** Cognition, context and theory building. In *College Composition and Communication*, 40:3, pp282-311.
- Lynn Goldstein (2006).** Feedback and revision in second language writing: contextual, teacher, and student variables. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: contexts and issues*. Cambridge University Press: Cambridge, UK.
- John R. Hayes (1996).** A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The Science of writing*. Lawrence Erlbaum Associates: Mahwah, N.J, USA.
- John S. Hedgcock & Natalie Lefkowitz (1992).** Collaborative oral/aural revision in foreign language writing instruction. In *Journal of Second Language Writing*, 1:3, pp255-276.
- John S. Hedgcock & Natalie Lefkowitz (1994).** Feedback on feedback: assessing learner receptivity to teacher response in L2 composing. In *Journal of Second Language Writing*, 3:2, pp141-163.
- Guangwei Hu (2005).** Using peer review with Chinese ESL student writers. In *Language Teaching Research*, 9:3, pp321-342.
- Ken Hyland (2003).** *Second language writing*. Cambridge University Press: Cambridge, UK.
- Ken Hyland & Fiona Hyland (2006a).** *Feedback in second language writing: contexts and issues*. Cambridge University Press: Cambridge, UK.
- Ken Hyland & Fiona Hyland (2006b).** Feedback on second language students' writing. In *Language Teaching*, 39, pp83-101.
- George M. Jacobs, Andy Curtis, George Braine & Su-Yueh Huang (1998).** Feedback on student writing: taking the middle path. In *Journal of Second Language Writing*, 7:3, pp307-317.
- Charles Lockhart & Peggy Ng (1995).** Analysing talk in ESL peer response groups: stances, functions and content. In *Language Learning*, 45:4, pp605-655.

- Kate Mangelsdorf (1992).** Peer reviews in the ESL composition classroom: what do the students think? In *ELT Journal*, 46:3, pp274-284.
- Mary E. McGroarty & Wei Zhu (1997).** Triangulation in classroom research: a study of peer revision. In *Language Learning*, 47:1, pp1-43.
- Cássia O. Mendonça & Karen E. Johnson (1994).** Peer review negotiations: revision activities in ESL writing instruction. In *TESOL Quarterly*, 28:4, pp745-769.
- Hui-Tzu Min (2005).** Training students to become successful peer reviewers. In *System*, 33:2, pp293-308.
- Mrudula Patri (2002).** The influence of peer feedback on self- and peer-assessment of oral skills. In *Language Testing*, 19:2, pp109-133.
- Trena M. Paulus (1999).** The effect of peer and teacher feedback on student writing. In *Journal of Second Language Writing*, 8:3, pp265-289.
- Paul Rollinson (2005).** Using peer feedback in the ESL writing class. In *ELT Journal*, 59:1, pp23-30.
- Jane Stanley (1992).** Coaching student writers to be effective peer evaluators. In *Journal of Second Language Writing*, 1:3, pp217-233.
- Amy B. M. Tsui & Maria Ng (2000).** Do secondary L2 writers benefit from peer comments? In *Journal of Second Language Writing*, 9:2, pp147-170.
- Frank Tuzi (2004).** The impact of e-feedback on the revision of L2 writers. In *Computers and Composition*, 21:2, pp217-235.
- Olga S. Villamil & Maria C. M de Guerrero (1996).** Peer revision in the L2 classroom: social-cognitive activities, mediating strategies, and aspects of social behaviour. In *Journal of Second Language Writing*, 5:1, pp51-75.
- Olga S. Villamil & Maria C. M de Guerrero (2006).** Socio-cultural theory: a framework for understanding the socio-cognitive dimensions of peer feedback. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: contexts and issues*. Cambridge University Press: Cambridge, UK.
- Lev S. Vygotsky (1978).** *Mind in society: the development of higher psychological process*. Cambridge University Press: Cambridge, UK.
- Miao Yang, Richard Badger & Zhen Yu (2006).** A comparative study of peer and teacher feedback in a Chinese EFL writing class. In *Journal of Second Language Writing*, 15:3, pp179-200.

Appendix A: Student writing sample

Comments by Wayne: = not well organized.

Behaviors that separate us

-----my review of "Crash"

There are still many controversies about racial and ^{interpersonal communication} ~~intercourse~~ problems. Paul ^{is this what you mean?}

Haggis opens out the origin of the terrible interpersonal relationship by film language. ^{not a good word to use here}

In real life, people sometimes become impulse, ^{I can't understand this} ~~crankiness~~ ^{we cranky or} even vicious after they ^{the reason} ~~suffering~~ from frustration, pain and harm. People then often release such pressure ^{their} ~~accumulated in heart~~ ^{irrelevant. what does this word mean here.} ~~onto~~ ^{meaning?} ~~without~~ ^{this} considering other people's feeling.

Why ~~they~~ ^{who?} chose L.A as the setting of the story is because of its ~~specialty~~ ^{meaning?}.

Multi-ethnic culture ^{are} ~~is~~ "crashing" in this city, so it is much easier to bring ~~new~~ ^{meaning?} conflict. In order to foil the idealistic atmosphere of this film, the director deals with L.A as ~~the world hell.~~ ^{this means to prevent or stop (as a verb)}.

A group of disparate Angelenos of all races, classes and religions become enmeshed with each other through a series of ~~increasingly~~ ^{increasingly} contrived events in "Crash," a promising but ultimately disappointing drama about California angst.

I think many tragedies happened ~~because~~ ^{communication (everyday)} of racial prejudice and lacking of ~~intercourse~~ ^{intercourse}. Daily, we witness the negative effects of racial prejudice. It ~~creates~~ ^{brings about.} inequality, exclusions, and an atmosphere of rejection which prevent ^{need to find a better word. what do you mean} some groups of people from ~~being allowed into~~ mainstream American life. Prejudice is like a terrible cancer, engulfing the entire body, mind, and spirit, often defying the skills of those who wish to intervene. ^{a policeman in the movie}

Set an example, ~~the cop~~ ^{aa} wanted to get ~~a better medical treatment~~ ^{for his suffering} for his father, but no matter how begged, the black office clerk ~~charged~~ ^{in charge of} this case ~~didn't~~ ^{wouldn't} sign.

Notes for Data Set 3.2

- (1) Comments in blue were provided by the peer;
- (2) Comments in green were offered by the teacher;
- (3) The teacher also provided comments on peer feedback, as described in Figure 3.1. He crossed out the peer feedback with which he disagreed and underlined the peer feedback with which he agreed. Take the last sentence of Data Set 3.2 as an example. He crossed 'had' in the peer feedback by showing the student writer to change the sentence to 'no matter how he begged' rather than 'no matter how had he begged'. He underlined the peer feedback 'in charge of' to show that he agreed with the peer to change 'charged' to 'in charge of'.

Appendix B: Coding scheme of written feedback

Student comments in parentheses, tutor comments in square brackets.

Categories	Sub- categories	Definition	Examples
Feedback focus	Grammar	refers to singular/ plural form, adjective/adverb, conjunction, preposition, article, agreement, first person, transitive/intransitive verb, direct/indirect object, etc.	It was impossible because of his low marks.
	Wording	refers to the choice of words, phrases, or idioms	He waited and waited for a message that he had been admitted allowed to change majors.
	Mechanics	refers to spelling, punctuation, capitalisation, indenting, double space, typology, abbreviation and the format of particular genres such as letters or poems	No noise, no pollution; it is really a good place to preserve our health.
	Sentence structure	refers to grammatical arrangements of words in sentences, such as the word order, the relationship between words, word classes and other sentence elements	Most citizens have , on average, have -less housing area than villagers do.
	Style	refers to a particular word, sentence or passage which is not the most effective although grammatically correct for a particular writing task, such as register (e.g. formal or informal), tone (e.g. stress) and voice	<u>I guess</u> (don't guess. Offer an informed opinion.) one of the possible causes of the rise of the unemployment rate in the U.S. was the economic depression since 1989.
	Organisation	refers to the logic development of ideas and structures between sentences or paragraphs, such as paragraphing, the order of sentences, cohesion and coherence	
	Content	refers to what is written about, such as ideas expressed, evidence or examples used in writing	
Feedback type	Direct feedback	refers to the feedback specifying what need to be changed and how to change	Though it was not a wealthy family (I think it is simple.)
	Indirect feedback	refers to the feedback indicating the necessity of revision without providing actual revision solutions	When he got to the university, (?) everything was new
Tutor's agreement with peer feedback	Agreement	the tutor agrees with peer feedback	<u>God won't let these bad men pass.</u> (I don't understand.) [I can't either.]
	Partial agreement	the tutor agrees part with peer feedback	It turns out in most cases we miss it (even) without [even] (being) aware of that.
	No comment	the tutor provides no comment on peer feedback	
	Disagreement	the tutor disagrees with peer feedback	It united people like <u>nothing else.</u> (mean what?) [I think it is ok as it is.]
Use of feedback	Fully used	feedback for which revision solutions presented by direct feedback or revision requests stated by indirect feedback were incorporated in students' revised drafts	
	Partially used	direct feedback for which revision solutions were not incorporated in students' revised drafts	
	No evidence of use	no evidence shown that feedback was used in students' revisions (including deletion)	

19

On-line and Off-line EFL Sentence Processing of Reduced and Unreduced Relative Clauses: Evidence from Think-aloud Protocols and Translation Task

Natsumi Tanaka

Graduate School, University of Tsukuba /

Japan Society for the Promotion of Science

natsumi_tanaka313@yahoo.co.jp

Background

Several studies have argued that L2 learners cannot use syntactic cues effectively, while they rely on semantic or pragmatic cues (Shallow structure hypothesis; Clahsen & Felser, 2006). These findings have been replicated repeatedly using an on-line method comprising self-paced reading using non-cumulative, moving-window technique, eye-tracking paradigm, or event-related potentials (ERP) (Tanenhaus, 2004). Though many prior studies utilized these methods to investigate the readers' on-line processing, these on-line methods could not specifically disclose what the readers thought during reading. Other methods that focused on the readers' qualitative thoughts were effective only for off-line processes (e.g., free written recall task after reading). On the other hand, the think-aloud method could focus on readers' specific thoughts during reading.

Think-aloud is a research method that requires learners to report what they are thinking while performing a task (Olson et al., 1984). It focuses on the processes or strategies that participants use during tasks. One of the tasks is reading passages. While some criteria of categories were introduced for L1 or L2/EFL readings (e.g., Horiba, 2013), few studies used the think-aloud method to examine sentence comprehension as some limitations were found when applying this method for sentence processing (Olson et al., 1984). First, oral protocols did not always reflect readers' on-line thoughts, especially because lower-level processing such as sentence parsing was based on unconscious and rapid processing, unlike passage reading. Second, the think-aloud method is a dual task that requires more attentional resources during reading. Sometimes, this method affected processing or task performance negatively or positively.

Olson described the possibility of using the think-aloud method in sentence processing research as follows: syntactic and semantic analysis of sentences may or may not be usefully analyzed. The present study shows which aspect of sentence processing could not be revealed through the think-aloud method.

Purpose

The study adopts think-aloud methods to describe the learners' on-line sentence processing. To confirm the effectiveness of think-aloud methods, this study compares temporal ambiguous sentences (i.e., reduced relative clause sentences) that cause syntactic processing difficulties in unambiguous sentences (i.e., unreduced relative clause sentences).

This study addressed the following research questions (RQs):

- RQ1: Can the think-aloud method effectively investigate sentence comprehension in learners?
- RQ2: Is the performance of on-line think-aloud method consistent with the scores of an off-line translation task?

Method

Participants

A total of 10 Japanese undergraduates participated in this study. Two of them failed to complete all the tasks, so the analysis was conducted using the remaining data. The number of participants was quite low, as research using the think-aloud method does not require many participants.

Materials

Two types of experimental sentences were prepared: seven temporal ambiguous sentences with reduced relative clause and another seven unambiguous sentences with unreduced relative clause. These short sentences consisted of approximately 10 words. Experimental sentences with reduced relative clauses (e.g., “*The woman sent a doll was very excited*”) caused a garden path effect because the first verb sent can be interpreted as the main verb in initial parsing. On the other hand, the other sentences with unreduced relative clauses included no syntactic ambiguity (e.g., “*The cook who was criticized by the customers used too much salt*”).

Various 40 filler sentences were also added but were not treated as target sentences by the participants. The filler sentences included the objective sentences with reduced relative clause (RRC) and the sentences with unreduced relative clause (URC) and other types of ambiguous and

unambiguous sentences (e.g., ambiguous conjunction sentences without commas: “*Before the mother ate the candies had disappeared*” versus “*Before the mother ate, the candies had disappeared*”).

Pilot Study

Before the experiment, a pilot study was carried out to verify the think-aloud procedure and revise the categories of protocols. Since the think-aloud method is rarely applied to studies on sentence processing, the present study revised the categories of the part of Horiba (1996). Horiba originally set the following categories for lower processing: *graphophonemic / graphomorphemic analysis, word recognition, and syntactic/semantic analysis of clause or sentence*. The pilot study decides which category should be deleted, added, or changed.

Procedure of the pilot study

In the pilot study, the four participants who had been involved in the experiment on learners read the same sentences as in the experiment: They read the sentences and performed the think-aloud procedure in which they orally reported what they were thinking. Before they started the procedure, they completed a practice session to get used to it. In the practice session, the participants listened to a sample recording demonstrating the think-aloud technique using two filler sentences. Then they practiced the same two sentences and two new sentences using the think-aloud method. When they finished thinking aloud about a sentence, they pressed a key to read the following sentence at their own pace. After the participants understood the method, they proceeded to the think-aloud session. The sentences were presented randomly, with at least one filler sentence intervening every two target sentences. The participants were not allowed to remain silent for longer than five seconds. Instead, they were asked to read aloud the parts of sentences if they pronounced them in their minds or to explain what they found difficult to understand. When the participants expressed no words for approximately five seconds, the experimenter asked them to express their thinking during both practice and think-aloud sessions.

Results of the pilot study

Through the pilot study, the categories of graphophonemic / graphomorphemic analysis in Horiba (1996) that never appear in learners’ performance were deleted. In addition, the category of word recognition was integrated into semantic analysis because the participants often referred to the word meanings in Japanese after reading their English translation. The detailed categories and examples of the protocols are shown in Table 1.

Procedure

The participants read the sentences and performed the think-aloud procedure in which they orally reported what they were thinking. The procedure of thinking aloud was identical to that in the pilot study and was accompanied by the following translation task: The participants translated the sentences into Japanese (their native language) to comprehend their meaning. No feedback on the translation task was provided.

Scoring

The participants' protocols were transcribed and then parsed into statements, which were almost equivalent to the clauses based on Horiba (1996, 2013). These statements were classified into six categories (see Table 1): (a) semantic analysis, (b) syntactic analysis, (c) context information, (d) metalinguistic comments, (e) reading aloud, and (f) other information. The number of protocols in each category used by the participants showed their sentence-processing strategy.

Category	Definition	Examples of protocols
A: Semantic analysis	Meaning of words or phrase, word meaning inferences	" <u>warned</u> means he warned."
B: Syntactic analysis	Identifying the subject and verb, segmentation, and modification.	"Phrases till <u>last night</u> must be the subject of this sentence..." " <u>the man</u> is modified by <u>warned last night</u> ."
C: Context Information	Identifying the inferences beyond phrases.	"He regretted what he did, though he was warned."
D: Metalinguistic comments	Identifying the grammatical term (for example, RRC)	"The relative pronoun is reduced here."
E: Reading aloud	Reading aloud the whole sentence, et cetera.	" <u>The man warned last night regretted what he did</u> . [Reading aloud without any pause]"
F: Others	Comments of readers or self-monitoring	"I do not understand what this word implicates very well..."

Table 1: Categories and examples of think-aloud protocols in the experimental sentence: "The man warned last night regretted what he did."

Note. The words in underlined italics refer to the original English sentences that helped to interpret the examples of the protocols. The other parts of the protocols were originally in Japanese.

For the translation task, a binary score system was used. When the learners successfully demonstrated an appropriate syntactic representation, they were given one point per sentence, even when the translation included some trivial inconsistencies in meaning. Finally, the scores on the translation task were summed; the maximum possible score was seven.

Results and Discussion

The data obtained from the think-aloud and translation tasks were analyzed, using repeated measures t-test to compare the sentence constructions.

Think-aloud protocols

The examples of typical think-aloud protocols are demonstrated in Table 2, and Table 3 shows the results derived from the quantitative data obtained from the think-aloud task.

[RCC] <i>The woman / sent a doll / was very excited.</i>	
Translation:	[Protocol 1 by Participant A]
Correct	(b) <i>The woman sent a doll was...</i> / (b) Because verb appeared again in this sentence, <i>sent a doll</i> might modify <i>the woman</i> . / (b) <i>Was very excited</i> is a predicate of this sentence. / The woman who was given, sent, given a doll was very excited.
Translation:	[Protocol 2 by Participant B]
Incorrect	(e) <i>The woman sent a doll was very excited.</i> / (a) <i>The woman</i> means the woman. / (d) As for <i>sent a doll</i> , I think that such relative pronouns as “who is” before <i>sent</i> is reduced right here. / <i>Sent a doll</i> means that she did send a doll. / (a) <i>Was very excited</i> means she was very excited. / (b) Now, I get to know this sentence consists of SVC construction. / In the final analysis, the woman who did send a doll was very excited.
[URC] <i>The cook / who was criticized by the customers / used too much salt.</i>	
Correct	[Protocol 3 by Participant A] (a) <i>The cook</i> . The dish, ah, it’s not true, it means a chef. / (d) <i>Who</i> is a relative pronoun / (a) and <i>was criticized</i> means he was criticized. / (a) <i>by the customers</i> means customers. / (a) <i>used too much salt</i> . Eh, he used salt more than somebody. / The cook who was criticized by customer used salt more than somebody. [Protocol 4 by Participant C] (e) <i>The cook who was criticized by the customers used too much salt.</i> The cook who was criticized by the customers used too much salt.
Incorrect	N.A.

Table 2: Examples of typical think-aloud protocols

Note. The words in underlined italics refer to the original English sentences that helped to interpret the examples of the protocols. The other parts of the protocols were originally in Japanese. The sentences in bold are answers to the translation task. The code at the beginning of each sentence refers to the categories of the protocols.

When the participants noticed the temporal ambiguities in the RRC sentences, they attempted to solve the problems with their grammatical knowledge, regardless of their performance on the translation task. On the other hand, with the URC sentences, the participants usually accumulated the meanings of words or phrases and finally integrated those meanings to perform the translation task.

Note that participant C showed an irregularly different performance than that which the others showed: She did not report her thoughts in detail. For almost every sentence, she performed the translation immediately after reading the sentence aloud in English. Because she was the most proficient reader among the participants, it seemed that her high processing rate caused her difficulty in verbalizing her thoughts.

	A	B	D	E	F
RRC	18.93 (11.71)	38.05 (22.65)	1.90 (3.43)	39.58 (32.29)	1.53 (1.91)
URC	32.00 (0.19)	35.00 (19.40)	9.00 (10.40)	23.00 (31.30)	1.00 (2.40)

Table 3: The Means with standard deviations of frequencies of each category in data obtained from think-aloud task (%)

Note. Since the data was limited, category C was excluded from the analysis of think-aloud protocols.

The results of t-tests indicated two significant patterns in ambiguous and unambiguous sentences in Category A, $t(7) = -2.49$, $p = .047$, $d = 0.84$ and E, $t(7) = 2.90$, $p = .028$, $d = 0.52$. The other categories remained insignificant (Category B: $p = .758$, $d = 0.15$; Category D: $p = .158$, $d = 0.92$; and Category F: $p = .905$, $d = 0.25$). The results of think-aloud protocols showed different patterns in RRC and URC sentences. The protocols of category A (semantic analysis) showed higher frequency in the URC sentences than in RRC sentences. This suggests that the participants could concentrate on the semantic processing because the URC sentences did not include any syntactic ambiguities. On the other hand, the category of reading aloud (E) showed higher frequency in the RRC than in the URC sentence processing. Although the participants did not express their specific analysis, they might have decided on or engaged in total parsing while reading aloud the whole sentence.

Translation task

While the URC sentences were translated almost perfectly ($M = 7.00$, $SD = 2.96$), the scores on the RRC sentences were relatively low ($M = 4.89$, $SD = 2.37$). The t-test which had the same design as its preceding one, revealed that the participants failed to translate the RRC sentences as

accurately as they translated the URC sentences, $t(7) = -8.90$, $p = .023$, $d = 0.79$. It shows that the URC sentences were better interpreted than the RRC sentences.

Efficiency of think-aloud method for sentence comprehension (RQ1)

To answer the RQ1, the advantages and disadvantages of the think-aloud method are discussed as follows:

Regarding the advantages of the method, it explicitly and qualitatively described the causes of the learners' failure at syntactic analysis; it also indicated that the researcher should inference the learners' difficulties in reading-time or eye-tracking studies. In addition, the protocols showed the changes in the interpretation or the learners' use of their grammatical knowledge. As for the disadvantages of the think-aloud method, the individual differences in terms of the automatization of syntactic parsing might affect the performance (for example, participant C). To analyze certain types of performances such as reading aloud, researchers have to infer how the participants process sentences. Therefore, the think-aloud method is partially effective for investigating readers' specific processing or errors, but its analysis using other on-line or off-line data, or both, is essential to understand which situation is similar to that of other on-line methods.

Relation between think-aloud protocol and translation task (RQ2)

To examine the relation between the two tasks, two correlation analyses were conducted for each category of protocols, and the accuracy of the translation task involving the RRC sentences and the same variables in URC sentences was measured. The results indicated that only the frequency of Category A was highly correlated with the translation task involving the RRC sentences ($r = .73$, $p = .026$). In the URC sentences, no significant correlation was found. It suggested that the learners who applied their limited cognitive skills to comprehend ambiguous RRC sentences performed well on the translation task.

Conclusion

This study adopted the think-aloud method, which has frequently been used to reveal how EFL learners process sentences while reading. Although this method was considerably effective, the data on the protocols should be considered carefully, as the data were not always interpreted as a direct reflection of the learners' processing.

Some limitations were seen. First, the number of participants was too low to generalize the findings statistically. Second, the experimental sentences were not manipulated in terms of animacy; the subject and object of some sentences were animate (e.g., man and cat), but other sentences included inanimate nouns (e.g., boat and newspaper). Although this study utilized short simple sentences, further research on the think-aloud method might need more complex or longer sentences requiring higher-level processing.

Acknowledgement

The given data was obtained from the analysis of sentence processing conducted by the author using the think-aloud method (for other types of ambiguous conjunction sentences with reduced comma, see Tanaka, 2015). This research was supported by the Japan Society for the Promotion of Science (JSPS).

References

- Harald Clahsen & Claudia Felser (2006).** Grammatical Processing in Language Learners. In *Applied Psycholinguistics*, 27, pp3-42.
- Yukie Horiba (1996).** Comprehension processing in L2 reading: Language competence, textual coherence, and inferences. In *Studies in Second Language Acquisition*, 18, pp433-473.
- Yukie Horiba (2013).** Task-induced Strategic Processing in L2 Text Comprehension. In *Reading in a Foreign Language*, 25, pp98-125.
- Jung H. Lim & Kiel Christianson (2013).** Second Language Sentence Processing in Reading for Comprehension and Translation. In *Bilingualism: Language and Cognition*, 16, pp518-537.
- Gary M. Olson, Susan A. Duffy, & Robert L. Mack (1984).** Thinking-out-loud as a Method for Studying Real-time Comprehension Processes. In D. E. Kieras & M. A. Just (eds), *New Methods in Reading Comprehension Research*. Lawrence Erlbaum Associates, Hillsdale, NJ, USA, pp253-286.
- Natsumi Tanaka (2015).** Case Study of Japanese EFL Learners' Syntactic Parsing: Data from the think-aloud protocols. In *IRICE PLAZA*, 25, pp31-39.
- Michael K. Tanenhaus (2004).** On-line Sentence Processing: Past, present and future. In Manuel Carreiras & Charles Clifton Jr. (eds), *The On-line Study of Sentence Comprehension: Eyetracking, ERPs and beyond*. Psychology Press: New York, NY, USA, pp371-393.

Breaking Theory: New Directions in Applied Linguistics

*Proceedings of the 48th Annual Meeting
of the British Association for Applied Linguistics*

3-5 September 2015

Aston University, Birmingham

Edited by
Tilly Harrison, Ursula Lanvers & Martin Edwardes



ISBN: 978-0-9559533-8-5