

## Japan Association for Language Teaching

## Articles

- 119 Learning to be a Teacher: Development of EFL Teacher Trainee Interactional Practices  
— *Yuri Hosoda and David Aline*
- 149 Interest and Prior Knowledge in Second Language Reading Comprehension — *John Eidswick*
- 169 多肢選択式項目の項目形式が文法テストパフォーマンスに与える影響について [The Effects of Multiple-Choice Item Formats on Grammar Test Performance]  
— 熊澤孝昭 (*Kumazawa Takaaki*)

## Perspectives

- 189 Reconsidering the Effectiveness and Suitability of PPP and TBLT in the Japanese EFL Classroom  
— *Rintaro Sato*

## Reviews

- 201 *The Handbook of Language Teaching* (Michael H. Long and Catherine J. Doughty, Eds.)  
— Reviewed by Thomas Amundrud
- 204 *Researching Collocations in Another Language: Multiple Interpretations* (Andy Barfield and Henrik Gyllstad, Eds.) — Reviewed by Dale Brown
- 207 *Global English Teaching and Teacher Education: Praxis and Possibility* (Seran Dogancay-Aktuna and Joel Hardman, Eds.) — Reviewed by Justin Harris
- 210 *Vocabulary Matrix: Understanding, Learning, Teaching* (Michael McCarthy, Anne O'Keeffe, and Steve Walsh)  
— Reviewed by Natsuko Imaoka
- 213 *Teaching Second Language Listening* (Tony Lynch)  
— Reviewed by Joseph Siegel
- 216 *Second Language Learning and Identity: Cracking Metaphors in Ideological and Poetic Discourse in the Third Space* (Mika Yoshimoto) — Reviewed by Yuan Yuan

# JALT Journal

Volume 32 • No. 2

November 2010

## Editor

Darren Lingley  
*Kochi University*

## Consulting Editor

Ian Isemonger  
*Kumamoto University*

## Reviews Editor

Bill Perry  
*Kobe University*

## Japanese-Language Editor

Ken Urano  
*Hokkai-Gakuen University*

## Editorial Board

- |  |   |
|--|---|
| David Aline<br><i>Kanagawa University</i>                              | Tim Murphey<br><i>Kanda University of International Studies</i> |
| David Beglar<br><i>Temple University—Japan Campus</i>                  | Mary Goebel Noguchi<br><i>Kansai University</i>                 |
| William Bradley<br><i>Ryukoku University</i>                           | Aek Phakiti<br><i>University of Sydney, Australia</i>           |
| James Dean Brown<br><i>University of Hawai'i, Manoa, USA</i>           | Cynthia Quinn<br><i>Konan University</i>                        |
| Charles Browne<br><i>Aoyama Gakuin University</i>                      | Timothy J. Riney<br><i>International Christian University</i>   |
| Christine Pearson Casanave<br><i>Temple University—Japan Campus</i>    | Carol Rinnert<br><i>Hiroshima City University</i>               |
| Eton Churchill<br><i>Kanagawa University</i>                           | Gordon Robson<br><i>Showa Women's University</i>                |
| Steve Cornwell<br><i>Osaka Jogakuin College</i>                        | Hideki Sakai<br><i>Shinshu University</i>                       |
| Anne Ediger<br><i>Hunter College, City University of New York, USA</i> | David Shea<br><i>Keio University</i>                            |
| Rod Ellis<br><i>University of Auckland, New Zealand</i>                | Gregory Sholdt<br><i>Kobe University</i>                        |
| Peter Gobel<br><i>Kyoto Sangyo University</i>                          | Donna Tatsuki<br><i>Kobe City University of Foreign Studies</i> |
| Eli Hinkel<br><i>Seattle University, USA</i>                           | Deryn Verity<br><i>Osaka Jogakuin College</i>                   |
| Yuri Hosoda<br><i>Kanagawa University</i>                              | Yoshinori J. Watanabe<br><i>Sophia University</i>               |
| Nicholas O. Jungheim<br><i>Waseda University</i>                       | Sayoko Yamashita<br><i>Meikai University</i>                    |
| Yuriko Kite<br><i>Kansai University</i>                                | Jack Yohay<br><i>Seifu Gakuen, Osaka</i>                        |
| Setsuko Mori<br><i>Kinki University</i>                                |   |

**Additional Readers:** Yo In'nami, Hidetoshi Saito

**JALT Journal Production Editor:** Aleda Krause

**JALT Journal Proofreading:** Mayumi Fujioka, Seiji Fukazawa, Megumi Kawate-Mierzejewska, Aleda Krause, Greg Rouault, Joseph Sheehan, Jack Yohay

**JALT Publications Board Chair:** Steve Brown

**JALT Journal Layout:** Pukeko Graphics **Cover:** The Word Works

**JALT Journal on the Internet:** [www.jalt-publications.org/jj/](http://www.jalt-publications.org/jj/) **Website Editor:** Theron Muller

# Contents

November 2010  
Volume 32 • No. 2

- 115 In this Issue
- 116 From the Editor

## Articles

- 119 Learning to be a Teacher: Development of EFL Teacher Trainee Interactional Practices — *Yuri Hosoda and David Aline*
- 149 Interest and Prior Knowledge in Second Language Reading Comprehension — *John Eidswick*
- 169 多肢選択式項目の項目形式が文法テストパフォーマンスに与える影響について [The Effects of Multiple-Choice Item Formats on Grammar Test Performance] — 熊澤孝昭 (*Kumazawa Takaaki*)

## Perspectives

- 189 Reconsidering the Effectiveness and Suitability of PPP and TBLT in the Japanese EFL Classroom — *Rintaro Sato*

## Reviews

- 201 *The Handbook of Language Teaching* (Michael H. Long and Catherine J. Doughty, Eds.) — Reviewed by Thomas Amundrud
- 204 *Researching Collocations in Another Language: Multiple Interpretations* (Andy Barfield and Henrik Gyllstad, Eds.) — Reviewed by Dale Brown
- 207 *Global English Teaching and Teacher Education: Praxis and Possibility* (Seran Dogancay-Aktuna and Joel Hardman, Eds.) — Reviewed by Justin Harris
- 210 *Vocabulary Matrix: Understanding, Learning, Teaching* (Michael McCarthy, Anne O’Keeffe, and Steve Walsh) — Reviewed by Natsuko Imaoka
- 213 *Teaching Second Language Listening* (Tony Lynch) — Reviewed by Joseph Siegel
- 216 *Second Language Learning and Identity: Cracking Metaphors in Ideological and Poetic Discourse in the Third Space* (Mika Yoshimoto) — Reviewed by Yuan Yuan

## JALT Journal Information

- 219 Information for Contributors (English and Japanese)

All materials in this publication are copyright ©2010 by JALT and their respective authors.

# Japan Association for Language Teaching

## A Nonprofit Organization

The Japan Association for Language Teaching (JALT) is a nonprofit professional organization dedicated to the improvement of language teaching and learning in Japan. It provides a forum for the exchange of new ideas and techniques and a means of keeping informed about developments in the rapidly changing field of second and foreign language education. Established in 1976, JALT serves an international membership of approximately 3,000 language teachers. There are 35 JALT chapters, all in Japan, along with 20 special interest groups (SIGs) and three forming SIGs. JALT is one of the founders of PAC (Pan-Asian Consortium), which is an association of language teacher organizations in Pacific Asia. PAC holds regional conferences and exchanges information among its member organizations. JALT is the Japan affiliate of International TESOL (Teachers of English to Speakers of Other Languages) and is a branch of IATEFL (International Association of Teachers of English as a Foreign Language).

JALT publishes *JALT Journal*, a research journal; *The Language Teacher*, a bimonthly periodical containing articles, teaching activities, reviews, and announcements about professional concerns; and *JALT International Conference Proceedings*.

The JALT International Conference on Language Teaching and Learning and Educational Materials Exposition attracts some 2,000 participants annually and offers over 600 papers, workshops, colloquia, and poster sessions. Each JALT chapter holds local meetings and JALT's SIGs provide information and newsletters on specific areas of interest. JALT also sponsors special events such as workshops and conferences on specific themes, and awards annual grants for research projects related to language teaching and learning. Membership is open to those interested in language education and includes automatic assignment to the nearest chapter or the chapter you prefer to join, copies of JALT publications, and reduced admission to JALT-sponsored events. JALT members can join as many SIGs as they wish for an annual fee of ¥1,500 per SIG. For more information, contact the JALT Central Office or visit the JALT website at <[www.jalt.org](http://www.jalt.org)>.

### JALT National Officers, 2009

*President:* ..... *Caroline Lloyd*  
*Vice President:* ..... *Cynthia Keith*  
*Auditor:* ..... *Tadashi Ishida*  
*Director of Treasury:* ..... *Kevin Cleary*  
*Director of Records:* ..... *Aleda Krause*  
*Director of Programs:* ..... *Philip McCasland*  
*Director of Membership:* ..... *Nathan Furuya*  
*Director of Public Relations:* ..... *Emi Itoi*

### Chapters

Akita, East Shikoku, Fukui, Fukuoka, Gifu, Gunma, Hamamatsu, Himeji, Hiroshima, Hokkaido, Ibaraki, Iwate, Kagoshima, Kitakyushu, Kobe, Kyoto, Matsuyama, Miyazaki, Nagasaki, Nagoya, Nara, Niigata, Oita, Okayama, Okinawa, Omiya, Osaka, Sendai, Shinshu, Shizuoka, Tokyo, Toyohashi, West Tokyo, Yamagata, Yokohama

### Special Interest Groups

Bilingualism; Business English (forming); College and University Educators; Computer-Assisted Language Learning; Critical Thinking (forming); Extensive Reading; Framework and Language Portfolio; Gender Awareness in Language Education; Global Issues in Language Education; Japanese as a Second Language; Junior and Senior High School Teaching; Learner Development; Lifelong Language Learning; Materials Writers; Other Language Educators; Pragmatics; Professionalism, Administration, and Leadership in Education; Study Abroad; Task Based Learning (forming); Teacher Education; Teachers Helping Teachers; Teaching Children; Testing and Evaluation

### JALT Central Office

Urban Edge Building, 5F 1-37-9 Taito, Taito-ku, Tokyo 110-0016, Japan

Tel.: 03-3837-1630; Fax: 03-3837-1631; Email: [jco@jalt.org](mailto:jco@jalt.org);

Website: <[www.jalt.org](http://www.jalt.org)>

# In this Issue

## Articles

In this issue, we feature three research articles. We open with a contribution from **Yuri Hosoda** and **David Aline**. Their paper takes a microanalytical Conversation Analysis (CA) approach to study the classroom interactional practices of in-service teacher trainees in elementary school EFL classes. Key findings from their longitudinal study include evidence of teacher development in two main areas – the provision of assessments to students and increased initiative in providing directives. Their research highlights the importance of teacher training programs in shaping identities as professional teachers. In our second article, **John Eidswick** addresses a neglected area of research in second language learning by examining the influence of interest and prior knowledge on reading comprehension. Research findings suggest that while prior knowledge seems to play a more influential role than interest, interaction between these two factors on overall reading comprehension is complex. Our third paper, a Japanese-language contribution by **Takaaki Kumazawa**, reports findings of a study on the effects of multiple-choice item formats on grammar test performance to determine test reliability.

## Perspectives

In this section, **Rintaro Sato** considers two approaches to language teaching, presenting a case for the relative effectiveness of the traditional PPP approach, with some revisions, in light of the realities of English language education in Japan, which he contends pose a challenge to the adoption of Task-Based Language Teaching.

## Reviews

In the first of our six book reviews, **Thomas Amundrud** reports on a state-of-the-art volume of collected language teaching papers. In the second, **Dale Brown** reviews an edited volume of papers based on studies on L2 collocation. **Justin Harris** considers another edited volume which looks at the continuing development of world English and the impact of this on teachers. Our fourth review, by **Natsuko Imaoka**, deals with a book in the area of L2 vocabulary acquisition. The fifth of our reviews, by **Joseph Siegel**, examines a book on listening in a second language. Finally, **Yuan Yuan** reports on a book which considers the connection between language learning and identity.

## From the Editor

This issue of *JALT Journal* marks a time of considerable change in the Editorial. **Ian Isemonger**, who has served as Editor for 2 years, is stepping down. While working under Ian as Associate Editor these past couple of years, I have seen firsthand the time and energy he has invested in *JALT Journal*. My time with Ian has been a model learning experience, and transitioning to Editor has been all the smoother because of Ian's behind-the-scenes striving to improve the editorial process, most notably his work in piloting the newly created Journal Production Editor position. Any submitting author over the past 2 years, whether successful or unsuccessful in negotiating the process, will have benefited from Ian's insightful comments on research design, his keen attention to argumentation, and his precision and meticulousness in feedback. These are big shoes to fill. We are very fortunate that he will remain with *JALT Journal* in the newly created position of Consulting Editor. As the incoming Editor, it is a great relief to be able to formally rely on people like Ian for input on *Journal* matters, especially for counsel and perspective in times of difficulty; and, above all, for his specialization in quantitative research and the often neglected area of measurement. We plan to appoint a second Consulting Editor in the near future. This new position is especially important considering that this is the first time, at least since I can remember, that *JALT Journal* will not have an Associate Editor. Our open search for a new Associate Editor continues and details can be found on our website at <http://jalt-publications.org/positions/>.

We also bid farewell to **Yoshinori J. Watanabe** who has served admirably as Japanese-language Editor since 2005. This is a noteworthy commitment to *JALT Journal* and we will miss his editorial direction on Japanese manuscripts and excellent communication skills – thank you very much, Josh! Taking over as Japanese-language Editor is **Ken Urano**. Ken comes to us with an excellent background and a strong skill set for Editorial work. He has served *JALT Journal* as a reader over the past few years. We warmly welcome Ken to the team.

This issue of *JALT Journal* also sees changes in the Editorial Advisory Board. We are pleased to welcome **Hideki Sakai** to the Board. Hideki has been a reader for some time now and we look forward to his continued support at the *Journal*. We also see the departure from the Board of **Sandra Fotos** whom, on behalf of the *Journal*, I would especially like to thank for long and outstanding service to *JALT Journal*, most notably as a previous Editor. Also leaving the Board is **Thomas Robb** to whom we are grateful not

only for his long service on *JALT Journal's* Editorial Advisory Board but also as one of the founding members of the JALT organization. To the standing members of the EAB, I very much look forward to working with you and I thank you for your hard work in helping to bring this issue to publication.

In this issue, we also formalize the position of Journal Production Editor by appointing **Aleda Krause**. Aleda has been working informally in this role for a year now and we are pleased to officially welcome her. Those of you who have worked with Aleda in her various JALT capacities will already know of her many qualities. Her organization and communication skills, her keen eye, and her sound judgment make her perfectly suited to this job. She heads a proofreading and copy editing production team consisting of **Greg Rouault, Joseph Sheehan, and Jack Yohay**. This production team plays a critical role, a role filled with grueling work that all too often goes unseen. Thank you, and I very much look forward to working with you all.

Finally, we take this opportunity to remind our readers that, effective immediately, we are moving to the Sixth Edition of the *Publication Manual of the American Psychological Association*. Please have a close look at our submissions guidelines as well – these have been recently revised.

*Darren Lingley*



# Articles

## Learning to be a Teacher: Development of EFL Teacher Trainee Interactional Practices

Yuri Hosoda (細田由利)

*Kanagawa University*

David Aline

*Kanagawa University*

Applying Conversation Analysis (CA), this longitudinal study examines how teacher trainees developed their classroom interactional practices in an elementary school and constructed their identities as teachers over a 19-month period. The analysis focuses on two Japanese university students participating as teacher trainees in English Activities classes at an elementary school. Over 30 hours of classroom interaction were video recorded and analyzed in detail. Observation of the data revealed that through participation in this particular community, the trainees developed classroom interactional practices in two areas: provision of assessments to students and initiative in giving directions. As the trainees' interactional practices evolved, there were changes in their social actions in becoming teachers. This study demonstrates the positive effects of on-site training, and delineates the processes fledgling trainees undergo in transition to becoming fully participating teachers.

本研究では小学校の英語活動の授業において英語教育サポーターとして活動する大学生が19ヶ月間の間にいかにして教室相互行為の慣行を身につけ、教師としてのアイデンティティを構築していったかを会話分析の手法を用いて長期に渡って検証した。分析は横浜市の某小学校で英語活動に英語教育サポーターとして参加する大学生2名に焦点を当てた。検証の結果、19ヶ月間の間に児童の発話に対する口頭評価の仕方とおよび児童への口頭指示の仕方において2名の英語教育サポーターの成長がみられた。また、教室でのやりとりを学んでいくと共に、教師とし

てのアイデンティティも彼らの相互行為活動の中に垣間見られるようになった。本研究は現場での教育実習の有効性を示すと共に英語教育サポーターがその教育現場で徐々に駆け出しのサポーターから一人前の教師のようになっていく様子を描写した。

**A**pplying Conversation Analysis (CA) as an analytical resource, which requires microanalysis of interaction from a data-driven emic (participants') perspective, this longitudinal study examines how teacher trainees developed their classroom interactional practices in an elementary school and constructed their identities as teachers over a period of 19 months.

In response to repeated calls for employing CA to attempt to understand learning from a longitudinal perspective (see Hall, 2004; Kasper, 1997, 2004; Mori & Markee, 2009), some recent research applying CA has examined second language learners' change in participation over time through a microanalysis of institutional language learning interaction (see Hellermann, 2006, 2007, 2009; Young & Miller, 2004). Hellermann (2006) tracked two learners for a period of 30 weeks and demonstrated how their interactional practices developed from peripheral to more engaged participation. In a further study, analyzing data covering 18 to 27 months, Hellermann (2007) outlined the changes learners made in interactional practices as they incorporated, into their own production, classroom discursive practices for opening pair tasks. Likewise, Young and Miller (2004), observing writing tutorials, provided analysis showing how, over four writing conferences, a learner moved to fuller participation in an unfamiliar discursive practice.

Drawing on perspectives of situated learning (Lave & Wenger, 1991), the above studies conceptualized learning as resulting from learner participation in and adaptation to the social practices of a community. In this view, learners become competent participants as they move from peripheral to more engaged participation. The present study takes a further step in this line of research by using CA to examine how the interactional practices and identity orientations of teacher trainees change over time. Informed by research in the area of language socialization (e.g., Ochs, 1989, 1996; Ochs & Schieffelin, 1984; Schieffelin & Ochs, 1986), the present study is premised on the view that learning can beneficially be perceived as change in participation patterns. It also seeks to elucidate the transition of novice interactants from peripheral participation to full participation in a linguaculture and thereby ascertain how they evolve into expert members in a specific community.

From a CA perspective, this paper discusses changes over time in deployment of assessments and directives by two teacher trainees as manifested

by their identity orientations. In CA research, the identity orientations of participants in interaction are discussed only when such identities are “demonstrably relevant to the participants, and at that moment – at the moment that whatever we are trying to provide an account of occurs” (Schegloff, 1992, p. 109). Therefore, our a priori knowledge that one participant is an experienced worker and the other participant is a novice worker should not influence our analysis, which is founded upon stable characterizations of the participants throughout interaction. Whether such characterizations are relevant at any given moment in the interaction is determined by the participants themselves through their observable talk and other conduct. In that sense, CA studies take a radically emic perspective (Markee & Kasper, 2004). Another facet to consider when we examine identities is that someone who displays a certain set of characteristics can be treated as a member of the category or identity with which those characteristics are commonly associated (Antaki & Widdicombe, 1998). In institutional talk, “specific action choices can index particular institutional stances, ideologies and identities that are being enacted in the talk” (Heritage & Clayman, 2010, p. 18). Thus, in language classrooms, particular ways teachers and students talk is one aspect of how the institution of the classroom is realized or reproduced, and “talking in these ways is a part of being a teacher or a student .... These roles are enacted by talking in these ways” (Heritage & Clayman, p. 32).

In the present study, we will demonstrate through microanalysis of interaction how two teacher trainees demonstrably transform their identity orientations as teachers from “novice” to “expert” or “near-expert” by explicating their deployment of assessments and directives.

## Data

The analysis focused on two Japanese university students from among a group of approximately 14 participating as teacher trainees in English as a foreign language classes at an elementary school in Yokohama. The city government of Yokohama, going beyond the national curriculum, established guidelines which requested elementary schools to hold English Activities classes for all grades (Grade 1 to Grade 6) from the academic year 2007. As current Japanese elementary school teachers lack experience in teaching English (see Hogan, 2004; Kelly, 2002; Murphey, Asaoka, & Sekiguchi, 2004; Takagaki, 2003) and there are not enough native-speaking teachers available, some elementary schools have requested assistance from local universities.

Among the 14<sup>1</sup> university students who participated as teacher trainees at one elementary school in Yokohama city, two English majors, Shota (SHO) and Makoto (MAK) (pseudonyms, both male), assisted in English classes 2 to 3 days per month for 19 months. At the time of the study, the student population of the elementary school was over 600, with an average of 35 students per class.

Shota and Makoto worked with over 20 of the regular Japanese homeroom teachers (JHTs), data of six of whom are shown in the extracts presented here. In each class session, there was one Japanese homeroom teacher and in most of the classes also a non-Japanese expert speaker of English or assistant language teacher (ALT) present. (The ALT was not present in Shota's Time 2 class presented here.) During each visit to the school, the teacher trainees assisted in two or three English Activities classes.

Over 30 hours of classroom interaction were video recorded and analyzed in detail. Subsequently, the classroom interaction of the six classes presented here was transcribed using the transcription system developed by Jefferson (2004), the one most commonly used in Conversation Analysis. In this longitudinal study, three time slots each from 40 classes taught by Shota and 70 classes taught by Makoto were selected to capture the 19-month time frame of the project. Then classes were randomly selected from the pool of recorded classes within each of those three time slots. The exact dates of each time slot are shown in Table 1.

**Table 1: Dates of Visits**

Teacher trainee	Time slot	Date	Number of visit	Grade level
Shota	1	Sept. 20, 2007	1st visit	5th grade
Shota	2	Sept. 25, 2008	10th visit	1st grade
Shota	3	Jan. 29, 2009	15th visit	3rd grade
Makoto	1	Sept. 25, 2007	1st visit	3rd grade
Makoto	2	July 2, 2008	15th visit	4th grade
Makoto	3	Feb. 16, 2009	30th visit	2nd grade

## Method

To analyze the data, we mainly employed the framework of CA but also incorporated aspects of sociocultural theory and language socialization.

Conversation Analysis developed in part as a reaction to the a priori theorization and ad hoc analysis of mainstream American sociologists in the 1960s. A fundamental principle of CA methodology maintains that it makes no a priori assumptions about the data. Approaching the data with preconceived notions or coding schemes, no matter how incisive the researcher may be, constrains observation of the actual phenomena. According to Pomerantz and Fehr (1997), "it (CA) rejects the use of investigator-stipulated theoretical and conceptual definitions of research questions" (p. 66). Consequently, research questions arise from the data itself and its particularities.

As a counter to the deficiencies of a priori theorization, CA methodology employs as a basis for data analysis unmotivated exploration: the systematic inspection of audio and video recordings along with finely detailed transcripts of uncontrived interaction, focusing on revealing participants' micropractices through their orientations to sense making and shared understanding of their own interaction as embodied in the minute details of their talk and other conduct.

Initially, the data in this study were transcribed, then a number of general observations were framed about the interaction through analysis of both the transcribed and recorded data, following which an interactional phenomenon of interest was identified through reanalysis of the data.

However, as He (2004) noted in her application of CA to research on L2 learning and teaching, CA does not directly address learning and CA's exclusive focus on observable behavior at the moment of social interaction may pose difficulties for documenting change of behavior over an extended period of time. Therefore, for the purposes of this study, we adopted some concepts from a sociocultural theory (SCT) of learning. Like CA, SCT views knowledge and cognition as being socially shared and distributed among interactants (e.g., Donato, 2004; Lantolf & Appel, 1994; van Lier, 1998). Unlike CA, SCT directly tackles the issue of learning: Based on the assertions of Activity Theory, SCT proponents argue that social interaction leads to language acquisition as learners internalize language through the creative construction of language.

Another approach to the issue of learning in situ is to incorporate the ideas of language socialization. For example, He (2004) integrates language socialization (e.g., Ochs, 1989; 1996; Ochs & Schieffelin, 1984; Schieffelin & Ochs, 1986) into her CA-inspired study of Chinese language classrooms. Although language socialization emphasizes the analysis of interaction as

much as CA does, the basic concern of language socialization is human development and growth. According to Schieffelin and Ochs (1986), language socialization works bidirectionally as language functions as a medium or tool in the socialization process: “socialization to use the language” and “socialization through the use of language” (p. 167). In the socialization process, novices acquire the knowledge, orientations, and practices that enable them to participate effectively and appropriately in a particular community. Acquisition is realized mostly through the use of language, the major symbolic medium for communicating and negotiating cultural knowledge (Garrett & Baquedano-López, 2002).

The current CA study also incorporates conceptualization of language socialization and documents the teacher trainees’ socialization process through the use of language in a particular activity. Specifically, the study demonstrates how teacher trainees develop their identities as teachers through the deployment of directives and assessments in elementary school English Activity classrooms over an extended period of time.

## Analysis and Discussion

Emerging from the analytical practice of unmotivated looking, analysis revealed development of the trainees’ classroom interactional practices over time in two areas: (a) provision of assessments to students, and (b) initiative in giving directives. The trainees’ interactional practices and identities as teachers are shown to be transformed through actual participation in the classroom community. The identity orientations of the trainees were manifested in the changes in their classroom interactional practices as they functioned with the homeroom teachers, the assistant language teacher (the same ALT throughout this data set), and students. The transition from peripheral participation as visitors to more engaged participation as near-expert teachers was publicly demonstrated in the classroom interactions.

## Assessments

Assessment deployment was one aspect of the interaction that came to the fore during analysis of the video and transcripts. In spontaneous, naturally occurring first language interaction, an activity common to speakers and recipients is some sort of evaluation of referents: the people, objects, and events being discussed (Goodwin & Goodwin, 1987). These evaluations, commonly termed *assessments*, can be analysed as structural units such as adjectives; suprasegmental phenomena such as intonation; and speakers’

acts, or assessment actions, in which the emphasis is on an entire speech act. An example of assessment in ordinary conversation is provided in Extract 1 from Pomerantz (1984, p. 57). In this extract, following J's suggestion to test the water, R makes a series of assessments.

(1) (VIYMC 1:4)

J: Let's feel the water. Oh, it...

R: It's wonderful. It's just right. It's like bathtub water

According to Pomerantz (1984), "assessments are produced as products of participation; with an assessment, a speaker claims knowledge of that which s/he is assessing" (p. 57). Therefore, the teacher trainees' assessment of the student performance purportedly displays their full participation in the classroom activity at hand as well as their sufficient knowledge of what the activity is for and about. Moreover, as Heritage and Raymond (2005) and Raymond and Heritage (2006) demonstrated, production of assessment reflects epistemic rights to evaluate the third party (e.g., a grandmother of the person being assessed has greater rights to assess than her friend). The speaker can declare greater rights to assess by producing assessments first or upgrading the assessments at times when the other interactant has produced assessment first. As demonstrated below, at later stages of this study, the teacher trainees began to claim their rights to assess student performance by occasionally proffering assessments first, or by upgrading their follow-up assessments when the other teacher(s) produced assessments first. In the institutional interaction of the classroom, assessments have traditionally been seen as being proffered by the teacher in the feedback turn of Initiation-Response-Feedback (IRF) sequences (Sinclair & Coulthard, 1975). Therefore, providing assessments to student responses shapes one's identity as a teacher.

In the transcripts, SHO refers to Shota, MAK refers to Makoto, ALT refers to the near-native assistant language teacher, JHT refers to one of six Japanese homeroom teachers (but only one in each class), EV refers to everybody in chorus, and S refers to a student.

### **Shota**

Shota displayed significant change in the way he used assessments in this data set. At Time 1, he produced no verbal assessments at all during any in-

teraction in the classroom and only aligned with the assessments of the ALT and JHT through applause following their initiation, or displayed embodied positive assessments through head nods.

In Extract 2, Shota models a dialog with a student at the front of the class. However, even though Shota is interacting with the student, it is the JHT and ALT who implement the assessments.

(2) [Shota, Sept. 20, 2007, 5:00-17]

01 SHO: how are you.

02 (.)

03 S1 : I'm fine thank you >and you?<

04 SHO: I'm fine, thank you.

05 (0.2) ((SHO head nod))

06 JHT: very goo:d

07 ALT: [( ?) ((ALT does high-five with S1))

08 EV : (([applau[se]))

09 Ss : [oh, yah yea

10 ALT: [(goo::d okay )

((ALT says 'okay' to JHT))

Prior to this extract, the ALT requests Shota to say “how are you” to one of the students. Upon the ALT’s request, Shota asks the question to S1 (line 1). By asking Shota to perform some of the classroom teaching tasks, the ALT treats Shota as an assistant, and by implementing the ALT’s request, he himself shows his orientation to being a teacher trainee through his assistance. As Shota completes the dialog in line 4, he begins to move out of the front of the classroom to the side, giving a slight head nod to the student. At this point, the JHT provides an assessment with “very goo:d” and the ALT initiates and co-completes an embodied assessment with the student. Shota then joins in with the applause initiated by the JHT; clearly he is not orienting to the same type of teacher identity as the JHT and ALT as the feedback turn of the traditional IRF sequence is noticeably absent from his interaction. Shota’s nod in line 5 shows that he treats the student’s response as appropriate. However, he does not verbalize the assessment. By refraining from giving assessment verbally, it is possible that he does not view himself as a person who possesses the right to judge the student’s performance. On

the other hand, the JHT verbalizes the positive assessment, publicly demonstrating that she is in a position to judge the student performance.

Later in the program, at Time 2, however, Shota does provide assessments: when working with students on a one-to-one basis; when following up on the JHT's assessments by translating them into English; and when put in charge of leading the class, as shown in Extract 3.

- (3) [Shota, Sept. 25 2008, 2:12-16]  
 ((SU refers to another trainee))  
 01 ((Shota flashes a number card))  
 02 Ss : *si::[:x* ((individually, not in chorus))  
 03 SU : [ what's *thi*s.) ((points to card))  
 04 SHO: *six* very good.  
 05 (.)  
 06 SHO: *so:::*  
 07 ?? : ( )  
 08 SHO: *oka::y*

In Extract 3, after Shota flashes a number card and the students call out the number, he receipts (receives and acknowledges) their response through repetition (a process identified in Greer, Bussinger, & Mischlinger, 2009), and then deploys an assessment, "very good." Although both Shota and SU were put in charge of the teaching activity, it was Shota who produced the assessment. This shows that Shota orients to the fact that it is he, not the other trainee, who has a right to judge students' performance, at least when the trainees are in charge of teaching. This is an interesting contrast with Time 1 when he did not produce any assessments even when he was in charge of some classroom teaching tasks. However, in Time 2, while Shota often gave assessments during individual work with students, he only provided assessments in teacher-fronted activities when he was clearly assigned to lead the class by the ALT or JHT. Even so, at this stage these were still rarely heard.

By Time 3, Shota displays greater variety in assessments when working individually with students and when following up on the JHT's and ALT's assessments. His follow-up assessments are always upgrades, just as second assessments are in first-language conversations (Pomerantz, 1984). Moreover, he now displays greater participation and identity as one of the group of teachers rather than as a visitor or assistant in that he deploys assessments

even when the JHT and ALT are clearly in charge of the main classroom activity (Extract 4).

(4) [Shota, Jan. 29, 2009, 2:55-3:00] ((Ss read off numbers from flash cards))

01 ALT: ((turns over another flash card))

02 Ss : eleve::[::::~n ((multiple scattered and overlapping responses))

03 ALT: [ele:ve:[n

04 SHO: [↑wo:w, (.) >very g[ood<]

05 ALT: [oka:~y,

06 (.)

07 ALT: ele:ven

08 Ss & SHO: ele:~ve:n

Here, Shota is kneeling off to the side as the JHT and ALT manage the main activity of turning over flash cards of numbers. As the students call out the number, the ALT repeats it through repetition, line 3. Shota then produces “↑wo:w,” in line 4, which appears to be what Goffman (1981) termed a *response cry*, a brief emotional expression, before accounting for the surprise with a compressed assessment,<sup>2</sup> which is overlapped by the ALT’s basic assessment token “oka:~y,” line 5. By producing an assessment of student performance even when the JHT and ALT are standing in front of the classroom managing the main activity, Shota displays his orientation to being in a team of teachers along with the JHT and ALT. At this point in the program, Shota repeatedly produced assessments in a greater variety of sequential environments whether or not he was in charge of the main activity; his follow-up assessments were always upgrades, and almost all of his assessments were rushed as he attempted to squeeze them into the fast-moving interaction of the classroom.

Through examination of Shota’s development in assessments, a clear change in his identity orientation was observed. In Time 1, by not giving assessment even when he was in charge of a teaching activity, he treated himself as someone who did not possess the right to judge student performance, but was rather just a teaching assistant who was trying to accomplish an assigned task. In Time 2, by giving some assessments when the ALT or JHT was not the main teacher of the activity but Shota was, he oriented to himself as someone who possesses the right to make judgments of student

performance. And in Time 3, by deploying assessments even when the ALT or JHT was the main teacher of the activity, he treated himself as one of the teachers in the classroom, in possession of the right to give feedback on student performance.

### **Makoto**

In Time 1, the assessments Makoto produced were limited to clapping as one form of positive assessment (for more on forms of assessment, see Hosoda & Aline, 2010a, 2010b), and simple assessment tokens such as “good job” and “very good,” which were exact copies of assessments produced by the ALT and JHT.

In Extract (5), Makoto has been assigned by the ALT to ask a student a question, and he chooses S13 to answer the question.

- (5) [Makoto, Sep. 25, 2007, 20:07-12]
- 01 MAK: ((walks toward Ss)) okay.  
 02 <what fruit do you like.>  
 03 S13: I: li:ke orange.=  
 04 JHT: =orange oka:y ((claps hands))  
 05 ALT: a:: very goo[d.((claps hands))  
 06 MAK: [very good. ((claps hands))

Although in line 2 it is Makoto who directs the question to S13, in line 3 the JHT is the first to accept S13’s answer. The ALT’s assessment follows the JHT’s assessment and then in line 6, Makoto’s assessment is exactly the same form as the ALT’s assessment, “very good.” and clapping.

In this extract, by assigning Makoto to do some of the teaching, the ALT treats Makoto as her teaching assistant, while by complying with the assignment and carrying out the task, Makoto orients to being an assistant for the class. The JHT’s deployment of assessment that immediately follows student performance in line 3 demonstrates her orientation to being in charge of the class even when somebody else is leading the class. The ALT’s assessment also displays her orientation to being a teacher by producing the assessment. As Makoto is the one who is leading the class and who asked the question in line 1, he is in a position to produce the assessment in the turn following S13’s response. However, Makoto provides assessment only after

assessments by the JHT and ALT, and it is in the same format as the ALT's assessment. Providing assessment in this format and position may show his uncertainty about having the right to produce assessments and how assessments should be produced. However, his deployment of the assessment demonstrates his orientation to being in a team of teachers.

In Time 2, Makoto started producing a greater variety of assessments, such as "excellent," "perfect," "good," and "good job." These assessments were not simply copies of those produced by the ALT or JHT, and his assessments occasionally occurred sequentially earlier than the ALT's or JHT's assessments. (Extract 6)

(6) [Makoto, July 2, 2008, 09:01-10]

01 ALT: u:o::kay. what's this?

02 EV : head.

03 ALT: ((touches her shoulders))

04 EV : shoulder

05 ALT: ((touches her knees))

06 EV : knees

07 ALT: ((touches her toes))

08 EV : toes

09 MAK: excellen[t

10 ALT: [very g- okay nex<sup>o</sup>t<sup>o</sup>

In Extract 6, in lines 1 to 8, the ALT asks the whole class what each part of the body is called in English and students answer the question in chorus. However, in line 9, when the activity is finished, it is not the ALT but rather Makoto who first produces an assessment token, "excellent". In line 10, the ALT starts to produce an assessment, "very g-" but cuts it off and produces "okay nex<sup>o</sup>t<sup>o</sup>", actions which indicate a shift of activity. As it occurs after Makoto's assessment, the ALT's assessment is considered a follow-up assessment. Pomerantz (1984) notes that assessments following somebody else's assessments are usually designed to be upgrades. Therefore, in this position, the ALT's assessment should be an upgrade. One reason the ALT cuts off her own assessment may be that her assessment, "very good," is not an upgrade of Makoto's assessment, "excellent". Through the way that Makoto produced his assessments, he demonstrated his orientation to being one of

the teachers even when he was not the main teacher of an activity. Moreover, the ALT's cutoff of her assessment and shift of the activity demonstrates her acceptance of Makoto's assessment and therefore her acceptance of his role as a teacher.

Makoto continued to use a variety of assessment tokens in Time 3. He often produced assessment tokens by himself, without the JHT or ALT having produced one, or produced assessments before the ALT or JHT produced theirs. Moreover, when the ALT's or JHT's assessments occurred before his assessments, Makoto upgraded his assessments, as shown in Extract 7.

(7) [Makoto, Feb. 16, 21:57-22:14]

((JHT has asked students how many marbles are in the envelope and students are making a guess.))

01 ALT: a:: hidaka:ç

02 S1 : ((stands up)) five?

03 ALT: five ah >okay okay okay<  
((stops S1 from sitting down))

04 JHT: up.

05 S? : *wakannai.*  
"I don't know"

06 S1 : °five°

07 ALT: °°six°°=

08 S1 : six.=

09 ALT: =six?

10 (0.2) ((JHT looks inside the envelope))

11 JHT: ye:s. [that's right.

12 ALT: [u[OH:::::::::: ((claps hands))

13 MAK: [OH:::::::::: ve:ry good. excellent.  
((claps hands))

In Extract 7, the JHT is standing at the front of the classroom holding an envelope that contains marbles and asks students to guess how many marbles there are in the envelope. The ALT is also standing at the front of the classroom assisting the JHT. Makoto is standing behind the students. In line 8, S1 makes a second guess and utters "six." In response to this answer,

the JHT checks the inside of the envelope and says, “ye:s.” Hearing the JHT’s “ye:s.,” both the ALT and Makoto produce prolonged *ohs* while the JHT is producing an assessment token, “that’s right.” After producing “oh:.....:,” Makoto adds assessment tokens, “very good.” and “excellent.” which are upgraded positive assessments of the JHT’s version of assessment “that’s right.” Again, his assessment occurs when other teachers are leading the class, showing his orientation to being one of the teachers. Moreover, his upgraded follow-up assessment shows his full rights to provide and confidence in deploying the assessment (Heritage & Raymond, 2005; Raymond & Heritage, 2006) as well as his mastery of how to produce the assessment.

During the time period examined, the way Makoto produced assessments changed from doing assessments peripherally, following the JHT or ALT, to doing assessments fully by himself on his own initiative. Providing assessments by himself or upgrading the JHT’s or ALT’s assessments displayed his identity as a fully participating teacher.

Both Shota and Makoto moved from only infrequent simple assessments following on the ALT’s or JHT’s initiation in Time 1 to repeated and varied assessments in Time 3. In Time 3, both trainees produced assessments sequentially earlier than the ALT or JHT even when they were not the main teacher of the teacher-fronted activities. This change publicly demonstrates their transition from being guests or assistants in the classroom to being teachers.

## Directives

Another aspect of change we noticed over the three time periods was a change in the way Shota and Makoto provided directives in the classes. Traditionally, directives have been studied from the speech act perspective and have been examined as a single utterance or as a pair of utterances (directive-compliance/noncompliance). Directives have been seen as utterances produced to get someone else to do something (Austin, 1962; Searle, 1976), and include offers, requests, orders, prohibitions, and other verbal acts that ask for goods or attempt to bring about changes in activities of others (Ervin-Tripp, 1976; Ervin-Tripp, O’Connor, & Rosenberg, 1984). Recently, however, directives have been investigated from a CA perspective,<sup>3</sup> demonstrating that directives and their actions are better understood through examination of their trajectories constituted in extended sequential contexts (Cekaite, 2010; Goodwin, 1990, 2006; Goodwin & Goodwin, 2001).

Ervin-Tripp et al. (1984) noted that “a speaker who is high in esteem has the right to receive verbal deference from others and can make control

moves boldly, without offering deference to those who are lower in esteem” (p. 118). Providing directives to others through ordering and summoning is considered to be universally linked to a high ranking status (Ochs, 1996). Thus, the action of uttering directives and complying (or not complying) to directives can be taken as “category-bound activities” (Sacks, 1972) that invoke pairs of categories that consist of a person who is in a higher rank and a person who is in a lower rank, such as parent-child, boss-employee, doctor-patient, or teacher-student. In a classroom context, by issuing directives, the speaker orients to herself or himself as a teacher while constituting the receivers of the directives as students. As discussed below, this kind of orientation by Makoto and Shota was initially absent but manifested itself at later stages.

### **Shota**

During Time 1, Shota displayed an orientation to himself as an assistant: repeatedly shushing the students when the ALT looked to him for assistance in getting the students to become quiet, displaying bodily agreement with the ALT’s directives through head nods and hand gestures, or helping to form groups of students for small group activities. His only verbalized directives were produced outside of the teacher-fronted arena when he moved into the crowd of students and formed groups or directed them to stand up, as in Extract 8.

(8) [Shota, Sept. 20, 2007, 3:22-52]

01 ALT: so, everybody stand up. (.) stand up  
 ((stamps her foot, palms up lifting gesture))  
 02 Ss : (0.8) ((class noise. some Ss stand, some don’t))  
 03 JHT: stand u:pu,  
 04 (1.0)  
 05 S1 : ( )  
 06 (5.0) ((class noise, some Ss not standing))  
 07 ALT: everybody stand up. ( )  
 08 SHO: stand up, stand up. (.) everybody stand up.  
 09 Ss : ((seated Ss finally stand))



again, twice, this time with the politeness marker added. Therefore, at Time 2, Shota is still orienting to his identity as an assistant in that he repeats the directives of the main teacher, but he is taking a stronger stance with directives in that his repeat overlaps with the JHT's continued directive, and Shota is additionally orienting to his identity as an expert speaker of English by expanding on another's directives with marked politeness. The JHT too orients to this expertise by incorporating the marker into his own directive. The format of Shota's directives, which is slightly different from the JHT's directives, demonstrates his better understanding of how to formulate directives in the class as compared to Time 1. By producing directives when he is not leading the class he demonstrates his orientation to being one of the teachers in the room.

A significant change in Shota's orientation to identity manifested itself by Time 3. Shota still assists through repeats of directives with slight expansions or paraphrases (back to your seats / sit down please), and deploys directives when put in charge of leading the class. But more than this, he begins giving directives of his own. This is an incipient orientation to his identity as a full-fledged member of the teaching community, as evidenced in Extract 10.

(10) [Shota, Jan. 29, 2009, 4:55-5:08]

01 ALT: OKAY L OOK.

02 (.)

03 SHO: look, look okay?

04 (0.2)

05 SHO: uh:: (.) [.hhh

06 S1 : [ura:[: ((student noise))

07 SHO: [l et's [let's tr-

08 ALT: [OKAY everybody toge[ther  
oka:y̥

09 SHO: [°kay°

10 SHO: ((pulls up card with number 1))

11 EV : o::ne,

Here, the ALT along with the JHT, Shota, and another teacher-trainee are finishing explaining the rules for a small group game by showing for the

last time how to pick up and read off the number cards. The ALT marks the change with a very loud call for attention in line 1, following which Shota repeats the directive twice. Shota has some difficulty in proceeding as indicated by the perturbation “uh:” and micropause in line 5. After the pause and an in-breath, which shows he is preparing to produce a chunk of utterance, Shota initiates a directive, but it is not a repetition of another teacher’s directive. In line 8, Shota says “let’s”, and is at a position of greatest grammatical control, a place of least turn transition relevance in that a verb is expected next. He then recycles “let’s” and in overlap with his recycle the ALT produces a loud “OKAY” and stress on the first two syllables of “everybody”. In consequence, Shota relinquishes his turn and aligns with the ALT as shown by his production of “kay” in quiet speech. This extract is one example of Shota beginning to deploy his own directives, not just repeat the directives of other teachers. Furthermore, although Shota orients to an identity as a teacher by initiating directives, he also displays an orientation to the ALT’s expertness by dropping out of overlap in this crucial position of his own directives, crucial in two respects: in terms of the utterance level as he is at the point of maximal control of his utterance, and in terms of sequential position as it is at the point of initiating an activity.

### **Makoto**

In Time 1, Makoto did not produce any directives to start tasks or facilitate students’ performance in a task. He quietly listened to the ALT or JHT give directions to students. In Extract 11, the ALT gives directions to initiate a singing activity.

(11) [Makoto, Sep. 25, 2007, 5:01-12]

- 01 ALT: okay e:verybody. (0.3) okay let’s repeat with us  
 02 CD one time. okay. can you repeat after us.  
 03 Isabella sensei, Nagasawa sensei, and Makoto  
 04 sensei.[ hello:.  
 05 MAK: [(standing still))  
 06 Ss : ♪hello:..♪ ((waves hand))  
 07 ALT: uhhah:?(points to one student) ♪he[lllo:..♪  
 08 MAK: [ehhehheh  
 ((standing still))

09 Ss : ♪he[ll̩o:..♪ ((waves hand))  
 10 MAK: [ ((waves hand))  
 11 ALT: ♪hello:..♪  
 12 Ss : ♪hello:..♪  
 13 (0.5)  
 14 ALT: ♪hell̩[̩o:..♪ ((waves both hands))  
 15 MAK: [uhhuh [hhuh hhuh hhuh ((waves both hands))  
 16 Ss : [ ♪hello:..♪  
 17 ALT: ♪how're you:..♪ ((moves both hands))  
 18 Ss & MAK: ♪how're you:..♪ ((moves both hands))

In lines 1 to 4, the ALT produces a directive to students to start singing. Specifically, she directs students to repeat after the JHT, Makoto, and her. In this way, the ALT includes Makoto as one of the teachers in her directive in line 3. Although the ALT tells students to repeat after the JHT, Makoto, and her, Makoto does not perform the first pair part of the repetition sequence. In a repetition sequence, some utterance comes as a first pair part, and repetition of the first pair part should come as a second pair part. However, Makoto remains quiet up to line 15. Moreover, he does the second pair part nonverbally with the students, line 10 (overlapped with students' turn in line 9), reflecting his disorientation to being a teacher. Makoto's hand-waving in line 15 occurs just between the teacher's first pair part and students' second pair part, possibly indicating his uncertainty about whether he should act as a teacher or a student. Furthermore, in line 18, he produces the second pair part with the students as if he were one of them rather than one of the teachers. This demonstrates that he does not at this time orient to an identity as one of the teachers in the classroom.

In Time 2, Makoto frequently deployed a variety of directives both verbally and nonverbally. On some occasions, he took the initiative in providing directives, especially during singing activities. In those activities, Makoto collaboratively produced directives with the ALT and JHT by (a) repeating the directives produced by the ALT or JHT but in a louder voice, (b) slightly modifying what the ALT or JHT said, or (c) deploying directives in original forms. In Extract 12, the students are singing the "Seven Steps" song.

(12) [Makoto, July 2, 2008, 1:22-2:02]

01 ALT: ready:??

- 02 (1.0)
- 03 ALT: please si::ng.
- 04 MAK: ♪TAN TAN TAN, ♪ ((singing gestures with hands))
- 05 EV : ♪one two three four five six seven. one two
- 06 three four five six seven, one two three, one
- 07 two three, one two three four five six seven.
- 08 one two three, one two three, one two three four
- 09 five six seven. one two three, one two [three ]
- 10 one tw[o three] four five six seven♪
- 11 JHT: [utatte]  
"sing"
- 12 MAK: [UTATTE.] ((iconic singing gestures))  
"sing"
- 13 (0.4) ((music plays))
- 14 ALT: ready::?
- 15 MAK: utau yo: ((iconic singing gestures))  
"{We'll} sing"
- 16 (0.5) ((music plays))
- 17 MAK: si:ng.=
- 18 ALT: =okay let's sing the [°( )°
- 19 MAK: [ TA TA TA
- 20 EV : ♪one two three four five six seven♪

Right after the ALT's directive "please si::ng" (line 3), Makoto deploys a directive verbally and nonverbally in his own format ("♪TAN TAN TAN, ♪" produced with singing gestures). This directive indicates to the students when to start singing and thus facilitates the students' performance of the activity. Toward the end of the first round of the song, in line 11, the JHT, speaking in Japanese, directs the students to sing, "*utatte*" (sing), and Makoto repeats the directive but in a louder voice. Between the first and second rounds of the song, in line 15, Makoto again directs the students in Japanese to sing, but this time with a slight modification, "*utau yo*:" ({{We'll}} sing) with a singing gesture. After a pause, he produces a directive again but this time in English "si:ng." This directive by Makoto is a slight modification of the



to find the paper apples. By verbalizing the action the ALT has already performed nonverbally, Makoto demonstrates his orientation to co-teaching and his identity as being in the same teaching team as the ALT. Students also view Makoto as a teacher: following Makoto's directive, all the students look around the classroom and start counting (line 7). As mentioned earlier, providing directives to others is considered to be universally linked to a higher ranking status (Ervin-Tripp et al., 1984; Ochs, 1996). Thus, by producing directives in the classes, Makoto demonstrated a higher rank and authoritative position with the right and privilege to direct students. In CA terms, he talked this higher rank into being. Moreover, in the later stages, producing directives in different forms from the ALT or JHT or by himself, Makoto demonstrated his at least equal or perhaps sometimes greater control over the class than the ALT or JHT.

In the case of directives, at Time 1 neither trainee produced self-initiated directives in whole-class activities. By Time 3, both Makoto and Shota attempted to deploy self-initiated directives in their own forms, and thus displayed their orientations as fully participating teachers.

## **Conclusion**

This longitudinal study examined how two teacher trainees developed their classroom interactional practices in terms of assessments and directives and constructed their identities as teachers over a period of 19 months.

Emerging from the microanalysis of the data, observable changes were manifested by the two elementary school English teacher trainees in deployment of assessments during classroom activities. Initially, Shota aligned with the assessments given by the ALT and JHTs by joining in with applause and producing embodied assessments in the form of head nods, but without participating in any verbalization of the assessments. Makoto also echoed assessments deployed by the ALT and JHTs through his alignment with the applause they initiated, and also echoed their verbal assessments but only in the same form. By Time 2, a year later, Shota verbalized assessments when interacting individually with students, when translating the JHT's productions from Japanese to English, and when taking charge of whole-class activities. Makoto manifested change through deployment of a greater variety of assessments, and occasionally by deploying some assessments sequentially earlier than the ALT or JHT. Finally, over the period of 19 months covering the data collection analyzed here, both Shota and Makoto deployed assessments in a greater variety of sequential environments, many self-initiated or

sequentially prior to the other teachers, while their follow-up assessments were always upgrades.

This study also described the two trainees' significant changes in deployment of directives. At the beginning, neither trainee verbalized directives in whole-class activities. Shota displayed bodily agreement with the ALT's directives through head nods and hand gestures, and he verbalized directives only outside of the teacher-fronted arena in the form of repeats of the ALT's or JHTs' directives, while Makoto did not publicly verbalize any directives in the class. In the later stage, the two trainees gradually began to display better understanding of how to formulate directives not only by repeating the ALT's or JHTs' directives but also by paraphrasing or expanding the ALT's or JHTs' directives; and by Time 2, Makoto occasionally deployed verbal and nonverbal directives in original forms. By the end of this investigation period, both Shota and Makoto expanded their repertoires of directives, and their verbal and nonverbal deployment of directives became more frequent. Furthermore, both of them at times attempted to initiate directives of their own.

As the trainees' interactional practices evolved, there was a change in their social orientations as teachers. Initially, the trainees participated peripherally and oriented mainly to themselves as assistants. However, as they were socialized into classroom culture and changed their interactional practices, they frequently displayed characteristics that are conventionally associated with teachers. Their orientations to their own identities as fully participating teachers were clearly displayed in the later periods. Moreover, the trainees' orientation in the later periods to being one of the teachers occasionally had some visible effects on interaction: The other teachers, as well as students, treated each of the trainees' assessments and directives as an action carried out by one with equal status as a member of the teaching staff. For instance, the other teachers let the trainees have opportunities to produce assessments and directives first and initiated new sequences after the trainees' assessments, and the students invariably complied with the trainees' directives. Over the 19-month period of this longitudinal study, the teacher trainees learned through interaction how to deploy assessments and directives in socially and culturally meaningful ways through participating in this particular community.

## **Implications and Future Research**

While the importance of practical training in all fields of education is well established, this study provides evidence for the importance of

teaching-training programs from a microanalysis of actual interaction. It demonstrates the positive effects of on-site teacher training, and delineates the processes fledgling trainees undergo in making the transition to fully participating teachers in learning to deploy previously learned language in actual classroom situations.

This research project also demonstrates the efficacy of the system of sending university students to participate as teacher trainees in elementary schools in Japan as their change over time leads to greater support for the elementary school teaching staff, better teaching for the elementary school students, and a benefit to society through an increase in the number of trained and competent teachers.

Future research will need to (a) examine the consequences of such teacher-training programs for the classroom teachers themselves as they assist the trainees in learning to be teachers, (b) appraise the educational impact on the students as they interact with the teachers and teacher trainees, and (c) scrutinize the interaction during the students' development, the trainees' evolution, and the responses of the teachers.

This study has further demonstrated that a microanalysis achieved through the application of CA methodology can reveal step-by-step changes in language deployment in a specific situation over time.

## Acknowledgments

The research presented here was supported by Grant-in-Aid #20520554 from the Japan Society for the Promotion of Science. We wish to thank the elementary school students and staff for allowing us to visit their classes, the teacher trainees who participated in the study, and two anonymous *JALT Journal* reviewers. Any inconsistencies in analysis or presentation remain ours.

## Notes

1. The exact number of teacher trainees varied over the course of this 19-month study.
2. A *response cry* is an exclamatory interjection that provides a clue as to the state of the speaker. Goffman (1981) described it as "a natural overflowing, a flooding up of previously contained feeling, a bursting of normal restraints, a case of being caught off guard" (p. 99). He added that speakers may utter "a mocked-up response cry" (p. 117) to evoke

the image of someone who has need of a particular response cry. In Extract 4 presented in this paper, Shota's cry, "wo:w" (line 3) may not be a response cry that displays an eruption of spontaneous emotion, but a mocked-up response cry which evokes the image of someone who is being surprised. Shota may be deploying this mocked-up response cry deliberately (vs. spontaneously) to transmit a particular message to the students: He is purposely showing surprise at the unexpectedly good performance of the students, and by doing so he accentuates the positive assessment ">very good<" that follows the cry.

3. In CA studies of ordinary conversation, the term "requests" has been used more often than "directives." However, in studies of conversation in which there is an asymmetrical relationship among participants such as parent-child interactions (Cekaite, 2010; Goodwin, 2006) the term "directives" has most often been utilized.
4. Lerner (1993) uses the term "association" to describe the "team-ness" of interactants when one speaks of "any assemblage of co-present individuals that are in any of various ways cast as a collectivity" (p. 214). A speaker may demonstrate membership of association by subsequently joining in the ongoing action. Shota, in Extract 8, subsequently joins in the ongoing action by producing the directives "stand up, stand up. (.) everybody stand up."
5. See Schegloff (2000) for resources interactants employ for resolving overlaps.

*Yuri Hosoda*, EdD, is an Associate Professor at Kanagawa University. Her research interests include conversation analysis and sociolinguistics.

*David Aline*, EdD, is a Professor at Kanagawa University. His research interests include conversation analysis for second language learning and use.

## References

- Antaki, C., & Widdicombe, S. (1998). Identity as an achievement and as a tool. In C. Antaki & S. Widdicombe (Eds.), *Identities in talk* (pp. 1-14). London: Sage.
- Austin, J. (1962). *How to do things with words*. Oxford: Clarendon Press.
- Cekaite, A. (2010). Shepherding the child: Embodied directive sequences in parent-child interaction. *Text & Talk, 30*, 1-25.

- Donato, R. (2004). Aspects of collaboration in pedagogical discourse. *Annual Review of Applied Linguistics*, 24, 284-302.
- Ervin-Tripp, S. M. (1976). Is Sybil there: Some American English directives. *Language in Society*, 5, 25-66.
- Ervin-Tripp, S. M., O'Connor, M. C., & Rosenberg, J. (1984). Language and power in the family. In C. Kramer, M. Schulz, & W. O'Barr (Eds.), *Language and power* (pp. 116-135). Beverly Hills, CA: Sage.
- Garrett, P. B., & Baquedano-López, P. (2002). Language socialization: Reproduction and continuity, transformation and change. *Annual Review of Anthropology*, 31, 339-361.
- Goffman, E. (1981). *Forms of talk*. Philadelphia: University of Pennsylvania Press.
- Goodwin, C., & Goodwin, M. H. (1987). Concurrent operations on talk: Notes on the interactive organization of assessments. *IPRA Papers in Pragmatics*, 1(1), 1-54.
- Goodwin, M. H. (1990). *He-said-she-said*. Bloomington and Indianapolis: Indiana University Press.
- Goodwin, M. H. (2006). Participation, affect, and trajectory in family directive/response sequences. *Text & Talk*, 26, 515-543.
- Goodwin, M. H., & Goodwin, C. (2001). Emotion within situated activity. In A. Duranti (Ed.), *Linguistic anthropology: A reader* (pp. 239-257). Malden, MA: Blackwell.
- Greer, T., Bussinger, V., Butterfield, J., & Mischinger, A. (2009). Receipt through repetition. *JALT Journal*, 31, 5-34.
- Hall, J. K. (2004). Language learning as an interactional achievement. *Modern Language Journal*, 88, 607-612.
- He, A. W. (2004). CA for SLA: Arguments from the Chinese language classroom. *Modern Language Journal*, 88, 568-582.
- Hellermann, J. (2006). Classroom interactive practices for developing L2 literacy: A microethnographic study of two beginning adult learners of English. *Applied Linguistics*, 27, 377-404.
- Hellermann, J. (2007). The development of practices for action in classroom dyadic interaction: Focus on task openings. *Modern Language Journal*, 91, 83-96.
- Hellermann, J. (2009). *Social actions for classroom language learning*. Clevedon, UK: Multilingual Matters.
- Heritage, J., & Clayman, S. (2010). *Talk in action: Interactions, identities, and institutions*. Malden, MA: Wiley-Blackwell.

- Heritage, J., & Raymond, G. (2005). The terms of agreement: Indexing epistemic authority and subordination in talk-in-interaction. *Social Psychology Quarterly*, *68*, 15-38.
- Hogan, J. (2004). Are concerns for the addition of English warranted? *The Language Teacher*, *28*(2), 3-8.
- Hosoda, Y., & Aline, D. (2010a). Positions and actions of classroom specific applause. *Pragmatics*, *20*, 133-148.
- Hosoda, Y., & Aline, D. (2010b). Teacher deployment of applause in interactional assessments of L2 learners. In G. Kasper, H. Nguyen, D. R. Yoshimi, & J. K. Yoshioka (Eds.), *Pragmatics and language learning* (Vol. 12) (pp. 253-274). Honolulu: University of Hawaii Press.
- Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp. 13-31). Amsterdam: John Benjamins.
- Kasper, G. (1997). "A" stands for acquisition: A response to Firth and Wagner. *Modern Language Journal*, *81*, 307-312.
- Kasper, G. (2004). Participant orientations in German conversation-for-learning. *Modern Language Journal*, *88*, 551-567.
- Kelly, C. (2002). Training Japanese elementary school teachers to teach English, *The Language Teacher*, *26*(7), 31-33.
- Lantolf, J. P., & Appel, G. (1994). *Vygotskian approaches to second language research*. Westport, CT: Ablex.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.
- Lerner, G. H. (1993). Collectivities in action: Establishing the relevance of conjoined participation in conversation. *Text*, *13*, 213-245.
- Markee, N., & Kasper, G. (2004). Classroom talks: An introduction. *Modern Language Journal*, *88*, 491-500.
- Mori, J., & Markee, N. (2009). Language learning, cognition, and interactional practices: An introduction. *International Review of Applied Linguistics*, *47*, 1-9.
- Murphey, T., Asaoka, C., & Sekiguchi, M. (2004). Primary teachers co-learning English with their students. *The Language Teacher*, *28*(2), 15-18.
- Ochs, E. (1989). *Culture and language development: Language acquisition and language socialization in a Samoan village*. Cambridge: Cambridge University Press.

- Ochs, E. (1996). Linguistic resources for socializing humanity. In J. J. Gumperz & S. L. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 407-437). Cambridge: Cambridge University Press.
- Ochs, E., & Schieffelin, B. B. (1984). Language acquisition and socialization. In R. Shweder & R. Levine (Eds.), *Culture theory: Essays on mind, self, and emotion* (pp. 276-320). Cambridge: Cambridge University Press.
- Pomerantz, A. (1984). Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 57-101). Cambridge: Cambridge University Press.
- Pomerantz, A., & Fehr, B. J. (1997). Conversation analysis: An approach to the study of social action as sense making practices. In T. A. van Dijk (Ed.), *Discourse as social interaction* (pp. 64-91). London: Sage.
- Raymond, G., & Heritage, J. (2006). The epistemics of social relations: Owning grandchildren. *Language in Society, 35*, 677-705.
- Sacks, H. (1972). On the analyzability of stories by children. In J. J. Gumperz & D. Hymes (Eds.), *Directions in sociolinguistics: The ethnography of communication* (pp. 325-345). New York: Holt, Rinehart, and Winston.
- Schegloff, E. A. (1992). On talk and its institutional occasions. In P. Drew & J. Heritage (Eds.), *Talk at work* (pp. 101-134). Cambridge: Cambridge University Press.
- Schegloff, E. A. (2000). Overlapping talk and the organization of turn-taking for conversation. *Language in Society, 29*, 1-63.
- Schieffelin, B. B., & Ochs, E., (1986). Language socialization. *Annual Review of Anthropology, 15*, 163-191.
- Searle, J. (1976). The classification of illocutionary acts. *Language and Society, 5*, 1-24.
- Sinclair, J., & Coulthard, M. (1975). *Towards an analysis of discourse*. Oxford: Oxford University Press.
- Takagaki, T. (2003). A critical look at elementary school English in Japan from the perspective of the core French program in Ontario, Canada. *The Language Teacher, 27*(6), 17-19.
- van Lier, L. (1998). The relationship between consciousness, interaction and language learning. *Language Awareness, 7*, 128-145.
- Young, R. F., & Miller, E. R. (2004). Learning as changing participation: Discourse roles in ESL writing conferences. *Modern Language Journal, 88*, 519-535.

## Appendix

### *Transcription Conventions*

[        ]	overlapping talk
=	latched utterances
(0.0)	timed pause (in seconds)
(.)	a short pause
co:lon	extension of the sound or syllable
co::lon	a more prolonged stretch
.	fall in intonation (final)
,	continuing intonation (non-final)
?	rising intonation (final)
CAPITAL	loud talk
<u>underline</u>	emphasis
↑	sharp rise
↓	sharp fall
°        °	quiet talk
°°       °°	quieter talk
♪        ♪	singing
<       >	slow talk
>       <	fast talk
hh	audible aspirations
(hh)	laughter within a word
((    ))	comment by the transcriber
(    )	problematic hearing that the transcriber is not certain about
{    }	translation of ellipsis
“    ”	idiomatic translation of Japanese utterances



# Interest and Prior Knowledge in Second Language Reading Comprehension

John Eidswick

*Kwansei Gakuin University*

The influence of *interest* on learning is well-established in educational psychology, but little interest research has been conducted in the field of second language (L2) learning. This study examined interest and an associated factor, prior knowledge, in relation to reading comprehension. Intermediate-level sophomore students in an English program at a private university in Japan ( $N = 23$ ) rated their interest in and knowledge of 11 topics. Correlations between interest and prior knowledge were significant only for topics related to famous individuals. Participants read expository texts about selections from the topic list with high interest-high prior knowledge (HH), high interest-low prior knowledge (HL), and low interest-low prior knowledge (LL), then took multiple-choice comprehension tests. Analysis of Variance (ANOVA) procedures revealed significant differences between scores for the HH text and those of the HL and LL texts. Implications of correlations and test score differences are discussed, and suggestions for future research are offered.

興味・関心が学習に影響を与えることは教育心理学において十分に立証されているにもかかわらず、第二言語(L2)の学習においては、興味・関心に関する研究は従来ほとんど行われてこなかった。本研究では、英文読解に関し、興味・関心およびそれに関連する要因である事前知識についての調査を行った。日本の私立大学の英語プログラムにおいて、中級レベルのクラスで学ぶ2年次の学生(23名)が、11のトピックに対して自分が持つ興味・関心の高さど知識の量の格付けを行った。興味・関心の高さと事前知識の量の間には、著名な人物に関するトピックの場合にのみ、有意の相関が見られた。学生たちは前述のトピックの中から、興味・関心が高く事前知識が多い(HH)、興味・関心が高く事前知識が少ない(HL)、興味・関心が低く事前知識が少ない(LL)トピックを選び、それらに関する説明文を読んだ後、選択問題形式の読解テストを受けた。テストの成績に対して分散分析(ANOVA)を行った結果、HHの説明文に対する成績が、HLおよびLLの説明文に対する成績と有意に異なっていることが明らかになった。これらの相関およびテストの成績の相違が意味するものについて考察し、今後の研究への提言を述べる。

Decades of research in educational and cognitive psychology have confirmed a strong relationship between *interest* and learning, particularly related to reading comprehension and text-based learning. It is surprising then that so few studies have been conducted to investigate the influence of interest on second language (L2) learning.

Individuals interested in a domain, activity, topic, or thing (henceforth “object”) are more persistent, engaged, and attentive when interacting with the object of interest (Ainley, Hidi, & Berndorff, 2002; Hidi, 1990, 2000; Renninger, 2000). Interest affects choices for the direction and duration of attention (Hidi, 1990; Schraw & Lehman, 2001). Both reader- and text-based interest have been found to positively and consistently influence comprehension and learning in a wide range of conditions and reading proficiency levels (for a review, see Hidi, 2001). The positive influence of interest has been shown to be independent of text difficulty, reading ability, and age (Schiefele, 1996). Interest likely derives its beneficial relationship with comprehension and learning from the increases in attention and engagement it generates (see Sadoski, 2001). Despite indications of a strong link between interest and comprehension in domains occurring in students’ first language (L1), the results of the scant studies of interest and L2 reading comprehension (e.g., Brantmeier, 2006; Eidswick, 2009) have not clearly indicated a positive relationship between the two factors.

Interest is commonly divided into *individual interest* and *situational interest*. Individual interest is considered to be a stable and enduring inclination to reengage with specific objects (Bergin, 1999; Hidi, 1990; Hidi, Berndorff, & Ainley, 2002; Krapp, Hidi, & Renninger, 1992; Schiefele, 1999). Reengagement with an object of interest produces increases in knowledge about the object. Knowledge acquired from engaged, self-determined activities related to individual interest is likely to be qualitatively different from knowledge gained from academic, institutionally determined activities.

Situational interest is a usually short-lived emotional state triggered by qualities of an object. Among the qualities found to evoke situational interest are visual imagery, novelty, surprise, concreteness, and intensity (Hidi & Baird, 1986, 1988), along with ease of comprehension (Mitchell, 1993). In developing a model of L2 learner reading interests, Brantmeier (2006) identified ease of recollection, engagement, cohesion, emotiveness, and prior knowledge as interesting in texts. In a study by Eidswick, Praver, & Rouault (2010), intermediate university students of English in Japan, responding to selected graded readers, endorsed as interesting those books that possessed qualities that could be described with the following adjectives:

tives: exciting, mysterious, romantic, heartwarming, strange, unpredictable, funny, scary, and easy to comprehend.

Hidi and Renninger's (2006) Model of Interest Development describes a continuum with four discrete phases: *triggered situational interest*, *maintained situational interest*, *emerging individual interest*, and *well-developed individual interest*. These phases are sequential, commencing with a spark of situational interest that can lead to further contacts with an object, and eventually to strongly committed individual interest. For L2 learners, this process might commence with an initial exposure to a new language that arouses their interest, and which can be sustained by subsequent interaction with interesting materials and activities with which the language is presented, or lost when such materials or activities fail to inspire or maintain learners' interest.

### **Prior Knowledge**

Research has demonstrated the importance of prior knowledge to a range of skills conducive to L2 text-based learning. These include the efficiency of attention allocation to input (Robinson, 2003), inference skills (Barry & Lazarte, 1998), and reading comprehension (Carrell, 1987; Chen & Donin, 1997; Hudson, 1982).

Prior knowledge can contribute to reading comprehension to the extent that it poses problems of test bias (Alderson, 2000). If a student does well on a reading test, is it because of strong reading ability or because of background knowledge of the reading topic or domain? A quality shaping the contribution of prior knowledge to L2 reading comprehension is the degree of reader familiarity with culturally related information within a text. Individuals who are unfamiliar with culture-specific textual elements are less likely to understand a text than individuals who are familiar with them (Johnson, 1981; see also Chihara, Sakurai, & Oller, 1989).

### **Prior Knowledge and Interest**

The condition of knowledge accumulation being a consequence of and an antecedent for interest presents complications for studies whose results signify interest effects on learning or comprehension. Does an individual's differential performance on an assessment instrument related to a high-interest object demonstrate influences of interest or of prior knowledge? While it has been shown that both prior knowledge and interest play roles in learning, research into conjunctions of these two factors has not clarified

completely how or to what extent they are correlated in specific domains and interact to contribute to human development (for a review, see Alexander, Kulikowich, & Jetton, 1994). Difficulty in examining the relationship between interest and prior knowledge arises in part from complexities in the constructs themselves, both of whose dimensions change depending on qualities of objects of interest and the individuals experiencing them, as well as the variable length and frequency of time of exposure and qualities of social environments in which such objects are experienced.

Results of the few interest-prior knowledge studies that have been conducted in L2 learning show little concordance. In a study of interest sources and reading comprehension with advanced native-English-speaking university students of Spanish in the United States, Brantmeier (2006) used factor analysis to designate prior knowledge as a component of situational interest, explaining 13% of its variance. Brantmeier also found that prior knowledge, as a constituent of situational interest, was positively related to reading comprehension on the 10-item multiple-choice section of a comprehension test, but not on the total recall or sentence completion sections. Bügel and Buunk (1996) explored the role of interest and prior knowledge as related to sex differences in reading comprehension among English as a foreign language (EFL) high school students in the Netherlands. While the researchers found significant differences between females and males in prior knowledge and interest, and claim their findings are evidence of sex-related differential effects of the two variables on reading comprehension, it is possible that reading proficiency emerged as a confound of the results. In a study examining interest, prior knowledge, and reading comprehension among English as a second language (ESL) students at an American university, Carrell and Wise (1998) found no significant effect for interest or prior knowledge on scores from 10-item multiple-choice reading comprehension tests. The researchers also found no significant correlation between prior knowledge and interest. Citing speculation by Baldwin, Peleg-Bruckner, and McClintock (1985), whose research had produced similar results among American middle-school students reading in their L1, Carrell and Wise attributed this finding to the condition that in school students sometimes are made to acquire knowledge about academic topics in which they have little interest. As the researchers provided largely academic topics (e.g., computers, human evolution, Islamic art), it is likely that when participants in this study designated their level of prior knowledge about a topic as high, the prior knowledge they identified had been acquired from experiences related to required academic study instead of individual or situational interest. This

might in turn have contributed to a lack of significance in the comprehension test scores as they related to interest and prior knowledge.

Carrell and Wise (1998) identify as a limitation in their study a characteristic shared by many ESL classes: the heterogeneity of the participants' native languages. Evidence appeared suggesting that within groups of students from the same country (e.g., Russia), differences existed with regard to (at least) gender and reading comprehension that could have differentiated specific national groups from others and influenced the results of the study overall. The researchers recommended future studies be conducted with participants that share a common L1.

The present investigation incorporates elements of the design used in the Carrell and Wise study and includes two important modifications. First, this study was conducted with EFL participants with the same L1; and second, reading topics were used that students were unlikely to have studied without having a personal interest.

The following research questions were explored. Among EFL learners:

1. What is the relationship between interest and prior knowledge?
2. What topic qualities evoke varying configurations of interest and prior knowledge?
3. Do variable configurations of interest and prior knowledge influence reading comprehension differently?

## Method

Research was conducted during the fall semester of 2009. Participants were 23 sophomore students (TOEFL ITP test scores of 430-525), majoring in humanities, an intermediate-level four-skills English class in an accelerated English program at a private university in Nishinomiya, Hyogo, Japan. Three of the students were male and the rest were female. Entry to the program is competitive, based on TOEFL scores. Students who complete the intermediate course can satisfy their 2-year university English requirement in one year.

Participants received 11 short descriptions of articles (examples in Appendix A), which they used to rate, via a 6-point Likert scale, their anticipated interest in reading. Participants were also asked to estimate their knowledge of the article topics. From the online encyclopedia Wikipedia, topics were selected whose thematic variation seemed likely to generate a wide range of interest and prior knowledge and were unlikely to be topics students had been required to study in school. Two of the topics, *Takafumi Horie* (the unconventional founder of the Livedoor web portal who was

arrested for securities fraud) and *Akashiya Sanma*, a famous Japanese comedian, were chosen because as these figures had appeared regularly on Japanese television, participants were likely to possess prior knowledge of them, regardless of interest level. Descriptive data for participant responses to the 11 topics is presented in Table 1.

**Table 1. Descriptive Statistics for Perceived Interest and Prior Knowledge**

Topic	Interest		Prior Knowledge	
	M	SD	M	SD
Akashiya Sanma	4.39	1.44	3.04	1.11
Candlemaking	4.48	1.08	2.43	1.12
Chocolate	4.57	1.12	2.70	1.11
Concrete	2.57	1.61	1.70	0.93
Ecocities	3.61	1.23	2.04	1.02
Ghosts	4.91	1.08	2.35	1.30
Japanese cuisine	3.74	1.18	2.78	1.24
Michael Jackson	4.96	1.33	4.83	1.30
Perfume	4.48	1.20	2.22	1.00
Takafumi Horie	4.48	1.08	4.04	1.28
Viruses	3.61	1.23	2.04	1.02

Topics that rated the highest on interest and knowledge (HH), the lowest on interest and knowledge (LL), and the highest on interest yet lowest on knowledge (HL) were chosen. To be used for the study, the HH and LL topics were required to have interest and prior knowledge ratings that were higher (in the case of HH) or lower (in the case of LL) when the means were added together than any other interest-prior knowledge configurations. To qualify as HL, a topic had to have a greater difference between the means for interest and prior knowledge than other configurations. Also, the high and low figures, respectively, of the scores had to fall within the higher or lower

half of the pertinent 6-point Likert scale. Using these criteria, the following topics were chosen: *Michael Jackson* (HH), *Ghosts* (HL), and *Concrete* (LL). Since no pairs of ratings met the requirements for low interest-high prior knowledge (LH), this configuration was not used.

Articles based on the three topics were prepared based on texts drawn from the entries for these topics on Wikipedia (example in Appendix B). The articles were shortened, and footnotes and hypertext links were removed. The articles were otherwise unmodified. Table 2 provides structural information about the three articles, including Flesch-Kincaid readability scales measures. The readability statistics place all the texts within the 3rd grade of the Flesch-Kincaid scale.

**Table 2. Article Lengths and Flesch-Kincaid Readability Scores**

Article	Number of words	Words per sentence	Readability
Michael Jackson (HH)	761	19.3	60.8
Ghosts (HL)	877	23.7	54.4
Concrete (LL)	552	19.7	56.4

*Note.* HH = high interest, high prior knowledge, HL = high interest, low prior knowledge, LL = low interest, low prior knowledge

In a counterbalanced manner, students were given one of the three articles in class each week, over a 3-week period. During a given session, each of the three articles was read by a third of the students, followed by randomized rotations of articles on subsequent weeks. Students were permitted to read for 15 minutes and to use dictionaries. After each reading, the articles were collected, and the participants completed a comprehension test containing 10 multiple-choice items (example in Appendix C).

## Results

### *Pre-Test Topic Ratings*

A Kolmogorov-Smirnov (K-S) procedure indicated that the interest and prior knowledge ratings for 10 out of the 11 proposed topics were non-

normally distributed. For these 10 topics, Kendall's *tau* ( $\tau$ ), a procedure for evaluating correlation between paired variables with nonparametric distribution and one especially appropriate for small data sets (see Field, 2005) was conducted. A Pearson correlation coefficient ( $r = .12$ ) was produced for the only topic (*Chocolate*) with normally distributed data. Table 3 shows correlations for student ratings.

**Table 3. Correlations Between Interest and Prior Knowledge Ratings of Article Topics**

Topic	$\tau$
Akashiya Sanma	.46*
Candlemaking	-.02
Concrete	-.15
Ecocities	.25
Ghosts	.32
Japanese cuisine	.12
Michael Jackson	.45*
Perfume	.16
Takafumi Horie	.64*
Viruses	-.14

Note. \*  $p < .05$

\* Significantly high correlation was found between interest and prior knowledge ratings on three topics: *Michael Jackson*, *Akashiya Sanma*, and *Takafumi Horie*.

### **Comprehension Test Results**

The results of another Kolmogorov-Smirnov (K-S) procedure were insignificant, indicating that all three sets of test scores were normally distributed. Reliability measures (*Michael Jackson*:  $a = .740$ ; *Ghosts*:  $a = .717$ ; *Concrete*:  $a = .713$ ) for the tests indicated acceptable reliability. Descriptive statistics are provided in Table 4. Mauchly's test was not significant ( $\chi^2(2) = 5.17, p < .05$ ), indicating the assumption of sphericity was not violated.

**Table 4. Descriptive Statistics of Comprehension Scores for Interest-Prior Knowledge Configurations**

Article	N	M	SD
Michael Jackson (HH)	23	7.73	1.87
Ghosts (HL)	23	6.53	2.10
Concrete (LL)	23	4.93	2.05

*Note.* HH = high interest, high prior knowledge, HL = high interest, low prior knowledge, LL = low interest, low prior knowledge

A one-way repeated-measures Analysis of Variance (ANOVA) was conducted to examine the relationship between reading comprehension test scores and the three configurations of interest and prior knowledge. The ANOVA showed a significant effect,  $F(2, 28) = 6.61, p < .05, w^2 = .27$ .

A Bonferonni post hoc test revealed that the *Michael Jackson* (HH) article had significantly higher comprehension scores than either of the *Concrete* and *Ghosts* articles ( $p < .05$ ), but no significant differences were found between the scores of *Ghosts* and *Concrete*. A summary is displayed in Table 5.

**Table 5. Contrasts Between Interest-Prior Knowledge Configurations**

Comparisons	df	MS	F	R
HH vs. HL	1	21.6	5.56	.53*
HL vs. LL	1	38.4	3.55	.44
HH vs. LL	1	117.6	9.66	.65*

\*  $p < .05$

## Discussion

The first two research questions of this study concerned (1) the nature of the relationship between interest and prior knowledge, and (2) the topic qualities that influenced varying configurations of interest and prior knowledge. Correlations between the two factors showed significance on only 3 of the original 11 topics. These were *Michael Jackson*, *Akashiya Sanma*, and

*Takafumi Horie*. The topic which earned the highest interest-prior knowledge rating was *Michael Jackson*; also relatively high were *Akashiya Sanma* and *Takafumi Horie*. In other words, the topics which had statistically significant correlations were also those with the highest interest-prior knowledge configurations. This finding accords well with the conceptualization of individual interest as a cognitive and affective state which gives rise to increases in knowledge about objects of interest, and of situational interest as a phenomenon whose constituents include prior knowledge.

It is intriguing, and a finding likely of interest to any teacher who wishes to better incorporate interest as a means of facilitating classroom learning, that these three topics, in addition to garnering the highest correlations and interest-prior knowledge ratings, were also the only ones that focused on named people. It is possible that topics based on human biography heighten interest because of their perceived personal relevance or value; readers are perhaps intrinsically drawn to reading about human beings whose personal concerns or areas of endeavor they see as immediately pertinent to their own, and who become foci of social interaction via widespread media exposure. It is perhaps unsurprising that two of these three named people were Japanese celebrities who are familiar to Japanese television viewers. A reasonable conjecture would be that for many students, these famous people were objects of emerging or well-developed individual interest. Potential existed, in terms of their longevity in public life (especially in the cases of Akashiya Sanma and Michael Jackson) and relevance to the lives of Japanese university students (especially in the case of Takafumi Horie) for deliberate repeated and engaged experiences via television and other media sources. The saturation news coverage of the criminal cases involving Michael Jackson and Takafumi Horie, and, more recently, the former's death, might also have sparked situational interest components such as novelty, intensity, excitement, mystery, sentimentality, unpredictability, and (in the case of Michael Jackson) strangeness.

It is interesting to note that another topic relevant to Japanese culture, *Japanese Cuisine*, garnered relatively low ratings for prior knowledge. It is possible that students in this age group (who are popularly reputed to lack cooking skills) interpreted their prior knowledge of *Japanese Cuisine* to mean knowledge of cooking rather than eating it.

*Ghosts* was the topic whose responses evidenced the widest divide between high interest and low prior knowledge. Students expressed relatively high interest in three other topics that they professed relatively low knowledge about: *Candlemaking*, *Chocolate*, and *Perfume*. The mysterious, scary

nature of ghosts could well have provoked sensations of situational interest, but the topics of candlemaking, chocolate, and perfume perhaps seemed interesting to participants because of personal relevance. Since chocolate and perfume are common objects, familiar to everyone, but participants claimed low prior knowledge of these objects, it is probable that prior knowledge, as in the case of *Japanese Cuisine*, was interpreted to mean procedural knowledge as to how the object was produced. The appearance on the alphabetical-order topic list of *Chocolate*, *Perfume*, and *Japanese Cuisine* after *Candlemaking* might have prompted participants to perceive *Chocolate* as chocolate making and *Perfume* as perfume making. By whatever means the interpretation of these topics arose, it is also possible that since candlemaking, chocolate, and perfume are objects that are both exceedingly ordinary yet possess origins of manufacture that are mysterious to many people, situational interest regarding these origins was aroused.

*Concrete*, the topic for which participants indicated the lowest interest-prior knowledge configuration, appeared to lack elements of situational interest, and it is doubtful that any participants had experienced many engaged encounters with the object of this topic, whose text, like others on the list, seems to have been anticipated as an explanation of a procedure. Few students claimed to have much familiarity with concrete, despite having walked on it while coming to class. Future research might examine the strategies that learners use to perceive intended meanings of concepts that possess multiple potential interpretations.

The results of the one-way repeated-measures ANOVA supported an affirmative answer to the third research question: configurations of interest and prior knowledge do influence reading comprehension differently. As noted, comprehension test results for the three configurations showed significant differences between the scores for *Michael Jackson* (HH) and both *Ghosts* (HL) and *Concrete* (LL), but not between *Ghosts* (HL) and *Concrete* (LL). Put another way, the two articles with high interest ratings produced test scores that were significantly different from each other, and one each of the high interest and low interest topics produced test scores that were not significantly different. Prior knowledge, on the other hand, consistently corresponded with the significant test score differences, suggesting that prior knowledge played a different and more influential role in text comprehension than did interest. Another possibility is that high interest and high prior knowledge interacted in a way that fortified text comprehension, and that this interaction differed qualitatively from interactions between the factors in other HL and LL configurations. Evidence for this interaction might lie in

the finding that the interest and prior knowledge ratings for the pre-reading topic *Michael Jackson* correlated significantly, whereas the interest and prior knowledge of other configurations did not.

## Conclusion

This study examined the bases for three configurations of high and low interest and prior knowledge prompted by nonacademic text topics, and whether these configurations differentially related with comprehension for expository texts. Participant responses to a list of text topics prompted varying interest and prior knowledge evaluations. These evaluations differed depending on whether characteristics of topics prompted perceptions of interest components such as personal relevance, mysteriousness, intensity, excitement, and unpredictability. Among the selection of 11 topics, only three related to named people, with high profiles in the media. These three topics also rated the highest on the list for high interest-high prior knowledge, suggesting that biography-related topics, especially about famous people, might be more likely to be experienced as both more interesting and more familiar than other topic categories. Analysis of multiple-choice comprehension test results revealed significantly higher scores for the high interest-high prior knowledge (HH) text than for those of the other texts, but no significant differences between the test scores for high interest-low prior knowledge (HL) and low interest-low prior knowledge (LL) texts. One interpretation of these results is that they provide support for a positive influence of prior knowledge on reading comprehension, and do not provide support for a similar influence by interest. Another interpretation is that, given that the HH text was based on one of the three topics that rated highest for both interest and prior knowledge and for correlation between them, the interaction of interest and prior knowledge facilitated comprehension in ways in which insignificantly correlated configurations did not.

While the results offer insights into the relationship between interest, prior knowledge, and reading comprehension, certain limitations in the study make it difficult to make strong claims. If the number of participants had been larger, the statistical power of the procedures, and the ability to be confident of their results, would have been greater. Also, although the Flesch-Kincaid readability formula is widely used to evaluate text difficulty, and has been shown to be consistent in predicting text difficulty (Fry, 1989), the use of the formula with L2 readers is not without controversy (see Greenfield, 2004). Future L2-interest researchers might wish to examine the dimensions of the present study in the light of different approaches to text diffi-

culty measurements. Furthermore, because of time constraints, estimations of participants' prior knowledge in this study were limited to self-reports via Likert-scales. While agreement is lacking on the best ways to measure prior knowledge (Grabe, 2009), other approaches that examine knowledge with more depth, such as tests targeting participants' comprehension of topic-specific vocabulary, or knowledge of closely related domains, might provide a fuller picture of participant prior knowledge in future studies. Time constraints also prevented further attempts to find a topic that conformed to the prerequisites for the configuration of low interest-high prior knowledge after this configuration was not found in the ratings of the initial 11 topics. Research that incorporates this missing fourth configuration will likely extend and enrich the results of the present study, which appear to indicate differential effects for low and high prior knowledge. Limitations related to operationalization of factors in this study highlight areas for further exploration. No control was provided for the effects of covariation between factors preventing a clearer understanding of whether interactions between interest and prior knowledge could facilitate comprehension. This study operationalized text comprehension using only a 10-item multiple-choice instrument, without attempt to analyze separately and compare categories important to comprehension, such as general, specific, and inferential knowledge. In order to gain a better understanding of the relationship between interest, prior knowledge, and text comprehension, more research needs to be conducted that investigates connections between the two factors and multiple forms of text comprehension, with texts that use modalities outside of expository writing and with research designs which enable researchers to conduct more precise examination of interaction effects.

*John Eidswick* is an instructor of English at Kwansai Gakuin University and a doctoral candidate in TESOL. His research focuses include interest, learner autonomy, and L2 writing.

## References

- Ainley, M. D., Hidi, S., & Berndorff, D. (2002). Interest, learning, and the psychological processes that mediate their relationship. *Journal of Educational Psychology, 94*, 545-561.
- Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.

- Alexander, P., Kulikowich, J., & Jetton, T. (1994). The role of subject-matter knowledge and interest in the processing of linear and non-linear texts. *Review of Educational Research, 64*, 201-252.
- Baldwin, R. S., Peleg-Bruckner, Z., & McClintock, A. H. (1985). Effects of topic interest and prior knowledge on reading comprehension. *Reading Research Quarterly, 20*, 497-504.
- Barry, S., & Lazarte, A. (1998). Evidence for mental models: How do prior knowledge, syntactic complexity, and reading topic affect inference generation in a recall task for nonnative readers of Spanish? *The Modern Language Journal, 82*, 176-193.
- Bergin, D. A., (1999). Influences on classroom interest. *Educational Psychologist, 34*, 87-98.
- Brantmeier, C. (2006). Toward a multicomponent model of interest and L2 reading: Sources of interest, perceived situational interest, and comprehension. *Reading in a Foreign Language, 18*, 89-115.
- Bügel, K., & Buunk, B. P. (1996). Sex differences in foreign language text comprehension: The role of interests and prior knowledge. *The Modern Language Journal, 80*, 15-31.
- Carrell, P. (1987). Content and formal schemata in ESL reading. *TESOL Quarterly, 21*, 461-481.
- Carrell, P. L. & Wise, T. E. (1998). The relationship between prior knowledge and topic interest in second language reading. *Studies in Second Language Acquisition, 20*, 285-309.
- Chen, Q., & Donin, J. (1997). Discourse processing of first and second language biology texts: Effects of language proficiency and domain specific knowledge. *The Modern Language Journal, 81*, 209-227.
- Chihara, T., Sakurai, T., & Oller, J. W. (1989). Background and culture as factors in EFL reading comprehension. *Language Testing, 6*, 143-151.
- Concrete (n.d.). Retrieved August 5, 2009, from Wikipedia: <http://en.wikipedia.org/wiki/Concrete>
- Eidswick, J. (2009). The influence of interest on reading comprehension in EFL students. *Annual Research Report of the Language Center, 12*, 25-38. Kwansei Gakuin University.
- Eidswick, J., Praver, M., & Rouault, G. (2010). Extensive reading and reading circles for ESL/EFL students. *Extensive Reading in Japan Journal, (3)*1.
- Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage Publications.

- Fry, E. (1989). Reading formulas: Maligned but valid. *Journal of Reading, 32*, 292-297.
- Ghosts (n.d.). Retrieved August 5, 2009, from Wikipedia: <http://en.wikipedia.org/wiki/Ghosts>
- Grabe, W. (2009). *Reading in a second language*. Cambridge: Cambridge University Press.
- Greenfield, J. (2004). Readability formulas for EFL. *JALT Journal, 26*, 5-24.
- Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research, 60*, 549-571.
- Hidi, S. (2000). An interest researcher's perspective on the effects of extrinsic and intrinsic factors on motivation. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic motivation: Controversies and new directions* (pp. 309-339). New York: Academic Press.
- Hidi, S. (2001). Interest, reading, and learning: Theoretical and practical considerations. *Educational Psychology Review, 13*, 191-209.
- Hidi, S., & Baird, W. (1986). Interestingness—a neglected variable in discourse processing. *Cognitive Science, 10*, 179-194.
- Hidi, S., & Baird, W. (1988). Strategies for increasing text-based interest and students' recall of expository texts. *Reading Research Quarterly, 23*, 465-483.
- Hidi, S., Berndorff, D., & Ainley, M. (2002). Children's argument writing, interest, and self-efficacy: An intervention study. *Learning and Instruction, 12*, 429-446.
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*, 111-127.
- Hudson, T. (1982). The effects of induced schemata on the short-circuit in L2 reading: Non-decoding factors in L2 reading performance. *Language Learning, 32*, 1-32.
- Johnson, P. (1981). Effects on reading comprehension of language complexity and cultural background of a text. *TESOL Quarterly, 15*, 169-181.
- Krapp, A., Hidi, S., & Renninger, K. A. (1992). Interest, learning, and development. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 3-25). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Michael Jackson (n.d.). Retrieved August 5, 2009, from Wikipedia: [http://en.wikipedia.org/wiki/Michael\\_Jackson](http://en.wikipedia.org/wiki/Michael_Jackson)
- Mitchell, M. (1993). Situational interest: Its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology, 85*, 424-436.

- Renninger, K. A. (2000). Individual interest and its implications for understanding intrinsic motivation. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimum motivation and performance* (pp. 373-404). New York: Academic Press.
- Robinson, P. (2003). Attention and memory during SLA. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 631-678). Malden, MA: Blackwell Publishing.
- Sadoski, M. (2001). Resolving the effects of concreteness on interest, comprehension, and learning important ideas from text. *Educational Psychology Review*, 13, 263-281.
- Schiefele, U. (1996). Topic interest, text representation, and quality of experience. *Contemporary Educational Psychology*, 21, 3-18.
- Schiefele, U. (1999). Interest and learning from text. *Scientific Studies of Reading*, 3, 257-279.
- Schraw, G., & Lehman, S. (2001). Situational interest: A review of the literature and directions for future research. *Educational Psychology Review*, 13, 23-52.

## Appendix A

### *Please read the following titles and descriptions of articles*

In response to the statements, circle a number from 1 to 6 (“1” means that you disagree very strongly and “6” means you agree very strongly).

#### 1. “Michael Jackson”

This article is about singer Michael Jackson, his rise to fame and his descent into strange behavior and early death.

I think this article will be interesting to read.

1            2            3            4            5            6

I know much about the topic of this article.

1            2            3            4            5            6

#### 2. “Concrete”

This article describes how concrete is made (its ingredients and methods for making it) and what it is used for.

I think this article will be interesting to read.

1            2            3            4            5            6

I know much about the topic of this article.

1            2            3            4            5            6

## Appendix B

### *Michael Jackson*

Michael Jackson was an American recording artist. He made his debut as an entertainer in 1968 as a member of The Jackson 5. That successful career led to him being dubbed the “King of Pop.” Jackson’s 1982 album *Thriller* remains the world’s best-selling album of all time. His albums *Bad*, *Dangerous*, and *HIStory* also sold very well. One of the few artists to have been inducted into the Rock and Roll Hall of Fame twice, his other achievements include multiple *Guinness World Records*—including one for “Most Successful Entertainer of All Time”—13 Grammy Awards, 13 number one singles in his solo career—more than any other male artist in the Hot 100 era—and the sales of over 750 million records worldwide. Along with his musical success, Jackson’s bizarre personal life, which included a pet chimp named Bubbles, sleeping in a special oxygen chamber (supposedly to prolong his life), dramatic whitening of his skin (it was rumored that Jackson was bleaching his skin), radical changes to his facial appearance because of many plastic surgeries (which resulted in his nose actually collapsing), and accusations of sexual molestation of boys, has made him a part of popular culture for almost four decades. President George H. W. Bush presented the singer with The White House’s special “Artist of the Decade” award in recognition of Jackson’s musical influence in the 1980s.

### *Music*

In the early 1980s, he became a dominant figure in American popular music and culture. He was the first African American entertainer to appear regularly on MTV. Videos such as “Black or White” and “Scream” made Jackson an enduring staple on MTV well into the 1990s. Jackson popularized some complicated dance techniques, such as the robot and the moonwalk, with his elaborate stage performances. His stunning performance on MTV’s show celebrating the fiftieth anniversary of Motown Records has been compared to the Beatles’ legendary appearance on the Ed Sullivan Show. Jackson donated and raised millions of dollars through his Heal the World Foundation.

### *Abused Child*

Jackson was abused by his father, Joseph Jackson, from a young age, enduring whippings and name-calling. In one altercation—later recalled by Jackson's brother Marlon Jackson—Joseph held Michael upside down by one leg and “punched him over and over again with his hand.” Joseph would also trip or push his male children into walls. One night while Jackson was asleep, Joseph climbed into his room through the bedroom window. Wearing a fright mask, he entered the room screaming and shouting. Joseph said he wanted to teach his children not to leave the window open when they went to sleep. For years afterwards, Jackson said he suffered nightmares about being kidnapped from his bedroom. When Michael Jackson was on trial for child molestation years later, a psychiatrist who interviewed him said that because of this abuse, the singer was a “regressed 10-year-old,” a child in a man's body, and was probably not capable of being a molester.

### *Personal Life*

Michael Jackson's personal life generated significant controversy. Though he was accused of child sexual abuse in 1993, the criminal investigation was closed due to lack of evidence. Jackson was reported to have secretly paid his 13-year-old accuser 22 million dollars. During a documentary in 2003, Michael Jackson stated that he often had boys (visitors to Jackson's Neverland Ranch theme park) sleep overnight with him in his bed, and these were non-sexual “slumber parties.” Jackson was later accused by two other boys of sexual abuse. He was tried and acquitted of abuse in 2005. While making a commercial for Pepsi, the singer's hair caught fire and his head was burned. Jackson later donated all of the 1.5 million dollars Pepsi paid him to a charity for burn victims. Falls while dancing caused a spinal injury and a broken nose, for which Jackson had his first experience with plastic surgery to his nose. Some report that he required much medicine for the pain from these injuries, which he began to take regularly. Over the years, he became very thin, and some claimed he was anorexic.

Jackson married twice (once to the daughter of Elvis Presley) and fathered three children. The mother of his first two children (Prince Michael Jackson and Paris Michael Katherine Jackson), Debbie Rowe, claims that she was impregnated by artificial insemination from an “unknown man.” The mother of Jackson's third child (Prince Michael II) has never been identified. The singer died on June 25, 2009, after suffering cardiac arrest. Jackson's attempts to control the symptoms of injuries were probably related to his death.

**Appendix C****Michael Jackson Comprehension Test**

1. The best-selling album of all time is called \_\_\_\_\_
  - a. Bad
  - b. Thriller
  - c. Dangerous
  - d. HIStory
  
2. Michael Jackson was given an “Artist of the Decade” award by \_\_\_\_\_.
  - a. The Rock and Roll Hall of Fame
  - b. President George W. Bush
  - c. The Guinness Book of World Records
  - d. President George H. W. Bush
  
3. Some claimed that Jackson slept in a special oxygen chamber to
  - a. whiten his skin
  - b. help his breathing
  - c. extend his life
  - d. have private time with boys
  
4. Joe Jackson wore a fright mask at night \_\_\_\_\_
  - a. to teach Michael Jackson to keep the window shut
  - b. to push his children to not be lazy
  - c. because he enjoyed abusing his children
  - d. because Michael was having nightmares
  
5. In total, \_\_\_\_\_ boys accused Michael Jackson of molestation.
  - a. 1
  - b. 2
  - c. 3
  - d. 4
  
6. From this article, we can understand that \_\_\_\_\_.
  - a. Michael sometimes slept with a chimp.
  - b. in the 70s, MTV did not feature black musicians.
  - c. Michael was a child molester.
  - d. Michael was embarrassed about being African American.
  
7. The article implies that Joe Jackson \_\_\_\_\_
  - a. was a child molester
  - b. was jealous of his son’s talent
  - c. was not violent with his daughters
  - d. had a bizarre sense of humor

8. The article implies that \_\_\_\_\_
- Jackson died because of use of pain drugs.
  - Debbie Rowe is the mother of Jackson's third child, but wants this to be a secret.
  - Michael gambled with his money and lost much of it.
  - Michael's pain medicine was expensive.
9. Michael Jackson \_\_\_\_\_.
- was willing to pay lots of money to charity.
  - was in financial trouble because of his relationships with boys
  - had spent too much money on plastic surgery
  - had money problems because his albums didn't sell well after *Thriller* and *Bad*.
10. Debbie Rowe implied that \_\_\_\_\_
- she didn't want to be married to Michael Jackson
  - she wanted custody of two of the children.
  - she was angry about Michael having a third child without her.
  - Michael Jackson was not the father of any of the children.

# 多肢選択式項目の項目形式が文法テストパフォーマンスに与える影響について

## The Effects of Multiple-Choice Item Formats on Grammar Test Performance

熊澤孝昭 (くまざわたかあき)

関東学院大学

Item formats are a facet of testing that influences examinees' test performance. In this study, six types of item formats were adopted, and 55 multiple-choice grammar items were developed and administered to 608 first-year university students for placement and diagnostic purposes. The research questions were: to what extent do items function for placement and diagnostic purposes, to what extent do item formats differ in terms of difficulty, to what extent do item formats contribute to the total score variance, and to what extent is the grammar test reliable. Based on the item analyses, most items functioned for placement and diagnostic purposes. FACETS analysis revealed that the six item formats differed in terms of difficulty. The generalizability study showed that 3% of the variance components was due to the item formats. The decision study indicated that the generalizability coefficient and dependability index were satisfactory for placement and diagnostic purposes. Based on the results, implications are discussed.

**Keywords:** grammar item format, multiple-choice items, test methods, placement test, diagnostic test, generalizability theory, multifaceted Rasch model

項目形式はテストパフォーマンスに影響する一つの要因だといわれている。本研究では6種類の項目形式を用いて55項目の多肢選択式項目を文法テストとして作成し、608名の大学一年生にブレイスメントテストと診断テストとして実施した結果を報告する。研究目的は、項目の機能、項目形式の困難度、項目形式などの分散成分の割合、文法テストの信頼性の検証である。項目分析の結果、ほとんどの項目がブレイスメントと診断テストの項目として機能していたことがわか

った。FACETS分析の結果、6種類の多肢選択式項目の形式はそれぞれ困難度が異なることがわかった。一般化可能性研究の結果、項目形式の違いによって生じた分散成分が若干あり、文法テストパフォーマンスに影響する要因であることも明らかとなった。決定研究の結果、一般化可能性係数は.81で、信頼度指数は.75であった。上述の結果を踏まえ、教育的示唆について論じた。

**キーワード:** 文法項目形式、多肢選択項目、テスト方法、プレイスメントテスト、診断テスト、一般化可能性理論、多相ラッシュモデル

## はじめに

テストを受験している際のパフォーマンスは得点という結果になるが、その得点に影響すると思われる要因は様々であり、言語テストの分野ではその要因の解明や影響の度合いが調査されてきた。その要因の中でも、テスト環境、テスト指示、インプット、予期する応答、インプットと応答の関係であるテスト方法(test method)は受験者のテストパフォーマンスに大きく影響する相(facet)だといわれている(Bachman, 1990)。予期する応答についてはさらに選択形式応答の項目(selected response item)と応答構築式型の項目(constructed response item)の二つの形式に分けられる。真意判定項目(true/false item [T/F])、多肢選択式項目(multiple choice item [MC])、整合項目(matching item)などは受験者がいくつかある選択肢から正解を一つ選ぶ形式なので、選択形式応答の項目である。空欄箇所補充項目(fill-in item)、短答形式の項目(short-response item)、クローズ項目(cloze item)などは受験者が英単語などの言語的なアウトプットを記述することで正解することができるため、応答構築式型の項目である。これらの項目形式は受験者のテストパフォーマンスに影響するため、テスト作成者はテストの目的に沿うようもとも妥当な形式を選ぶ必要がある。

応答構築式型の項目より選択式応答の項目のほうが入念に作成する必要があるが、後者は採点がマークシートリーダーなどを用いることで容易に採点処理ができるのでより実用的であり、利害関係が大きい(high stakes)テストであるセンター試験や入学試験から比較的に利害関係が小さい授業内テストまでよく用いられる。その中でも選択肢がいくつか設けられるMCがもっとも一般的に用いられている。MCにも多様な形式があり、どのMC項目形式が受験者には難しかったかとMC項目形式という要因がどの程度得点に影響するかを調査することはテスト開発者に重要な示唆を与えると考える。例えば、難しいと思われる項目形式を用いることでテストの難易度を上たり、項目形式が与える影響を最小限にすることでより妥当性が高いテストを開発できると考えられる。本研究では文法テストにある6種類のMC項目形式の困難度、MC項目形式が得点に与える影響の度合い、文法テストの信頼性を調査することを主な目的とする。

## 研究の背景

心理測定の分野ではMCを含む項目作成方法について書かれた文献がある(Haladyna, 1992, 2004; Haladyna, Downing, & Rodriguez, 2002等)。言語テストの分野でも、項目作成の際に避けるべき事項などの指針が記してある文献がある(Alderson, Clapham, & Wall, 1995; Bachman & Palmer, 1996; Brown, 2005等)。MC

に関する実証研究もあり、Shizuka, Takeuchi, Yashima, and Yoshizawa (2006) は、大学入学試験で多用される4択MCを3択MCと比較し、3択MCでも十分に機能すると示唆した。Dudley (2006) は語彙と読解としてのMCと多肢真意判定項目 (multiple true/false item) と比較した結果、多肢真意判定項目はMCとほぼ同様に機能すると示唆した。

項目形式が受験者の読解テストパフォーマンスに与える効果についての研究は Shohamy (1984)、Kobayashi (2002) がある。Shohamy (1984) はMCと自由回答式項目 (open-ended items) からなる読解テストを2000人の学習者に実施した結果、MCのほうが自由回答式項目より易しいという結果を得た。自由回答式項目の場合、英語を記述しないと正解にならないため、発表技能 (productive skill) が試されるので難しいとの理由である。また、英語力が高い受験者はさほど項目形式にパフォーマンスは影響されなく、逆に英語が低い受験者はより影響されると主張している。Kobayashi (2002) はテキストの種類と項目形式が読解テストのパフォーマンスに与える影響について調査した。項目形式は3種類あり、クローズ、自由回答式項目、要約ライティング (summary writing) で、信頼性は高かった。平均値をもとに項目形式の難易度を比較したところ、自由回答式項目、要約ライティング、クローズの順に難しかった。しかし、テキストの段落構成がなっていない場合、クローズが難しかった。教育的示唆としては、それぞれの項目形式は違う読解能力を測っているようで、自由回答式項目と要約ライティングはどちらかといえば大まかな内容を理解したかを測っているとのことであった。

聴解テストのパフォーマンスに与える影響に関しての研究として、Shohamy and Inbar (1991)、Yi'an (1998)、Brindley and Slatyer (2002) がある。Shohamy and Inbar (1991) では、テスト形式の他に、テキストの種類も一つの要因として調査している。結果として、ニュース放送が最も難しく、講義と会話は比較的易しいことがわかった。また、本文の内容にある事柄を理解したかを問う自由回答式項目は全体の内容を理解したかを問う項目より易しいことがわかった。Yi'an (1998) はMCからなる聴解テストを受験しているEFL学習者に回顧的言語化報告 (retrospective verbal report) を用いてテストを受けている際の頭脳処理を調査した。結果として、基本的には言語知識の処理を行っているが、言語知識を補うため非言語知識が使われ、MCは英語力が高い学習者が消去法などを駆使し正解できる場合があるので妥当性に欠けると主張した。Brindley and Slatyer (2002) は話す速度 (speech rate)、テキスト、項目形式が聴解タスクと項目困難度に与える影響について調査した。用いられた項目形式は空欄箇所補充項目 (sentence completion)、短答形式の項目 (short response)、表に必要情報を記入する表記式項目 (table answer) で、すべて応答構築式型の項目であった。結論として、会話速度を変えるなどして難しいと思われた項目は易しい場合があり、一概に特定の相が項目を難しくするというわけではなく、相の交互作用が項目を難しくすると主張している。

In'nami and Koizumi (2009) はMCと自由回答式項目が第一言語 (L1) の読解、第二言語 (L2) の読解、L2の聴解テストパフォーマンスに与える影響について言及した56もの研究を用いてメタ分析 (meta-analysis) を行っている。MCはL1読解とL2聴解テストにおいて自由回答式項目より易しいという結果になった。L2読解では特定の研究を用いて分析を行えば同様の結果が得られた。

語彙テストの研究として、Morimoto (2006) は3種類の40項目を含むMCからなるテスト形式 (test form) を英語力が同等だと思われる三つのグループに語彙知識を測る

目的として実施した。タイプAは英文中にある下線が引かれた英単語の意味にもっとも近い英単語を選ぶ形式で、文脈独立的 (context-independent) 項目で、文脈から答えを導き出せないものである。タイプBももっとも近い意味を選ぶ形式だが、文脈依存的 (context-dependent) 項目で、文脈から答えを導き出せるものである。タイプCは英文中にある空欄箇所にもっともあてはまる英単語を選ぶ形式で、文脈依存的項目である。三つのテスト形式の信頼性は高いといえる。平均値が最も高かったテスト形式はタイプBであった。つまり、文脈からより情報が得られるテスト形式の平均値が高かった。

文法テストの研究として、David (2007) は新たな文法項目形式であるマルチトラック (multitrack) を考案し、従来よく用いられる3種類の項目形式である文中にある空欄箇所補充MC (sentence-based)、クローズMC (text-based MC cloze)、文中にある空欄二箇所補充MC (double-blank sentence-based) の項目形式困難度を調査するため FACETS 分析を用いた。その結果、項目形式困難度の推定値のばらつき度合いを示す分離指数 (separation index) は13.00となり、そのばらつきが有意かを示すカイ二乗は有意な数値になり、マルチトラック、空欄二箇所補充MC、空欄箇所補充MC、クローズMCの順に困難度の推定値が高く、項目形式は文法テストのパフォーマンスに影響する要因であることがわかった。

### 研究の目的

テスト方法についての既存研究では予期する応答についての研究が主であり、選択形式応答と応答構築式型の項目の難易度など項目形式がパフォーマンスに与える効果について検証され、選択形式応答の項目形式のほうがより易しいとの結果となっている (In'nami & Koizumi, 2009; Shohamy, 1984)。また、MCの項目形式に関する研究は少数みられ、それぞれ難易度も違いがあることが指摘されている (David, 2007; Morimoto, 2006)。MCは授業内テストのような利害関係が小さいテストから入学試験のような利害関係が大きいものまで幅広く用いられる項目形式であり、テストのパフォーマンスに影響するので、さらにその相が与える影響について調査をする必要がある。しかし、その相が文法テストパフォーマンスに与える影響について言及した研究はDavid (2007) のみなので、さらなる調査が必要である。よって、本研究は以下の4点を研究目的とする。

1. 本研究で用いた文法MCはプレイズメントテスト目的と診断テスト目的として受験者にどの程度機能しているか。
2. 本研究で用いた文法テストにある6種類の項目形式はどの程度困難度の推定値には違いがあるか。
3. 受験者、項目、項目形式の相は文法テスト得点の分散に対してどの程度の割合であるか。
4. 本研究で用いた文法テストはどの程度信頼性があるか。

## 方法

### 対象者

本研究は関東地方にある私立大学で行われた。対象者はテストを受験する前に研究目的として結果を用いられることに同意した工学部に所属する2009年度入学した一年生で、608名の英語学習者である。対象者の中には中等教育で習った学習内容を完全に習熟していない学生がいるため、リメディアル英語教育にも力を入れている。よって、カリキュラムの主要な目標の一つは英語の基礎である文法知識を学習することである。

### 文法テスト

この文法テストの開発目的はプレイスメントテストと、到達度テスト (achievement test) としても実施できるテストを開発することである。よって、新学期がはじまる4月にプレイスメントテストとして実施し、テスト得点をもとにクラス分けを行い、一年後に同じテストを実施し、授業目的を学習したかを調査する目的で開発された。つまり、このテストはプレイスメントテストとしても用いられる目標標準準拠テスト (criterion-referenced test [CRT]) であるといえよう。やはり、プレイスメントテストの内容はカリキュラムの内容にある授業目的と合致したものがよいと考えた (Bachman, 1990; Brown, 1989)。

### 試行テスト

2008年12月に試行テストを行った。試行テストの作成手順は、まずリメディアル英語教育で使用する教科書にある文法項目や練習問題のみを、教科書の内容を補足する練習問題を作成し、練習帳として仕上げ冊子にした。扱った16の文法項目はbe動詞、一般動詞、未来形、助動詞、冠詞、代名詞、前置詞、接続詞、比較、進行形、to不定詞、動名詞、受動態、現在完了、関係詞、仮定法である。教科書と練習帳でよく用いられた項目形式は空欄箇所補充、英和訳、和英訳、並べ替え、MCなどである。試行テストでは教科書にある文法項目と項目形式を用いてテスト項目を作成することとした。セクションは四つあり、それぞれ項目形式は空欄箇所補充MC、英和訳MC、和英訳MC、間違い探しMCで、項目数はそれぞれ32、16、16、16問である。

試行テストは、工学部の9クラスと法学部の4クラスを履修した合計374名の2008年度入学した一年生に授業内で実施された。受験時間は一時間程度であった。文法テスト以外にも65問からなる語彙テストもあり、語彙から文法の順に受験したためか、受験者の中には時間が足りなかった人がいたとの報告もあった。しかし、項目についての良質なデータを得ることができた。分析の結果、平均点は80点満点中38点で、項目数が十分にあったためそれぞれのセクションの信頼性は高いほうであった ( $\alpha = .80, .74, .76, .70$ )。

### 本試験

2009年4月に本試験であるプレイスメントテストを実施した。そのテストを開発するため、試行テストの結果をもとにテスト項目を改良した。まず、項目容易度 (item facility)、項目弁別力 (item discrimination) など古典的テスト理論にもとづき項目

分析を行うコンピュータソフトであるITEMAN [version 3.6] (Assessment Systems Corporation, 1996)を用いた。項目の是非を決める基準は項目容易度が.70以下で項目弁別力が.20以上である。項目容易度が高いということはすでに受験者はその項目を学習済みであり、学習度合いを測るのに適していないといえる。また、プレースメントテストは得点のばらつきを生じさせたいので、項目弁別力が高いほうがよい。セクション1は16の文法項目につき2項目ずつ作成したので、二つの基準をより満たした項目を二つのうちから一つ選び、攪乱肢分析(distractor analysis)の結果なども参考にし、選択肢にある文言を変更するなどをして改良した。セクション2、3、4は試行テストでは16の文法項目について順に1項目ずつ作成した。練習帳では並べ替えの項目形式が用いられていたが、試行テストではなかった。新たに偶数の文法項目についての項目を8問作成し、セクション5とした。セクション6ではクローズMCを7問作成した。他のセクションにある項目数は8項目、または16項目なので、各セクションの項目数を均等にすべきだが、語彙テストが25問と文法テストが55問で、合計項目数を80問としたかったため、セクション6のみ7問とした。文法テスト以外に語彙テストとして25問出題されたが、本稿の主題に外れるので触れないこととする。

セクション1は空欄箇所補充MC ( $k = 16$ )で、文中の空欄にもっともあてはまる選択肢を一つ選ぶもので、例題は以下の通りである。

There ( ) a pen on the table.

- ① is    ② am    ③ are    ④ aren't

この項目はbe動詞についてで、正解は①である。セクション2は英和訳MC ( $k = 8$ )で、英文の意味にもっとも合う和文を選択肢①～④から一つ選ぶもので、例題は以下の通りである。

There isn't a pen on the table.

- ① ペンは机の上にある。  
 ② ペンは机の上にあった。  
 ③ ペンは机の上にはありませんか。  
 ④ ペンは机の上にはない。

これもbe動詞についての項目で、正解は④である。セクション3は和英訳MC ( $k = 8$ )で、和文の意味にもっとも合う文を選択肢①～④から一つ選ぶものであり、例題は以下の通りである。

彼は多くのレポートを書いた。

- ① He is writing many reports.  
 ② He is written many reports.  
 ③ He writes many reports.  
 ④ He wrote many reports.

これは一般動詞についての項目で、正解は④である。セクション4は間違い探しMC ( $k = 8$ )で、英文中にある文法的に間違いだと思ふ選択肢を①～④から一つ選ぶという形式で、例題は以下の通りである。

- ① My sons ② is ③ not at home ④ yesterday.

これはbe動詞についての項目で、正解は②である。セクション5は並べ替えMC ( $k = 8$ )で、文中にあるカッコ内の①～④を意味の通るよう文に並べ替え、3番目にくる選択肢を①④から一つ選ぶという形式であり、例題は以下の通りである。

Meg (① did ② cook ③ well ④ not).

これは一般動詞についての項目で、正解は②である。セクション6はクローズMC ( $k = 7$ )で、会話文中にある空欄箇所にあてはまる選択肢を①④から一つ選ぶというもので、例題は以下の通りである。

A: I went to Tokyo last week.

B: Really? How ( ) it?

A: It was nice.

① is ② are ③ was ④ were

これはbe動詞についての項目で、正解は③である。

### 分析の手順

まず、テスト結果を把握するためSPSS [version 17.0]を用いて記述統計と内部一貫性信頼性(internal consistency reliability)であるクロンバックアルファ係数( $\alpha$ 係数)を求めた。項目が受験者にプレイズメントテストとしてと、目標規準準拠テストとしても機能しているかを古典的テスト理論の観点からみるため、項目容易度、項目弁別力、 $B$ -指数( $B$ -index)をエクセルで算出した。 $B$ -指数とはいかに合格者と不合格者がある項目が区別しているかを示す指数である。

次に、FACETS [version 3.41] (Linacre, 2002)を用いて受験者能力値、項目困難度、項目形式困難度の三つの相の推定値を求めた。このソフトは項目応答理論の一種である多相ラッシュモデル(multifaceted Rasch model)にもとづいている(Linacre, 1989)。ラッシュモデル(Rasch model)とはある項目の項目困難度の推定値とある受験者の正解や不正解の応答にもとづき受験者能力値を推定する統計モデルである(Bond & Fox, 2001)。多相ラッシュモデルとは項目困難度や他の相の推定値とある受験者の得点をもとに受験者能力値を推定する統計モデルである(Linacre, 1989)。また、このFACETS分析はより受験者の能力や項目の困難度などに影響されない項目形式困難度の推定値を求めることができる。一般的にこのモデルは評定者が受験者の能力を採点するパフォーマンステストで用いられる(Coniam, 2008; Norris, Brown, Hudson, & Bonk, 2002; Schaefer, 2008等)。しかし、項目形式の困難度を求める目的でDavid (2007)もFACETS分析を行っている。

一般化可能性理論(generalizability theory [G theory])を用いて、テスト得点の分散に対して項目形式の相の分散の割合を一般化可能性研究(generalizability study [G研究])で推定し、古典的テスト理論では信頼性係数(reliability coefficient)に相当する一般化可能性係数(generalizability coefficient)と信頼度指数(dependability index)を決定研究(decision study [D研究])で求めた。一般化可能性理論に関してはBrennan (2001a)、ShavelsonとWebb (1991)が解説書を書いている。この理論は一般的には分散分析と古典的テスト理論を用いて算出する信頼性係数を合わせたようなものだといわれている(Gebril, 2009)。つまり、分散分析では各相の分散成分を推定し、その分散成分をもとに一般化可能性係数と信頼度指数を推定する。この理論もパフォ

一マンテストの分析によく用いられる (Lee & Kantor, 2007; Yamanishi, 2005; Xi, 2007 等)。Gebril (2009) はパフォーマンステストにおけるライティングタスクの違いによって分散が生じ、また一般化可能性係数に影響するとの結果を示している。また、MC型 TOEFL における受験者の第一言語の違いがどの程度テストの分散を生じさせているかを検証するため、Brown (1999) はこの理論を用いている。目標規標準準抛テストの信頼度を調査するため、この理論を用いた研究は Brown (1993)、Kumazawa (2009)、Kunnan (1992) がある。本研究での G 研究は受験者 ( $p$ )、項目形式 ( $f$ )、項目が項目形式に組み込まれており ( $i:f$ ) ネスト (nest) しているので三つの相を分析対象とする。よって、デザインは  $p \times X (i:f)$  となる。また、各項目形式に組み込まれている項目数が異なるためアンバランスデザインとなる。使用したソフトは mGENOVA [version 2.1] (Brennan, 2001b) である。D 研究のデザインは  $p \times X (I:F)$  で様々な項目数の組み合わせによる一般化可能性係数と信頼度指数の変動を検証するため行った。

## 結果

表1は記述統計の結果を記す。この文法テストは六つのセクションからなり、各セクションは異なる項目形式が用いられている。項目数はセクション1がもっとも多く、16個すべての文法項目についての内容を出題している。セクション2から5は8項目ずつで、セクション6のみ7項目である。最小値は全セクションとも0点で、最大値は満点である。もっとも平均値が高いのはセクション1の8.37で、これは主に項目数が多いためである。平均値が低いのはセクション6の2.37で、主に項目数が少ないためである。項目数がセクションごとに異なるため、平均値ではどのセクションが受験者にとって難しかったかなどを解釈するのは不可能である。標準偏差は得点のばらつきを示す。もっとも値が大きいのはセクション1で次にセクション5、セクション3の順である。 $\alpha$  係数はプレースメントテストとして機能しているかを判断するための重要な統計で、得点のばらつきなどに係数は左右される。セクション6の  $\alpha$  係数は.27と低く、測定誤差が大きいことを示す。これは項目数が少ないのが主な理由である。セクション4の  $\alpha$  係数も.36で、思わしくない。しかし、全体の  $\alpha$  係数は.84で、プレースメントテストとしては十分だと考える。また、平均値は中央値とほぼ同じの27.14で、歪度と尖度も誤差の範囲以内でほぼ標準分布をなしている。

表1. 記述統計の結果 (N = 608)

セクション	項目形式	k	最小	最大	M	SD	歪度	尖度	$\alpha$
1	空欄箇所補充	16	0	16	8.37	3.18	.13	-.58	.70
2	英和訳	8	0	8	5.17	1.48	-.60	.05	.41
3	和英訳	8	0	8	4.15	1.76	-.09	-.48	.46
4	間違い探し	8	0	8	3.13	1.59	.30	-.39	.36
5	並べ替え	8	0	8	3.95	1.86	-.11	-.57	.56
6	クローズ	7	0	7	2.37	1.41	.34	-.27	.27
合計		55	7	55	27.14	8.16	.17	-.55	.84

表2は項目分析の結果を記す。項目容易度はプレイスメントテストの場合、値が.30から.70の範囲にある項目がよいとされている(Brown, 2005)。17項目がその範囲以外で思わしくない値を示す。目標標準準拠テストで事前テストとして実施された場合、項目容易度は低いほうが望ましい。つまり、一定の値以上を示す項目は既に授業を受ける前に学習済みということで、学習する必要がない項目となり、機能していないとみなす。8項目が.70以上の値を示し、すでに学習済みであり機能していない。項目弁別力はいかにある項目が受験者の能力を弁別しているかを示し、プレイスメントテストの項目として機能しているかを判断するうえで重要な統計であり、.20以上を示す項目がよいとされる(Brown, 2005)。9項目の値が低く改善の余地がある。B-指数は合否分割点(cut-off score)を設定した目標標準準拠テストの項目がどの程度合格者と不合格者の能力を区別しているかを示す統計である。合否分割点を60点とした場合、指数が.20以下で機能していないと思われる項目は13項目あった。一方で、全体的にプレイスメントテストと事前診断テストとして機能している項目も多くあった。

図1はFACETS分析で求めた受験者、項目形式、項目の推定値をグラフ化したものである。グラフの左側から受験者、項目形式、項目の順に分布があり、グラフの上に行くほど能力値が高い、項目形式が難しい、項目が難しいということになる。受験者、項目形式、項目の分離指数はそれぞれ2.28、8.02、9.65で、ラッシュ信頼性係数は.84、.98、.99となった。つまり、難しい項目であれば難しく、簡単であれば簡単であるというばらつきを再生できる可能性が高いということを意味する。項目形式の分離指数が高いことから六つの項目形式には困難度のばらつきがあるということがわかる。また、このグラフからもわかるが、項目形式困難度の推定値にはばらつきがある。セクション6のクローズMCがもっとも難しいようである。また、その反対に簡単な項目形式はセクション2の英和訳MCであることがわかる。

FACETS分析を行い求めた項目困難度の推定値、誤差、インフィット平均二乗、アウトフィット平均二乗の結果が表2にある。項目困難度の推定値は-2.79から2.10で、広範囲の能力を測定できる項目がある点から、能力が多様な受験者が受けるプレイスメントテストとしては望ましい。項目1のbe動詞についての項目の困難度推定値は1.48で、これは難しい項目だといえる。be動詞は初期段階である中学一年に学習する項目で、易しいと思われる文法事項であるが、難しい項目となった。一方で項目17もbe動詞についてであるが、項目困難度の推定値は-2.79となり、易しい項目であるといえる。測定誤差は.19以下で誤差が少ない。インフィット平均二乗は実際のデータとモデルとの適合度(fit)を表し、.80から1.20以内であれば項目がモデルに適合しているといえる(Bond & Fox, 2001)。このテストにおいては全項目ともにモデルに適合している。アウトフィット平均二乗ははずれ値に敏感なので、設定した範囲から外れ、不適合(misfit)となった項目が7項目ある。これらの項目の項目弁別力をみると値が極端に高い、または低いことがわかる。どちらかといえば、アウトフィット平均二乗の値が1.20以上になり適合不足(underfit)と判断された5項目がより問題視すべきである。なぜならば文法知識が豊富な受験者がやさしい項目に不正解になったということを示すからである。

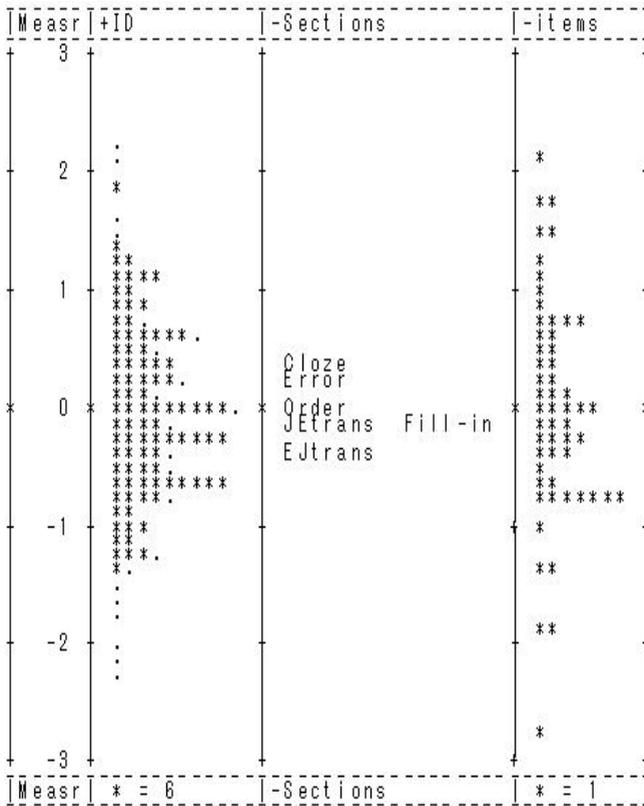


図1. グラフ化したFACETS分析の結果.

表2. 古典的テスト理論とFACETS分析を用いての項目分析の結果 (N = 608)

項目	文法項目	古典的テスト理論			FACETS分析			
		IF	ID	B-指数(.6)	Diff	SE	Infit MS	Outfit MS
1	be動詞	.22	.44	.38	1.48	.10	.90	.80
2	一般動詞	.64	.44	.34	-.61	.09	.90	.90
3	未来形	.29	.32	.31	1.08	.09	1.00	1.00
4	助動詞	.59	.43	.33	-.35	.09	.90	.90
5	冠詞	.67	.18	.16	-.75	.09	1.10	1.20
6	代名詞	.51	.33	.32	.03	.09	1.00	1.00
7	前置詞	.18	.36	.29	1.76	.11	.90	1.00

表2. 古典的テスト理論とFACETS分析を用いての項目分析の結果(N = 608)

項目	文法項目	古典的テスト理論			FACETS分析			
		IF	ID	B-指数(.6)	Diff	SE	Infit MS	Outfit MS
8	接続詞	.79	.40	.23	-1.41	.10	.90	.80
9	比較	.53	.39	.36	-.08	.09	1.00	1.00
10	進行形	.57	.45	.37	-.27	.09	.90	.90
11	to不定詞	.85	.34	.15	-1.83	.12	.90	.90
12	動名詞	.37	.37	.34	.69	.09	1.00	1.00
13	受動態	.59	.48	.42	-.34	.09	.90	.90
14	現在完了	.67	.33	.26	-.73	.09	1.00	1.00
15	関係詞	.43	.32	.34	.40	.09	1.00	1.00
16	仮定法	.48	.46	.40	.16	.09	.90	.90
17	be動詞	.95	.27	.07	-2.79	.19	.90	.70
18	未来形	.71	.40	.27	-.62	.09	.90	.90
19	冠詞	.18	.06	.05	2.10	.11	1.10	1.40
20	前置詞	.69	.38	.30	-.52	.09	1.00	.90
21	比較	.45	.34	.38	.61	.09	1.00	1.00
22	to不定詞	.88	.40	.14	-1.85	.13	.90	.70
23	受動態	.57	.17	.11	.09	.09	1.10	1.20
24	関係詞	.74	.37	.24	-.77	.10	.90	.90
25	一般動詞	.55	.52	.46	-.19	.09	.90	.80
26	助動詞	.59	.29	.26	-.37	.09	1.00	1.00
27	代名詞	.67	.39	.27	-.77	.09	.90	.90
28	接続詞	.73	.42	.31	-1.05	.10	.90	.90
29	進行形	.31	.19	.14	.97	.09	1.10	1.10
30	動名詞	.50	.32	.27	.05	.09	1.00	1.00
31	現在完了	.46	.34	.28	.23	.09	1.00	1.00
32	仮定法	.35	.20	.18	.74	.09	1.10	1.10
33	be動詞	.59	.37	.37	-.70	.09	1.00	.90
34	未来形	.61	.44	.37	-.80	.09	.90	.90
35	冠詞	.16	.05	.05	1.53	.11	1.10	1.60
36	前置詞	.20	.07	.04	1.27	.11	1.10	1.40
37	比較	.32	.24	.21	.57	.09	1.10	1.10
38	to不定詞	.47	.23	.21	-.16	.09	1.10	1.10
39	受動態	.42	.39	.40	.05	.09	1.00	1.00

表2. 古典的テスト理論とFACETS分析を用いての項目分析の結果(N = 608)

項目	文法項目	古典的テスト理論			FACETS分析			
		IF	ID	B指数(.6)	Diff	SE	Infit MS	Outfit MS
40	関係詞	.37	.35	.34	.28	.09	1.00	1.00
41	一般動詞	.77	.45	.30	-1.35	.10	.90	.80
42	助動詞	.40	.26	.19	.44	.09	1.10	1.10
43	代名詞	.49	.44	.42	.04	.09	.90	.90
44	接続詞	.18	.08	.10	1.71	.11	1.10	1.40
45	進行形	.66	.50	.33	-.77	.09	.90	.80
46	動名詞	.35	.30	.28	.70	.09	1.00	1.10
47	現在完了	.55	.40	.31	-.25	.09	1.00	.90
48	仮定法	.56	.39	.34	-.28	.09	1.00	.90
49	be動詞	.33	.28	.26	.41	.09	1.00	1.10
50	仮定法	.38	.27	.23	.15	.09	1.10	1.10
51	関係詞	.32	.31	.28	.46	.09	1.00	1.00
52	代名詞	.24	.23	.20	.90	.10	1.10	1.10
53	接続詞	.40	.16	.21	.05	.09	1.10	1.20
54	比較	.26	.01	-.06	.77	.10	1.20	1.30
55	to不定詞	.44	.25	.24	-.14	.09	1.10	1.10

注. IF = 項目容易度、ID = 項目弁別力、Diff = 項目困難度、SE = 標準測定誤差、Infit MS = インフィット平均二乗、Outfit MS = アウトフィット平均二乗

表3はFACETS分析を用いての項目形式困難度の推定値を示す。分離指数とラッシュ信頼性係数は8.02と.98で、高い値を示した。また、カイ二乗も有意な値を示し、項目形式困難度の推定値は有意なばらつきがみられた。平均値をみるだけでは項目数が違うなどの理由からどの項目形式が難しいかを解釈することはできないが、FACETS分析を用いるとすべての相の推定値が同じ間隔尺度上にあるので解釈が容易である。項目形式困難度の推定値をみると、クローズ、間違い探し、並べ替え、和英訳、空欄箇所補充、英和訳の順で難しい項目形式であるといえる。測定誤差も少なく、モデルの適合度もよい。

表3. 項目形式困難度の推定値

セクション	項目形式	セクション困難度	SE	Infit MS	Outfit MS
1	空欄補充	-.10	.02	1.00	1.00
2	英和訳	-.43	.04	1.00	1.00

表3. 項目形式困難度の推定値

セクション	項目形式	セクション困難度	SE	Infit MS	Outfit MS
3	和英訳	-.08	.03	1.00	1.00
4	間違い探し	.27	.03	1.00	1.00
5	並べ替え	-.03	.03	1.00	1.00
6	クローズ	.36	.03	1.00	1.00

注. SE = 標準測定誤差、Infit MS = インフィット平均二乗、Outfit MS = アウトフィット平均二乗

表4は $p \times (i:f)$ アンバランスデザインによるG研究の結果を記す。受験者の文法知識の差によって生じた分散成分の割合はテスト得点の分散成分を100%とした場合、7%であった。項目形式の違いによって生じた分散成分の割合は3%であった。項目によって生じた割合はより高く13%であった。受験者と項目形式の交互作用の分散成分の割合は1%であった。つまり、受験者によって得意不得意な項目形式が若干あり、テストパフォーマンスに影響したといえる。受験者と項目の交互作用の割合は76%で、受験者は項目によって異なるパフォーマンスをしたことになるが、これは受験者の特定の項目に対する得意不得意などから生じているのではないかと推測される。

表4.  $p \times (i:f)$ アンバランスデザインによるG研究の結果

相	分散成分推定値	分散成分%
受験者 ( $p$ )	.01789	7%
項目形式 ( $f$ )	.00655	3%
項目 ( $i:f$ )	.03186	13%
受験者 $\times$ 項目形式 ( $pf$ )	.00347	1%
受験者 $\times$ 項目 ( $p \times i:f$ )	.19202	76%
合計	.25179	100%

表5は $p \times (I:F)$ によるD研究の結果を表し、図2はその結果をグラフ化したものである。母得点(universe score [ $\sigma^2(\tau)$ ])は古典的テスト理論では真値(true score)に値するが、その分散は.01789であった。集団基準準拠テストに用いる誤差は相対誤差(relative error ( $\sigma^2(\delta)$ ))というが、その誤差は.00413であった。目標基準準拠テストに用いる誤差は絶対誤差(absolute error [ $\sigma^2(\Delta)$ ])というが、その誤差は.00593であった。母得点分散と相対誤差をもとに算出するのが一般化可能性係数(generalizability coefficient [ $\rho^2$ ])であり、集団基準準拠テストの信頼性の解釈の際に用いられ、その

係数は.81であった。母得点分散と絶対誤差をもとに求めるのが信頼度指数 (index of dependability [Φ]) であり、目標規準準拠テストの信頼度の解釈に用いられ、その指数は.75であった。六つのセクションとも4項目の場合、一般化可能性係数は.68で、信頼度指数は.63で、もとの係数と指数より大きく下回る。セクション1から8項目を削除し、セクション6に1項目追加し、全セクションともに8項目ずつにすると、係数は.80となり、もとの係数である.81とさほど変わらない。係数が.90に達するには144項目も必要になる。48項目からなるテストでは係数は.80になり、それ以降項目を追加しても係数には若干の変動しかみられない。

表5.  $p \times (I:F)$  デザインによるD研究の結果

D研究	空欄補充	英和訳	和英訳	間違い探し	並べ替え	クローズ	合計	$\rho^2$	$\Phi$
1	4	4	4	4	4	4	24	.68	.63
2	5	5	5	5	5	5	30	.72	.66
3	6	6	6	6	6	6	36	.75	.69
4	7	7	7	7	7	7	42	.78	.72
5	8	8	8	8	8	8	48	.80	.74
6	9	9	9	9	9	9	54	.82	.75
7	10	10	10	10	10	10	60	.83	.77
8	11	11	11	11	11	11	66	.84	.78
9	12	12	12	12	12	12	72	.85	.79
10	13	13	13	13	13	13	78	.85	.80
11	14	14	14	14	14	14	84	.86	.80
12	15	15	15	15	15	15	90	.87	.81
13	16	16	16	16	16	16	96	.87	.82
14	24	24	24	24	24	24	144	.90	.85
15	32	32	32	32	32	32	192	.92	.86

注.  $\rho^2$  = 一般化可能性係数、 $\Phi$  = 信頼度指数

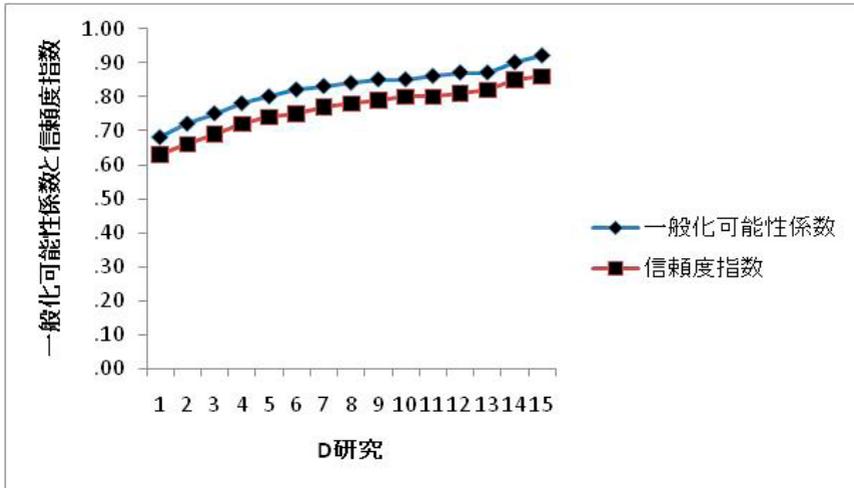


図2. グラフ化したD研究の結果

### 考察

本章では4点の研究目的について考察する。研究目的1は文法MCは受験者にどの程度機能しているかについてである。項目容易度の値から55項目中17項目がプレイスメントテストとしてはやさしすぎるまたは難しすぎるということとなった。しかし、プレイスメントテストにおいてより重視すべき統計は項目弁別力である。55項目中9項目のみ受験者の文法知識を弁別していないという結果になった。市販の集団基準標準拠テストの一種である英語運用能力テスト(English proficiency test)をプレイスメントテストとして用いた場合についての既存研究(Culligan & Gorsuch, 1999; Westrick, 2005)と比較しても、84%の項目が機能している本テストはかなり受験者に合った項目からテストが構成されているといえる。これはやはり受験者の英語力を想定し項目を作成し、一度試行テストとして実施し、その結果をもとに厳密に項目を選抜したからであろう。

この文法テストはプレイスメントテストとして実施したが、学習者がどの程度カリキュラム内容を学習しているかを測る目的でも実施しており、つまり診断テスト(diagnostic test)でもある。その目的を担う項目は項目容易度の値が高くなく、B-指数が高いものがよいが、その条件を満たしていない項目は項目11、17、22の3項目のみであった。これも試行テストの結果をもとに項目容易度が高い項目を削除したからであろう。いかに項目がカリキュラム内容の学習度合いを測定しているかを検証するには到達度テストとして事後テストを実施し、事後テストでの正答率から事前テストでの正答率の差を算出する差異指数(difference index)を求める必要がある。

項目困難度は文法項目に左右されると思うが必ずしもそうではなく、どちらかといえば項目の特性によるものという結果であった。例えば、be動詞に関する項目は項目1、17、33、49と4項目あるが、項目困難度の推定値はそれぞれ1.48、-2.79、

-.70、.41で、値は一定ではない。また、be動詞は文法項目の中でも易しい部類だと思うが、項目1の推定値は高く、難しい項目である。よって、同じ文法項目についての項目であっても攪乱肢の出来具合などの要因によって困難度は左右されるのではないか。今回、マークシートリーダーからは正解誤答を表すデータのみ入手できたので、攪乱肢がいかに機能しているかを調査することができなかったが、今後、受験者がある攪乱肢を選んだ%などを調査する目的で攪乱肢分析(distractor analysis)を行うことで、より攪乱肢の出来具合が項目困難度に与える影響がわかるであろう。

研究目的2は項目形式の困難度の推定値はどの程度違いがあるかであった。クローズMC、間違い探しMC、並べ替えMC、和英訳MC、空欄箇所補充MC、英和訳MCの順で難しい形式であるといえる。MCの項目形式は多種あり、既存研究(David, 2007; Morimoto, 2006)の結果と同様にそれぞれ困難度は異なるといえる。Morimoto (2006)によると語彙テストの項目は文脈がよりあり、前後の内容についての情報量がよりあったほうが正答率は上がると言っている。David (2007)の結果では文脈についての情報量がよりあるクローズMCがもっとも易しい形式であった。しかし、本研究ではクローズMCがもっとも難しい形式となった。理由として考えられるのは項目形式に不慣れであったという点である。間違い探しMCは過去TOEFLにあったが、日本の教育機関ではあまり使われない項目形式である。よって、この形式に慣れていない受験者がおり、困難度が上がったのではないか。逆に並べ替えMCは選択肢を文法的に正しい順に並べ替えないと正解を得られないので、より情報処理量が多いと思うが、この形式は入学試験でも用いられるので慣れていたためクローズMCより易しい形式になったのではないかと思われる。しかし、受験者が項目形式に慣れているかなどの練習効果(practice effect)が項目形式困難度に与える影響はさらなる調査が必要であろう。

研究目的3は受験者、項目、項目形式の相は文法テストの分散に対してどの程度の割合であるかである。まず、受験者の分散成分の割合は7%を占めた。つまり、プレイスメントテストを受験した学生の英語力には若干差があったといえる。これは55点中最小値は7点、最大値は48点であり、得点差が大幅にあることからいえる。次に、項目形式の違いによって生じた分散成分の割合は3%で、これは項目形式が6種類あり、それぞれの難易度が異なることから生じたといえる。項目の分散成分の割合は項目形式のものより大きくなった。これは、項目数は項目形式数より断然多く、また難易度の違いが項目は広範囲になっているため、よりテストパフォーマンスに影響したことを意味する。受験者と項目形式の交互作用は1%で、さほど大きくない。しかし、受験者と項目の交互作用は76%と大きい。これはある特定の相ではなく、いくつかの相の交互作用が大きくパフォーマンスに影響していると主張したBrindley and Slatyer(2002)を支持している。部分的測定テスト項目(discrete-point item)からなるテストを一般化可能性理論を用いて分析を行った既存研究の結果と比較すると、この文法テストでも受験者の分散成分の割合が少なく、項目と交互作用の割合が大きくなっており、同様の結果が得られた。Kumazawa(2009)では英語力がほぼ同じ学生に語彙診断テストとして実施したため、受験者の分散成分の割合が低く、代わりに項目の割合が大きくなってしまった。Brown(1999)の研究では英語運用力テストであるTOEFLの文法セクションを分析したが、やはり交互作用の割合が大きくなっていて、Zhang(2006)の研究はTOEICについてで、測定している技能が聴解と読解だが、交互作用の割合が大きくなっている。この交互作用の割合の大きさが意味することは、言語テストでは多様な要因がテストパフォーマンスに影響しているということで

あろう。よって、受験者などの他の相の分散成分の割合が比較的少なくなるのであろう。今後、より多くの相(文法項目、文中にある語数等)を分析の対象とし、それらがパフォーマンスに与える影響を調査する必要がある。

研究目的4は文法テストはどの程度信頼性があるかである。表1にあるセクションごとの $\alpha$ 係数には高低の差はあるが全体では.84で、比較的高いといえる。D研究ではセクションごとの項目数の設定を変えた場合の一般化可能性係数と信頼度指数の変動を検証した。実施時間が項目数に対して短いことを考えると全セクションを8項目で統一し、項目数を減らしたほうがよい。48項目とした場合、係数と指数はそれぞれ.81、.74となり、係数を.80以上に確保できる。この値は既存研究(Culligan & Gorsuch, 1999; Westrick, 2005)の結果より上回っている。もっとも妥当な項目数を決めるには信頼性と同時にテストの実用性(practicality)も考慮する必要がある。つまり、係数も確保でき、受験者の負担を軽減し、テスト実施時間内に終える項目数を出題すべきである。しかし、このテストはテスト開発の時間と労力はかかったが、マークシートリーダーで採点処理でき、教員が開発した学内テストなので、費用はさほどかからないため実用的ではある。

## 結論

本研究ではリメディアル教育の一環として行われる授業用の教科書と練習帳に準拠した目標規準準拠テストをプレイスメントテストとして実施した結果をもとにMCの機能、MCの項目形式の困難度、分散成分の割合、テストの信頼性について論じた。総括すると、この文法テストのほとんどの項目は機能していたこと、MCの項目形式には困難度のばらつきがあること、項目形式の分散成分が抽出され文法テストパフォーマンスに影響する要因であること、信頼性は比較的高かったことなどが主な結果として挙げられる。

教育的示唆としては、プレイスメントテストは集団基準準拠テストであり、事前診断テストのような目標規準準拠テストとは異なるが、その二つの実施目的を満たすテストを開発することは可能である点がいずれも挙げられる。この点についてはBrown(1989)でも支持する結果を示している。次に、MCの項目形式には多種あり、それぞれの困難度は異なるが、本研究での結果を踏まえ若干ではあるがテストの難易度を調節できる可能性が示された。例えば、文法テストの難易度をより上げたい場合、比較的易しい項目形式であった空欄箇所補充MCではなく、間違い探しMCを用いるということだ。最後に、カリキュラムに適したプレイスメントテストを実施するための最善策は独自にカリキュラム内容と受験者に合ったテストを開発することだと考える。その際、明確なテスト設計をし、試行テストを実施し、改良することは必須である。市販のテストを採用する場合でも必ず試行テストを行い、テストが受験者に適しているか検証することが推奨される。

本研究の欠点として挙げられるのは妥当性について具体的に言及していない点である。Kane(1992)が推奨する論証型アプローチ(argument-based approach)によると、妥当性を論証するには、観測(observation)、一般化(generalization)、外挿(extrapolation)、理論(theory-based inference)、決定(decision)について言及すべきだとある。本研究では、方法にある内容通りテスト開発過程、実施方法、実施目的などについて述べ、これは観測に相当するが、テスト細目(test specification)をもとに妥

当性論証に用いたほうがより適切ではある。一般化については、D研究で一般化可能性係数を求め、比較的值が高いことで論証できる。外挿とは文法テストにある項目が実際の授業で行う問題と類似しているかについてであるが、授業で使用する指定教材にある練習問題の種類などを調査したうえでその形式と同じ、または似たテスト項目を作成したことを一つの論証として用いることができる。理論とは文法テストが文法知識を測定していると思われる推論が妥当かということであるが、この点についてはさらなる統計分析を行い、それぞれの項目形式がどのような文法知識を測定していると思われるかの構想概念について調査する必要がある。決定とは文法テスト得点をもとに決定した事柄が妥当であったかということである。このテストはプレースメントテストと診断テストとして実施された。プレースメントテスト得点をもとにリメディアル教育を受けるかを判断することが妥当なのか、さらなる検証が必要である。また、プレースメントテスト得点でリメディアル教育を受ける必要があると判断された場合、それが学習者の英語学習にどのような影響を与えるかの波及効果も調査する必要がある。

### 謝辞

本研究は、関東学院大学授与学習支援GPプロジェクト(「校訓に基づく入学前～卒業後までの総合支援」)の一つである「リメディアル用教材開発」の一環として実施されたテスト結果を報告したものである。貴重なコメントをくださった工学部と法学部の先生方、及び査読者の方々に感謝いたします。

### 参考文献

- Alderson, C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Assessment Systems Corporation. (1996). *ITEMAN* (version 3.6) [computer software]. St. Paul, MN: Assessment Systems Corporation.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Brennan, R. L. (2001a). *Generalizability theory*. New York: Springer.
- Brennan, R. L. (2001b). *mGENOVA* (version 2.1) [computer software]. Iowa City, IA: The American College Testing Program.
- Brindley, G., & Slatyer, H. (2002). Exploring task difficulty in ESL listening assessment. *Language Testing*, 19, 369-394.

- Brown, J. D. (1989). Improving ESL placement tests using two perspectives. *TESOL Quarterly*, 23, 65-83.
- Brown, J. D. (1993). A comprehensive criterion-referenced language testing project. In D. Douglas & C. Chapelle (Eds.), *A new decade of language testing: Collaboration and cooperation* (pp. 163-184). Ann Arbor, MI: University of Michigan Press.
- Brown, J. D. (1999). Relative importance of persons, items, subtests and languages to TOEFL test variance. *Language Testing*, 16, 216-237.
- Brown, J. D. (2005). *Testing in language programs: A comprehensive guide to English language assessment*. New York: McGraw-Hill College Press.
- Coniam, D. (2008). An investigation into the effect of raw scores in determining grades in a public examination of writing. *JALT Journal*, 30, 69-84.
- Culligan, B., & Gorsuch, G. (1999). Using a commercially produced proficiency test in a one-year core EFL curriculum in Japan for placement purposes. *JALT Journal*, 21, 7-25.
- David, G. (2007). Investigating the performance of alternative types of grammar items. *Language Testing*, 24, 65-97.
- Dudley, A. (2006). Multiple dichotomous-scored items in second language testing: Investigating the multiple true-false item type under norm-referenced conditions. *Language Testing*, 23, 198-228.
- Gebriel, A. (2009). Score generalizability of academic writing tasks: Does one test method fit it all? *Language Testing*, 26, 507-531.
- Haladyna, T. (1992). The effectiveness of several multiple-choice formats. *Applied Measurement in Education*, 51, 73-88.
- Haladyna, T. (2004). *Developing and validating multiple-choice test items* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Haladyna, T., Downing, S., & Rodriguez, M. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, 15, 309-334.
- In'nami, Y., & Koizumi, R. (2009). A meta-analysis of test format effects on reading and listening test performance: Focus on multiple-choice and open-ended formats. *Language Testing*, 26, 219-244.
- Kane, M. (1992). An argument-based approach to validity. *Psychological Bulletin*, 112, 527-535.
- Kobayashi, M. (2002). Method effects on reading comprehension test performance: Text organization and response format. *Language Testing*, 19, 193-220.

- Kumazawa, T. (2009). Revision of a criterion-referenced vocabulary test using generalizability theory. *JALT Journal*, 31, 81-100.
- Kunnan, A. J. (1992). An investigation of a criterion-referenced test using G-theory, and factor and cluster analyses. *Language Testing*, 9, 30-49.
- Lee, Y., & Kantor, R. (2007). Evaluating prototype tasks and alternative rating schemes for new ESL writing test through G-theory. *International Journal of Testing*, 7, 353-385.
- Linacre, J. M. (1989). *Many-facet Rasch measurement*. Chicago: MESA.
- Linacre, J. M. (2002). *FACETS* (Version 3.41) [computer software]. Chicago: MESA.
- Morimoto, Y. (2006). Comparison between matching- and supplying-format in multiple choice vocabulary tests. *JLTA Journal*, 9, 73-85.
- Norris, J., Brown, J. D., Hudson, T., & Bonk, W. (2002). Examinee abilities and task difficulty in task-based second language performance assessment. *Language Testing*, 19, 395-418.
- Schaefer, E. (2008). Rater bias patterns in an EFL writing assessment. *Language Testing*, 25, 465-493.
- Shavelson, R., & Webb, N. (1991). *Generalizability theory: A primer*. Thousand Oaks, CA: Sage.
- Shizuka, T., Takeuchi, O., Yashima, T., & Yoshizawa, K. (2006). A comparison of three- and four-option English tests for university entrance selection purposes in Japan. *Language Testing*, 23, 35-57.
- Shohamy, E. (1984). Does the testing method make a difference? The case of reading comprehension. *Language Testing*, 1, 147-170.
- Shohamy, E., & Inbar, O. (1991). Validation of listening comprehension tests: The effect of text and question type. *Language Testing*, 8, 23-40.
- Westrick, P. (2005). Score reliability and placement testing. *JALT Journal*, 27, 71-94.
- Xi, X. (2007). Evaluating analytic scoring for the TOEFL Academic Speaking Test (TAST) for operational use. *Language Testing*, 24, 251-286.
- Yamanishi, H. (2005). 一般化可能性理論を用いた高校生の自由英作文評価の検討. (Using generalizability theory in the evaluation of L2 writing) *JALT Journal*, 27, 169-185.
- Yi'an, W. (1998). What do tests of listening comprehension test?-A retrospection study of EFL test-takers performing a multiple-choice task. *Language Testing*, 15, 21-44.
- Zhang, S. (2006). Investigating the relative effects of persons, items, sections, and languages on TOEIC score dependability. *Language Testing*, 23, 351-369.

# Perspectives

## Reconsidering the Effectiveness and Suitability of PPP and TBLT in the Japanese EFL Classroom

Rintaro Sato (佐藤臨太郎)

*Nara University of Education*

With the recent high-profile utilization of tasks in language classrooms, the long-established traditional teaching methodology based on the presentation-practice-production (PPP) model is now being replaced by task-based language teaching (TBLT) in SLA (e.g., Skehan, 1996, 1998; White, 1988; Willis, 1996, 2004). However, in the Japanese EFL learning environment, there is still considerable skepticism regarding the effectiveness of TBLT. This paper explores the suitability of TBLT and PPP in the Japanese secondary school context and discusses the relative effectiveness of PPP from the point of view of skill acquisition theory. Some suggestions for effective teaching procedures are also discussed.

実践的コミュニケーション能力育成の重要性が高まるなか、TBLT (task-based language teaching) の効果が最近日本の英語教育現場において注目されている。一方、伝統的な PPP (Presentation-Practice-Production) モデルが、第2言語習得研究において否定されつつあるが、日本の外国語としての英語学習状況においてはTBLTに対して懐疑的な見方があるのも事実である。本稿では、PPPの効果について、日本の英語学習環境やskill acquisition theoryの観点から分析し、TBLT の効果について、その日本の教室現場での活用における問題点を指摘する。さらにPPPとTBLTを比較検討することにより、日本の英語教育においてのより効果的なアプローチについて提案する。

The Ministry of Education, Culture, Sports, Science, and Technology (MEXT) has declared that one of the overall objectives of English education in junior high school is to develop basic communication abilities (MEXT, 2008), and furthermore to develop communication abilities to convey information and express ideas and opinions in senior high school (MEXT, 2009). MEXT (2008) clearly states that “students should be engaged in activities that will lead them to exchange their thoughts and feelings by actually using the English language” (p. 6). The utilization of task-based language teaching (TBLT), which is a logical development of communicative language teaching (Willis, 1996), has been recently gaining attention for English teaching in Japan (e.g., Matsumura, 2009; Takashima, 2000, 2005).

Ellis (2003) defines a task as follows:

1. A task is a work plan.
2. A task involves a primary focus on meaning.
3. A task involves real-world processes of language use.
4. A task can involve any of the four language skills.
5. A task engages the cognitive process.
6. A task has a clearly defined communicative outcome. (pp. 9-10)

In the presentation-practice-production (PPP) model, practice of specified target structures in production has a crucial role. The teacher starts by explaining some specific new forms and meanings of the L2. Skehan (1998) summarizes PPP as follows: “The first stage is generally focused on a single point of grammar which is presented explicitly or implicitly to maximize the chances that the underlying rule will be understood and internalized. This would essentially aim at the development of declarative knowledge” (p. 9). Then, learners move on to a practice stage, focusing mainly on accuracy, subject to the teacher’s careful supervision or control. In the practice stage, which is aimed at converting declarative knowledge to procedural knowledge, control is gradually loosened as learners move to the next stage, production, at which point they are provided with opportunities to produce the target form, sometimes through communicative activities. In this production stage, Skehan notes that “learners would be required to produce language more spontaneously, based on meanings the learner himself or herself would want to express” (p. 93).

Proponents of TBLT dismiss the traditional PPP approach (e.g., Skehan, 1996, 1998; White, 1988; Willis, 1996, 2004). Skehan (1996) claims that “the belief that a precise focus on a particular form leads to learning and au-

tomatization no longer carries much credibility in linguistics or psychology” (p. 18). Willis (1996) also asserts that “language learning rarely happens in an additive fashion” (p. 135). Skehan (1998) states that “such an approach [i.e., PPP] is now out of fashion” (p. 94) and White (1988) discredits the PPP approach as a meaning-impoverished methodology.

However, the effectiveness of TBLT, especially in teaching grammar, can be questioned: TBLT may not be effective in teaching pre-specified target structures; it is not designed for examinations; and the Japanese language is the primary medium in a Japanese classroom. Dismissing PPP completely from our classrooms may be premature since the effectiveness of this approach has not fully been tested by SLA researchers (DeKeyser, 1998). Taking account of the Japanese EFL situation in which students do not have much exposure to English and have little need for communication in English in their daily lives, it is crucial to reconsider the effects of the traditional PPP approach and compare them with those of TBLT.

## **Reconsidering the Suitability of TBLT in the Japanese EFL Classroom**

### ***Target Grammatical Structures***

In teaching English in Japanese secondary schools, teachers have to use textbooks authorized by the Japanese government. These textbooks require students to learn a target grammatical structure in each section. It may be true that language learning does not occur in a linear, additive way (Willis, 2004), but pedagogical activities in the textbooks are systematically arranged in a way that requires students to learn the target language in a step-by-step way. In a typical task activity, in which students focus on meaning or communication to complete the task, the target grammatical structure cannot always be used by the students.

To examine whether students actually use a target structure or not and how they feel about a given task, I conducted a very simple experiment with 21 university students majoring in English education at a national university in Japan. In the task, extracted from Takashima (2005), students were put in a situation where they were implicitly encouraged to use the present perfect. The following was the procedure:

1. Students made pairs (one student was paired with me).
2. Each of the students was given a sheet which included different information from their partner’s sheet.

3. Referring to the sheets which had (1) information on what they had eaten before the Golden Week (GW) holidays, (2) a restaurant recommended by their mother, and (3) plans after GW, they were required to talk with their partner to decide which restaurant to go to during the holiday.
4. One student from each pair reported their final decision and the reasons why they chose a particular restaurant.
5. Students then completed a brief survey which asked whether they had used the target present perfect grammar during the task, and they were asked to write comments about the task.

This brief experiment showed that 15 students did not produce sentences with the target grammar at all. In my particular pairing with a student, I found that my partner had no immediate need to use the target grammar to complete the goal of the task. Some of the comments written by the students further highlighted some of the issues with the task (comments were originally written in Japanese and translated into English by the author):

- Student A: *I really enjoyed doing the task. It was like playing a game.*
- Student B: *I used the past and future tense but I did not have to use the present perfect to complete the task.*
- Student C: *When I couldn't say in English what I really wanted to mean, I spoke in Japanese.*
- Student D: *I don't know if I learned something in the activity, but anyway it was fun.*

As the task allows learners to choose the language needed to achieve the outcome of the task (Ellis, 2003) and learners are given the freedom to decide which grammatical items to use, the unlikelihood that learners will produce the target items in a task is quite understandable. It is crucial for Japanese learners, most of whom have exposure to English only in an English class, to learn new items during the class. However, the task may not meet this need. Ellis (2003) admitted that students often regard communicative tasks as opportunities for communication rather than learning. As an advantage of TBLT, Willis (1996) stated that “the role of tasks is to encourage learners to activate and use whatever language they already have, both for comprehension and for speaking and writing” (p. 147), and “it provides learners with the motivation to improve and build on whatever

language they already have” (p. 1). This statement is consistent with Swain’s claim (2005) that TBLT is suitable for advanced learners. University students engaged in the experiment above can be regarded as relatively higher level students as they had obtained high scores in the entrance exams for the national university and were majoring in English, but the fact that the task was not effective even for those higher level learners—in that most of them did not have to use the target grammatical structure—implies that it might be even less effective for junior and senior high school students. If new language items cannot be learned purposefully through a task, and if we refer to Bruton’s (2005) conclusion that a task has limited applicability for EFL students, we can be skeptical of the appropriateness of task-based approaches for Japanese students. Takashima (2005) contended that many Japanese EFL learners might not be able to adequately use target vocabulary and grammatical structures in the task context of TBLT, and he introduced focused-task activities that he believed are more applicable in the Japanese secondary school learning environment.

### **Examinations and Tests**

Yashima (2000) pointed out that Japanese learners have dual goals, namely, a practical realistic goal related to tests, and a goal related to using English for communication; and these learners may attach a greater or lesser degree of importance to each of these goals. It seems that most Japanese students have test-related motivation rather than communication-related motivation (Yashima, Zenuck-Nishide, & Shimizu, 2004). It cannot be denied that most Japanese secondary students study English for tests that mostly measure accurate knowledge of English rather than communicative language ability. Willis and Willis (2007) stated that task-based teaching is not designed with examinations in mind, and that it is designed to produce learners who can use their English outside the classroom, even if they make grammatical mistakes or errors. Needless to say, examinations have been a key factor in Japanese secondary school students’ English learning and have had significant influence. Japanese students are still being given traditional high-stakes tests such as end-term tests or tests for entrance examinations, which include reading and sometimes writing and even listening, but not speaking. There is therefore a mismatch between examinations and task-based language teaching.

## English Classes Conducted in Japanese

In the TBLT classroom, “generally it is best to do all the classroom organization and instruction giving in English, as this creates a very real context and purpose for listening” (Willis & Willis, 2007, p. 220). The concern in Japanese secondary schools is that English is not the primary medium in the Japanese English classroom. According to the results of a survey conducted by MEXT, less than one-third of teachers in junior high school and less than 10% of senior high teachers conduct their classes mainly in English (Kan, 2006). This long-standing issue of the English class dominated by the Japanese language should certainly be addressed. However, as long as Japanese teachers of English use Japanese for communication, students will follow their teachers and overuse their shared mother tongue in pair or group work. A precondition for TBLT is the establishing of the English-for-communication class. Facing the fact that Japanese is the primary language in the English classroom, the introduction of TBLT may be premature, notwithstanding the stated MEXT objectives.

## Reconsidering the Utilization of PPP

### *Declarative Knowledge and Procedural Knowledge*

In the process of second language learning, learners learn or acquire two different types of knowledge, namely, declarative knowledge and procedural knowledge. Declarative knowledge is factual knowledge that is expressed explicitly, for example, knowing that the capital of Japan is Tokyo and its former name was Edo, or knowing that English has two articles: the definite article *the* and the indefinite article *a* or *an*. In contrast to this type of knowledge, procedural knowledge can only be performed. Knowing how to use the English articles *the* and *a*, or knowing how to swim or ride a bicycle are examples of procedural knowledge. In the input-scarce environment of Japan, students first acquire declarative knowledge in a formal class and then this knowledge develops into procedural knowledge through practice or exposure (Sharwood Smith, 1981). Anderson (1992) tells us that procedural knowledge is declarative knowledge that has been fully automatized, leading us to conclude that acquisition of procedural knowledge should be a final goal of second language learning.

### *Skill Acquisition Theory*

In his skill acquisition theory, Anderson (1993, 1995) claimed that second language learning starts out in declarative form, progresses to the stage of

proceduralization through extensive practice, and then knowledge becomes automatic. Anderson (1995) divided the process of skill acquisition into three stages, namely, a cognitive, an associative, and an autonomous stage. He noted that declarative knowledge corresponds to the cognitive stage, proceduralization of knowledge corresponds to the associative stage, and automatizing procedural knowledge is on a parallel with the autonomous stage. In the cognitive stage, students learn a set of facts relevant to the skill (declarative knowledge). In the associative stage, they strengthen the connections among the elements needed for successful performance, detecting and eliminating errors, and converting declarative knowledge into a procedural form. In the autonomous stage, students continue developing procedural knowledge, and perform the skill better and more automatically. Anderson (1993) conceded that not all knowledge starts with the declarative form, developing into the procedural, and that acquisition of procedural knowledge does not always mean the loss of declarative knowledge. However, the important point in the theory is that learners initially learn declarative knowledge and then develop it into procedural knowledge through engaging in the target behavior.

### ***PPP and Skill Acquisition Theory***

The effectiveness of the PPP approach can be examined from the point of view of skill acquisition theory. According to DeKeyser (1998), the theory implies that learners should be given explicit teaching of the target grammar first (cognitive stage), followed by activities or practice to develop their acquired or learned declarative knowledge into procedural knowledge (associative stage), and then less focused communicative activities to enhance proceduralization and automatization (autonomous stage). In a typical PPP lesson, a target language item is introduced by the teacher to clarify its meaning in the presentation stage, which corresponds to the cognitive stage in skill acquisition theory. In the practice stage, students repeat and practice target items or sentences with activities such as pattern practice, drills, and answering questions using a specified form. This stage, whose activities seem to be completely dismissed in TBLT, can be regarded as the associative stage in skill acquisition theory. The production stage, in which students are expected to produce language items they have just learned with other previously learned languages, can be compared to the autonomous stage. Byrne (1986) stated that the practice stage in PPP roughly corresponds to Anderson's proceduralization stage, and the production stage corresponds to automatization. Supporting Anderson's theory (1993), Yamaoka (2005,

2006) argued that imitation, repetition, and pattern practice are essential for the development of declarative knowledge into procedural knowledge in the Japanese EFL environment. DeKeyser (2001) concluded that it is desirable that activities be introduced after a rule-based model in which language is learned through a formula, because with explicit teaching students can notice the new structure and process its form-meaning link so that they can finally acquire it. It can safely be said that the PPP approach is compatible with skill acquisition theory in that the approach can enhance the transition of students' knowledge from declarative to procedural.

### **Suggestion**

It is definitely important for Japanese students, who rarely have opportunities to use English for communication outside the classroom, to learn, acquire, and use target structures within the class period. Anderson's skill acquisition theory (1993, 1995) claimed that practice or focused tasks have a role to play in learning, contributing to the development of declarative knowledge into procedural knowledge. With that as a reference, it seems that the traditional PPP type approach is still practical and beneficial for learning in Japanese secondary schools. Students are required to learn target structures effectively and efficiently within a limited amount of time.

However, there are some considerations to bear in mind when creating and conducting English classes based on the PPP approach. Yamaoka (2005) warned that learners should experience meaning-form connections during practice in order to develop procedural knowledge. He dismissed the simplistic, mechanical imitation and repetition often seen in the traditional PPP approach and advocated complex, active imitation and repetition, connecting form with meaning in language learning. He also stated that instead of simple, mechanical oral pattern practice, cognitive practice is required to establish form-meaning connections. Similarly DeKeyser (1998) claimed that mechanical drills or rushed output are far from ideal for the development of declarative and procedural knowledge. In skill-based language teaching, activities or practice are carried out after declarative explanation of target items, which is consistent with the PPP approach. In the practice stage, bearing in mind the disadvantages of disconnecting form and meaning (namely that it results in parrot-like repetitive practice or mechanical drills [see DeKeyser, 1998; Yamaoka, 2005, 2006]), a large amount of practice in context is required so that declarative knowledge can develop into procedural knowledge. A study by Gathbonton and Segalowitz (1988) concluded that communicative activities should be devised to allow learners to use the same expressions and formulas

repeatedly. Arevart and Nation (1991) introduced an activity in which learners tell the same story several times, with increasingly strict time limitations, and they noted the positive effect on acquisition of having the learners use the same vocabulary items and formulas many times in a row. It is desirable to provide practice which allows learners to use the target rules or expressions repeatedly in a context where a connection between form and meaning can be established (DeKeyser, 2001).

In the last production stage, more open activities and tasks focusing primarily on meaning that are not designed for the use of a specific form, such as opinion gap tasks, can be used. However, we could argue that focused activities, which intrinsically require learners to use the target items repeatedly, can still be effectively employed. Presumably, activities which give priority to producing the target structure accurately can be criticized from the TBLT point of view. Willis (1996), for example, argues that students sometimes overuse the target structures and that they are still “in practice mode” (p. 134) trying to show control of the structure rather than expressing their own meanings. However, considering that learners should be encouraged to learn the new grammatical structures rather than simply using structures they have already learned and internalized, the use of focused practice in the production stage is rational.

In the Japanese EFL context, where there is little or no practical need to use English outside the classroom, not a few students study English as a school subject rather than as a practical language, and their aim is just to obtain high scores in high-stakes tests. In competitive examination systems, accuracy-focused written tests are seen as a key factor affecting teaching and learning. There seems to be a mismatch between this situation and the kind of speaking-oriented communicative activities carried out in TBLT. It is therefore vital for reading, writing, and listening activities to be implemented even at the production stage. An example of a technique that can meet this requirement is the dictogloss devised by Wajnryb (1990). In this procedure, students listen to a text two or three times and then reconstruct it. The text is designed to draw students' attention to a specific target grammatical feature. This kind of focused task involving listening and writing can integrate production and the requirements of written tests.

## Conclusion

It is true that the task-based approach, which developed out of communicative language teaching, continues to attract keen attention. However,

as this approach does not take sufficient account of the particular English learning environment in Japan, it may not be as practical in its application as the PPP approach. One serious problem of the PPP approach in the Japanese classroom in general is that the last P (production) stage is not given enough time in order for students to improve their communication (e.g., Izumi, 2009; Takashima, personal communication, December 13, 2009). Additionally, if teachers put too much emphasis only on mechanical activities in the second P (practice) stage without context, it is likely this will not lead to learning (Yamaoka, 2005, 2006). Lucas (1984) has pointed out that use of teaching methods relying heavily on tightly controlled drills and exercises has led to the poor English speaking ability of Japanese students. With these criticisms in mind, some revisions to the traditional PPP approach are obviously needed. However, with these revisions, it may still be the most suitable approach here in Japan.

The effect of TBLT, however, should not be dismissed since it can improve learners' motivation and help develop true fluency in an L2 (DeKeyser, 1998) by putting students in a situation where they can use English for real communication. By being provided with opportunities to actually use the language, students are motivated to talk in English (e.g., Nakahira, Yashima, & Maekawa, 2010). Tasks also can activate the atmosphere of the English classroom, improving students' positive attitude for communication (Matsumura, 2009). With PPP used as the primary approach, tasks can be effectively used in the third stage, the production stage.

As Ellis (2006) mentions, there is not just one effective approach to teaching grammar: the acquisition or learning of grammar of a second or a foreign language is a complex process. We should recognize "what options are available, what the theoretical rationales for these options are, and what the problems are with these rationales" (p. 103). I hope this paper will provide a good opportunity for teachers to reconsider their own teaching in order to conduct more effective English classes in the Japanese secondary school context. Rather than considering the respective approaches as an unhealthy dichotomy, we need to make room for some kind of combination of methods, taking from each to better serve the contextual realities. However, given the particular language teaching realities in Japan, especially at the high school level, there is still a very strong case to be made for the effectiveness of the PPP approach.

*Rintaro Sato* is an Associate Professor in the Department of English Education of Nara University of Education. His research interests include intake and output processing, feedback, and negotiation of meaning.

## References

- Anderson, J. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum.
- Anderson, J. (1995). Developing expertise. In J. R. Anderson (Ed.), *Cognitive psychology and its implications* (4th ed.) (pp. 272-304). New York: W. H. Freeman and Company.
- Arevart, S. & Nation, P. (1991). Fluency improvement in a second language. *RELC Journal*, 22, 84-94.
- Bruton, A. (2005). Task-based language teaching for the state of secondary FL classroom. *Language Learning Journal*, 31, 55-68.
- Byrne, D. (1986). *Teaching oral English* (2nd ed.). Harlow, UK: Longman.
- DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42-63). Cambridge: Cambridge University Press.
- DeKeyser, R. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125-151). New York: Cambridge University Press.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Ellis, R. (2006). Current issues in the teaching of grammar: An SLA perspective. *TESOL Quarterly*, 40, 80-107.
- Gatbonton, E. & Segalowitz, N. (1988). Creative automatization: Principles for promoting fluency within a communicative framework. *TESOL Quarterly*, 22, 473-492.
- Izumi, S. (2009). *Focus on form wo toriireta atarashii eigo kyouiku* [New English education which includes focus on form]. Tokyo: Taisyuukan Shoten.
- Kan, M. (2006). *Nihon no eigo kyouiku no genjo to kadai* [The actual situation of English education in Japan and problems]. In Y. Otsu (Ed.), *Nihon no eigo kyouiku ni hitsuyou na koto* [Things required in English education in Japan] (pp. 171-182). Tokyo: Keiogijuku Daigaku Shuppankai.
- Lucas, J. (1984). Communication apprehension in the ESL classroom: Getting our students to talk. *Foreign Language Annals*, 17, 593-598.
- Matsumura, M. (2009). *Eigo kyouiku wo shiru 58 no kagi* [58 key factors for English education]. Tokyo: Taisyuukan Shoten.
- MEXT. (2008). *Course of study for junior high schools foreign languages (English)*. Tokyo: Tokyo Shoseki.
- MEXT. (2009). *Koutougakko gakushu shido yoryo* [Course of study for senior high schools]. Kyoto: Higashiyama Shobo.

- Nakahira, S., Yashima, T., & Maekawa, T. (2010). Relationships among motivation, psychological needs, FL WTC, and Can-Do statements of English language learning based on Self-Determination Theory: Preliminary study of non-English-major junior college students in Japan. *JACET Kansai Journal*, 12, 44-55.
- Sharwood Smith, M. (1981). Consciousness-raising and the second language learner. *Applied Linguistics*, 2, 159-169.
- Skehan, P. (1996). Second language acquisition research and task-based instruction. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp.17-30). Oxford: Heinemann.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Swain, M. (2005). Legislation by hypothesis: The case of task-based instruction. *Applied Linguistics*, 26, 376-401.
- Takashima, H. (2000). *Jissenteki communication noryoku no tame no eigo tasuku katsudo to bunpou shidou* [English task activities and grammar teaching for practical communication abilities]. Tokyo: Taisyuukan Shoten.
- Takashima, H. (2005). *Eigo no tasuku katsudo to tasuku*. [English tasks activities and tasks]. Tokyo: Taisyuukan Shoten.
- Wajnryb, R. (1990). *Grammar dictation*. Oxford: Oxford University Press.
- White, R. (1988). *The ELT curriculum*. Oxford: Blackwell.
- Willis, D., & Willis, J. (2007). *Doing task-based teaching*. Oxford: Oxford University Press.
- Willis, J. (1996). *A framework for task-based learning*. Harlow, UK: Longman.
- Willis, J. (2004). Perspectives on task-based instruction: Understanding our practices, acknowledging our different practitioners. In B. L. Leaver & J. Willis (Eds.), *Task-based instruction in foreign language education* (pp. 3-44). Washington, DC: Georgetown University Press.
- Yamaoka, T. (2005). From item learning to category learning: A learning process of procedural knowledge of learning. *ARELE*, 16, 21-30.
- Yamaoka, T. (2006). On the importance of imitation and repetition in foreign language learning. *ARELE*, 17, 1-10.
- Yashima, T. (2000). Orientations and motivation in foreign language-learning: A study of Japanese college students. *JACET Bulletin*, 31, 121-133.
- Yashima, T., Zenuk-Nishide, L., & Shimizu, K. (2004). The influence of attitudes and affect on willingness to communicate and second language communication. *Language Learning*, 54, 119-152.

# Reviews

***The Handbook of Language Teaching*. Michael H. Long and Catherine J. Doughty (Eds.). Oxford: Wiley-Blackwell, 2009. xix + 801 pp.**

*Reviewed by*

Thomas Amundrud

Ritsumeikan University, Kyoto

Every year, volumes are published that presume to be the sort that researcher-practitioners will regularly consult for years to come. To this end, *The Handbook of Language Teaching* largely succeeds. Editors Michael Long and Catherine Doughty, whose previous *Handbook of Second Language Acquisition* (2003) is already a standard in the field, have managed to cover a wide range of issues pertinent to language teaching in a concise, practical, and generally accessible manner.

*The Handbook of Language Teaching* is, as befits its subject, quite broad in scope. Therefore, readers may benefit from considering possible maps for their journeys through the volume, using either those provided by the sections and chapters, or making their own way through the subject index.

For those looking for general introductions, the book is conveniently categorized into seven content areas (social, political, and educational contexts; the psycholinguistic underpinnings of language learning; program design; materials writing and course design; teaching and testing; teacher education; and assessment and evaluation). Many of the 39 chapters simply provide an overview of the topic; these would be a convenient first step on an investigatory journey and could also provide useful comparisons for more knowledgeable readers. For instance, in "Investigating the Effects and Effectiveness of L2 Instruction," De Graff and Housen look at research issues in SLA regarding the role of instruction in language learning, consider the

investigative methods, and make suggestions for future pedagogically oriented research. While admitting the inconclusiveness of current research on what teaching methods, if any, lead to the best learner outcomes, this chapter clearly describes the arguments for and against whether L2 knowledge is best gained explicitly or implicitly. It also explains why there are no clear answers regarding what makes linguistic structures more or less teachable.

Some of the overviews discuss research areas related to language teaching about which little credible research has been as yet published that is directly relevant to the field. In "The Language Learning Brain," Beretta questions whether one can speak of "neurolinguistics," as it presupposes a unified body of knowledge that does not yet exist. Simply put, we do not know enough about brains to accurately say what they do with language. Through a discussion of the empirical evidence that does exist, Beretta dismisses the notion that it is possible at our current levels of knowledge to speak of "brain-compatible language learning" in any credible sense.

Other readers may be making a quick foray into the book, looking for specific guides on how to make particular changes to their classes or programs. Answering such needs is a particular strength of this handbook, as many chapters provide explanatory demonstrations of principles or step-by-step guides that readers can readily take and apply to their own work. J. D. Brown's chapter on foreign and second language needs analysis, for example, simplifies and explains a considerable array of research on what a needs analysis entails, distilling 10 basic steps from the previous literature.

However, individual section or chapter headings may prove overly restrictive for those whose journeys take more meandering paths. For these readers, there is a comprehensive subject index in the back. While by no means a novel feature, this index is particularly useful in allowing the reader to examine common methods and concerns across content areas. For instance, a search under "task-based learning/teaching" leads to discussion in areas as diverse as CALL and pragmatics. Teacher-researchers interested in particular debates or currents within the field can chart their own courses via the index to gain a broader picture than a simple scan of the chapter headings would allow.

Of course, as is inevitable in any large, comprehensive collection on a topic as diverse and controversial as language teaching, there are bound to be aspects of the field that are not covered in sufficient detail. There is only one article specifically on testing, Kunnan and Jang's "Diagnostic Feedback in Language Assessment," though the topic is covered in part by other chapters in the "Teaching and Testing" section. The chapter by Crookes on "Radi-

cal Language Teaching” seems pressed for space as it uses the term “radical” to bridge a panoply of approaches, from Freierian, Marxist-rooted critical methods to feminist pedagogies.

Despite these shortcomings, this handbook serves to introduce and document the state of the art in language teaching at the beginning of the 21st century. In addition, though English is by far the most highly represented language in the volume, most chapters include research and discussion of pedagogical implications for other languages. For example, Koda’s contribution on “Learning to Read in New Writing Systems” would be of special interest to both JSL teachers and European-language native speakers currently learning Japanese.

*The Handbook of Language Teaching* is a weighty volume in both content and price. Nevertheless, its encyclopedic treatment of issues in the field makes it worthy of perusal and, perhaps, purchase. Wiley-Blackwell generally releases paperback editions from its “Handbooks in Linguistics” series within 2 or 3 years after the hardcover is published. However, anyone interested in the state of the art in our field should page through the table of contents, index, and chapters. Readers starting or continuing journeys buffeted by the changing currents within language teaching will undoubtedly find this book to be a useful reference in their explorations.

## Reference

Doughty, C., & Long, M. (Eds.). (2003). *The handbook of second language acquisition*. Oxford: Wiley-Blackwell.

**Researching Collocations in Another Language: Multiple Interpretations. Andy Barfield and Henrik Gyllstad (Eds.). Basingstoke: Palgrave Macmillan, 2009. xix + 273 pp.**

Reviewed by

Dale Brown

Nanzan University

Collocation became something of a buzzword in language teaching in the 1990s. In recent years the idea has been less prominent, but research into L2 collocation has continued, and *Researching Collocations in Another Language: Multiple Interpretations* presents 12 studies from around the world. This is a carefully edited volume. In addition to a clear and useful introduction from the editors, Barfield and Gyllstad, the book's four equal sections (of three chapters) are each concluded by a commentary article by an established figure in the field before an excellent concluding article by Wray gives a wider perspective on the issues. It is very much a book rather than simply a collection of papers.

Section 1 looks at corpus research on L2 collocation. Groom (chapter 2) uses two sub-corpora from the Uppsala Student English Corpus to investigate the impact of time in an English-speaking environment on learners' use of collocations. Reppen (chapter 4) explores differences between two corpora of writing in English, by English L1 children and Navajo L1 children. The highlight in this section, though, is Lin and Adolphs's paper (chapter 3) on phonological coherence as a marker of formulaicity. Using a corpus of Chinese learners' speech, the researchers examine the extent to which the phrase *I don't know why* coincides with intonation unit boundaries. While the results are mixed, the study is unusual in that it focuses on spoken language, and more so in that the data were not elicited in an experimental setting. It thus serves as an interesting methodological model that future research can build on.

In Section 2, lexicographic and classroom materials research, Handl (chapter 6) investigates alternative ways that collocations dictionaries could organize and present their entries, and Jiang (chapter 8) discusses the response of Chinese learners to collocations-focused materials. Most interesting, however, is Komuro's study in chapter 7 of how Japanese learners actually use a collocations dictionary. Using both quantitative and qualitative data, Komuro reveals some of the difficulties learners face, in particular when collocations in the L2 are not structurally equivalent to those in the L1.

The book's third section deals with assessment of L2 collocation knowledge, with three studies on the evaluation and validity of new measurement instruments. Revier's CONTRIX (chapter 10) uses a sentence completion format. Eyckman's DISCO (chapter 11) has learners choosing the two actual collocations from among a set of three candidate strings. Gyllstad's COLLEX (chapter 12) similarly presents three candidate strings, but asks learners to choose the single collocation among them; while his COLLMATCH is a Yes/No format using verb + noun strings as items. All three studies discuss the motivations and the decisions involved in their instrument's construction, and present data from trials of the instruments. My disappointment with this section was that none of the studies really grapples with previous work on the testing of collocations. As the editors note in their introduction, there have been a number of instruments developed, and these three chapters certainly add to that body of work. However, what I feel would be more useful is direct comparisons of different measures. The fact that, as Shillaw (chapter 13) points out in his commentary, each measurement instrument defines collocation slightly differently certainly makes this difficult, but until different measures are compared, it is hard to see how the measurement of L2 collocation will move forward.

The final section of the book is on the learning of collocations. Ying and O'Neill (chapter 14) and Barfield (chapter 16) both use qualitative data and a longitudinal approach. The former investigate learners' reflections on a classroom approach focusing on awareness of the idea of collocation and on learning collocations themselves. The latter looks at how knowledge of collocations develops in parallel with learners' knowledge of an academic field and also at the learners' evolving methods of recording collocations. Peters, in chapter 15, examines the impact of instructions to learners. Given a reading task with both single words and collocations glossed, learners were told either (1) to focus on vocabulary, with the term being left undefined, or (2) specifically to focus on words and collocations. Surprisingly, given that Jiang (chapter 8), Ying and O'Neill (chapter 14), and Barfield (chapter 16) all report a strong tendency for learners to focus on single words, a subsequent test of the glossed items showed no differences between the two groups. Feedback from the participants showed that, despite the instructions, both groups had given attention to the collocations, something Peters ascribes to their advanced level. I think it would be interesting to repeat the experiment without glosses, in other words, to give learners a passage and ask them either to learn vocabulary from it or specifically tell them to look for both single words and collocations.

Looking over all 12 studies, one criticism is that many of them concentrate on the details of the experiments or procedures followed, but spend relatively little time discussing either the theoretical background to their work or the theoretical implications of it. It is to the book's credit that this point is actually made both by Henriksen and Stæhr in their commentary (chapter 17) and by Wray in her conclusion (chapter 18). The majority of the studies here are exploratory, but there are one or two that are very much limited pilot studies or very early reports of work in progress.

One of the book's strengths, in contrast, is that the subtitle is certainly apt, as multiple perspectives on the idea of collocation are provided. Barfield and Gyllstad's introduction highlights the two traditions in collocations research of frequency and phraseology, with most of the papers in this collection having a foot in both camps while leaning towards the former. Another tension clearly highlighted is between approaches that see collocation as a property of individual words and those that see collocations as lexical items in their own right and thus separate to an extent from the words that comprise them (an argument made most forcefully by Revier, in chapter 10). In her conclusion, Wray adds to the perspectives on display by viewing collocation as part of the broader phenomenon of formulaic language and by pointing out the need for connections to be made with researchers grappling with similar questions in areas such as artificial intelligence and communication disorders. It is a fitting end to a thought-provoking book that I would recommend to all those interested in collocation or formulaic language in general.

***Global English Teaching and Teacher Education: Praxis and Possibility.* Seran Dogancay-Aktuna and Joel Hardman (Eds.). Alexandria, VA: Teachers of English to Speakers of Other Languages, 2008. xxi + 198 pp.**

Reviewed by  
Justin Harris  
Kansai Gaigo Senmongakko

Growing knowledge of World Englishes (WE) has resulted in the realization by an increasing number of teachers that there is no single English norm to be aimed at by language learners and at the same time, that nonnative teachers “should not be defined only in terms of their non-nativeness” (Matsuda, 2003, p. 725). It is not uncommon, however, to find English teachers who believe that there is a fixed, unmoving, and “correct” English, which their learners should be unfalteringly striving toward. While there is continuing resistance to the release of the ownership of English, many nonnative teachers suffer from a lack of confidence in that many seem to equate language teaching skills with language ability. If attitudes are to change, nonnative teachers will be the key catalysts in the process.

*Global English Teaching and Teacher Education* is a much-needed addition to an area that has only recently started to gain the attention it deserves. The intention of this book is to provide examples of how nonnative teachers are approaching the teaching of English in their local contexts and to investigate the attitudes of nonnative teachers and teachers-to-be toward English and English teaching. The book is divided into three distinct, yet related areas. It flows well, first introducing the problems of teachers and students not accepting a localized version of English, opting instead for an imagined ideal. The book then addresses issues related to the people who can be instrumental in overcoming these problems, and concludes by highlighting the ways in which nonnative teachers are adapting their teaching styles to their local contexts.

The first section of the book deals with resistance to local Englishes and may hold the most interest for students of WE. The first three chapters introduce three countries formally under British control: Pakistan, Sri Lanka, and Ghana. Although different in many ways, the teachers in these countries are often in opposition to local varieties of English, even though these varieties are currently used in everyday life. In chapter 1, Mahboob and Tallat

discover that the English being tested in schools in Pakistan is not the local version and suggest that what is missing among teachers in Pakistan is not necessarily language-related, but rather a lack of global experience and training in educational methodology. Gunesequera (chapter 2) discusses how in Sri Lanka a majority of city-based teachers refuse to acknowledge a local variety of Sri Lankan English, instead opting for an RP (Received Pronunciation) version of British English dating from the mid-1950s. The reality, however, is that many Sri Lankans, especially in rural areas, are using a local version. National listening tests, typically created by more liberal academics, are recorded in the local pronunciation, putting city-based students who have learned RP at a disadvantage. In chapter 3, Wu and VanderBroek describe how they employed a number of quantitative techniques to investigate Ghanaian English teachers' opinions; they found that while many of these teachers were aware of Ghanaian features of English, they nevertheless aspired to British norms. Each chapter in this section provides good examples of the local English and how a given mother tongue and English cross-pollinate each other. While these chapters may paint a rather negative picture of the conflict between competing versions of English, each outlines ways in which institutions may address the problems.

The second section, entitled "Changing Attitudes Toward English," is an investigation into the attitudes of nonnative teachers toward their English language ability. These chapters may be most useful for teacher-trainers and are particularly relevant to the Japanese EFL context. In chapter 4, Dogancay-Aktuna interviews in-service teachers, while Atay, in chapter 5, interviews teachers-in-training. Dogancay-Atkuna found that Turkish in-service teachers have great confidence in their abilities to teach English. The majority saw their nonnativeness as an advantage because of their status as role models for their students, their understanding of the local context, and their ability to use both the L1 and the L2. Atay, on the other hand, surveyed Turkish pre-service teachers and found quite the opposite. Perhaps due to a lack of experience, they were not confident that nonnative English speakers could be effective teachers.

The third and final section deals with classroom interaction that avoids formulaic approaches to teaching and is responsive to the needs of each group of students in India, South Korea, and Germany. The Indian and South Korean studies (chapters 7 and 8) support the use of code switching in the English classroom. For example, in chapter 7, Vaish gives examples of teachers in India who use code switching as part of their teaching methodology. Vaish explains that many teachers avoid using only the L2 due to

their particular situation and makes a case for an indirect method (i.e., the use of the L1 in the classroom). In chapter 8, Kang argues for the use of code switching in the Korean elementary school classroom, where teachers translate most of the English spoken back into Korean, even though this practice runs counter to government policy. This could be seen as an attempt to support grammar-translation at an early stage of learning in the face of more modern communicative approaches, but the author makes the point that language teaching needs to be situational, and theories and methods of language teaching cannot simply be transferred to any local situation.

Finally, in chapter 9, Erling presents the findings of a quantitative study which investigated German students' beliefs about which "English" they aim to master. Although a majority of students aligned themselves with the traditional American or British Englishes, over 30% showed no affinity to either. The author interprets this as evidence that students are "well aware of the global dimensions of the English language" (p. 159). Erling also sees this as evidence of a move away from teaching English within a given national cultural framework—in this case, as part of British culture. The author also suggests that some students see English becoming a "vehicle to express local identities" (p. 159). This supports what Crystal (1997) has suggested may happen in the future when it becomes the norm to have two forms of English: a simplified and culturally neutral international version for global contexts on one hand, and a localized version reflecting the particular culture and identity of the speaker on the other.

*Global English Teaching and Teacher Education* provides an optimistic look at the state of English teaching for nonnative speakers. It highlights the positive changes that are taking place in the attitudes of these teachers and shows them to be taking ownership of the language. It is essential and thought-provoking reading for anyone who trains nonnative teachers of English and for students of WE because of its insights into the contexts for teaching different Englishes around the globe. The book is a welcome and much needed addition to the WE literature.

## References

- Crystal, D. (1997). *English as a global language*. Cambridge: Cambridge University Press.
- Matsuda, A. (2003). Incorporating world Englishes in teaching English as an international language. *TESOL Quarterly*, 37, 719-729.

***Vocabulary Matrix: Understanding, Learning, Teaching.*  
Michael McCarthy, Anne O’Keeffe, and Steve Walsh.  
Hampshire, UK: Heinle, Cengage Learning, 2010. ix + 165 pp.**

*Reviewed by*

Natsuko Imaoka (今岡奈津子)

Nihon University

In *Vocabulary Matrix*, McCarthy, O’Keeffe, and Walsh offer readers both theoretical and practical ideas to help them teach vocabulary more effectively in their L2 classrooms. Although the book is very thorough, its consistent organization helps readers familiarize themselves with important ideas of L2 vocabulary acquisition. Each chapter consists of three parts. Part A explains issues related to the topic; then the topic is investigated from two perspectives: the learner’s and the teacher’s in Parts B and C, respectively. Part C is particularly useful because practical ideas are sorted according to the students’ L2 abilities. Moreover, each chapter includes examples and short tasks to raise awareness, and ends with a review section. Answers with the authors’ comments and a useful glossary are provided at the end of the book.

The authors launch into this difficult task of teaching vocabulary in chapter 1 by focusing on word formation, including processes such as blending, clipping, initialism, and inflection. Chapter 2 follows with a focus on meanings of words such as denotations, connotations, and register.

Next, the authors shift their attention to word combinations. Chapter 3 looks at collocations, arguing that collocations lead to natural language use. Chapter 4 focuses on grammatical relationships of words, namely colligations. Two types of relationships are introduced in this chapter: paradigmatic relationships which look at words with similar meanings, and syntagmatic relationships which include collocations and colligations. Chapter 5 deals with multiword items such as compounds, prepositional phrases, phrasal verbs, and lexical chunks, while chapter 6 presents idioms, categorized according to their grammatical organization, frequency, whether they are fixed as a whole or partially fixed, and the type of meaning they have, transparent or nonliteral.

The authors then investigate L2 vocabulary in more depth in chapter 7, which deals with word relationships focusing on meaning. Knowing a word according to common sense relationships is not enough; students also need

to know homophones, homographs, and, especially, metaphors. Chapter 8 concentrates on how words are used in texts. The role of discourse analysis in studying longer, natural texts is explored with emphasis on cohesion, coherence, and the use of schemata. In chapter 9, L2 vocabulary acquisition is discussed. The mental lexicon, how the brain keeps vocabulary, is explained by metaphors based on theories of second language acquisition. Chapter 10 then discusses vocabulary in social contexts. We need to know that changes in the society lead to changes in vocabulary and recognize constraints in vocabulary use due to register and factors such as taboos, political correctness, and connotations. This chapter reminds us that sensitivity to using the language is a must since “the offence is in the ear and eye of the listener or reader” (p. 120).

We can understand what our students need to know in terms of L2 vocabulary by looking into native speakers’ vocabulary knowledge. In the field of L2 vocabulary acquisition, individual words were the main focus until researchers such as Pawley and Syder (1983) began paying attention to fixed expressions. Native speakers not only know individual words, but also possess knowledge of how they are combined with other words to form longer phrases. Readers may recall Pawley and Syder’s investigation of abilities of native speakers in which they discussed abilities such as *native-like accuracy*, *native-like fluency*, and *clause-chaining style*. Their work enabled Gass and Selinker (2001) to say “language learning is largely lexical learning” (p. 188). The advancement of technology has also brought about increased interest in word combinations, including “lexico-grammar” (DeCarrico & Larsen-Freeman, 2002).

Due to recent changes of our understanding of the nature of L2 vocabulary, chapters 3 to 6, which focus on word combinations, require careful reading. Especially important are chapters 3 and 4, where collocations and colligations are discussed. To introduce the concept of collocations in chapter 3, the authors ask us to distinguish word combinations that sound natural from those that sound unnatural. For example, although *strong* and *powerful* have a similar meaning, from choices such as *strong wind*, *strong car*, and *powerful car*, students should be able to identify the unnaturalness of *strong car*. There is no rationale for this perception—it is just intuitive, implicit knowledge of word combinations as used by native speakers. This instinct is the line between lower level and more advanced learners.

Chapter 4 focuses on colligations, the grammatical relationship of words. Those who have a knowledge of colligations can fill in the blanks of the following sentence: “Grace sat alone drinking \_\_\_\_ cup \_\_\_\_ tea \_\_\_\_ a chipped

mug" (p. 40). This requires not only vocabulary knowledge, but also grammatical knowledge. The authors stress that L2 vocabulary instruction includes not only learning meanings of new items, but also understanding the relationships of the items with hyponyms, subordinates, antonyms, synonyms, and cognates, and developing a knowledge of word combinations such as collocations and colligations. This is summed up nicely in chapter 10 as the authors argue that the teacher's goal is to help students become independent vocabulary learners, and that it is important to develop a depth of knowledge for the first 2,000 words rather than developing a breadth of knowledge.

*Vocabulary Matrix* is a concise but very informative book on L2 vocabulary acquisition. Although this book covers a wide variety of issues, readers do not need an extensive knowledge of TESOL to benefit from the ideas since theoretical explanations are kept short, and technical terms are clearly explained. Practical ideas for teaching vocabulary range from traditional exercises to new approaches using computers. Keeping a vocabulary notebook with the help of dictionaries, a rather traditional approach, is emphasized throughout the book. I believe this book serves as a good, practical reference book on teaching L2 vocabulary. I think it would be useful for all teachers who understand the importance of vocabulary in L2 acquisition and who are willing to improve their vocabulary teaching by incorporating ideas from the book into their own classroom approaches.

## References

- DeCarrico, J., & Larsen-Freeman, D. (2002). Grammar. In N. Schmitt (Ed.), *An introduction to applied linguistics* (pp. 19-34). London: Arnold.
- Gass, S. M., & Selinker, L. (2001). *Second language acquisition: An introductory course* (2nd ed.). Mahwah, NJ: Erlbaum.
- Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191-225). London: Longman.

***Teaching Second Language Listening*. Tony Lynch. Oxford: Oxford University Press, 2009. xix + 183 pp.**

*Reviewed by*

Joseph Siegel

Ritsumeikan Asia Pacific University

Listening is commonly identified as one of the four main language skills. However, developing the the ability to teach second language listening can be challenging for teachers. In *Teaching Second Language Listening*, Lynch provides guidance to designers and teachers of listening courses. From his own perspective as a second language listener, Lynch illustrates obstacles students face when learning to listen. As a second language teacher, he offers pedagogic recommendations that have proven worthwhile for him and his students.

As the title suggests, the book is targeted principally at those teaching second language listening. The book will most benefit those with little meta-knowledge of the listening process or listening pedagogy. Those already familiar with the fundamental concepts may need to look hard for new information.

*Teaching Second Language Listening* is divided into four sections, the first of which is "Background Issues." This section recognizes the solid impact of technology on the teaching of listening. Chapter 1 discusses various mobile listening devices such as iPods and laptops that are now available. A theoretical progression including communication theory and information processing are outlined up to the more recent acceptance of social constructivism, which "emphasizes the place of the individual in a social environment" (p. 12). In chapter 2, Lynch details features of spoken output that may prove problematic to second language listeners. Aspects such as redundancy and nonverbal communication are noted as being of potential use to listeners. Brief reviews of research on speed and accent give teachers something to consider when selecting listening texts.

Part Two, "Listening Processes," looks at "Recognition," "Interpretation," and "Participation" (chapters 3-5). In "Recognition" (chapter 3), parts of Levelt's model of spoken communication (Levelt, 1993, as cited in Lynch, p. 30) are used to demonstrate the internal decoding and formulation of meaning that takes place during listening. These components are loosely linked to Anderson's notions of perception, parsing, and utilization (Anderson, 1985, as cited in Lynch, p. 32). Additionally, Lynch emphasizes the importance of raising students' awareness of connected speech. Research into lexical and

syntactic challenges facing listeners is also reported, leading to a conclusion that vocabulary is more helpful to L2 listeners than grammatical aptitude.

Chapter 4 moves on to discuss ways in which listeners interpret a speaker's message. Resources available to listeners such as background and sociocultural knowledge are highlighted. Lynch adopts a cautious tone when discussing the listening process, as research in this area continues to develop. Chapter 5 centers on two-way listening. Lynch offers what he calls a framework for participation within which the negotiation of meaning is stressed. In other words, listeners should have rights to interrupt, ask for clarification, and make other conversational modifications. Listening course designers and teachers are strongly encouraged to include more two-way listening activities in order to prepare students for real life. Lynch places so much emphasis on the teaching of two-way listening that one-way listening instruction receives little attention. Yet, many second language students need one-way listening skills to achieve high scores on internationally recognized tests of English listening such as the TOEIC and the TOEFL.

Part Three deals with listening pedagogy and materials. In chapter 6, the debatable topic of listening strategies and skills is presented. Lynch also ponders a fundamental question in recent pedagogical discussions: Can listening strategies be taught? The pros and cons of strategy training are covered and lead to a tentative stalemate regarding the effectiveness of teaching strategies. Practical steps for teaching listening strategies are included for the benefit of interested teachers.

Listening materials and tasks are explored in the next chapter. The recent history of listening materials development is considered, along with matters of evaluating, adapting, and grading tasks and texts. Both listening teachers and materials developers are prompted to go beyond literal comprehension and lead learners to interpretation, inference, and reaction. Though materials creation is mentioned, this area deserves more attention. Teachers may need guidance in transferring the theories and anecdotes described earlier in the book to their classrooms. A more developed section with a procedural "how to" outline would have helped to distinguish this book from other works on the subject.

The book's main original contribution comes in chapter 8, "Integrating Listening with the Other Skills." Useful subsections describe how listening can (and should) be linked to other language skills. Connecting listening to other skills rather than attempting to isolate it is a worthwhile endeavor and signifies the future of listening pedagogy. Lynch points to links between listening and reading comprehension as well as to the symbiotic relation-

ship between speaking and listening. Sample activities clearly reflect an integrated approach. Chapter 9 moves on to assessment. Typical complications in gauging listening abilities are detailed and alternative evaluation methods are presented.

The two chapters in Part Four examine listening outside the classroom in self-access centers (SAC) and the real world. Lynch believes that SAC activities need to be differentiated from in-class activities and should not merely replicate classroom practices. In addition, extensive listening is identified as not only advantageous for students but also as fertile ground for research. This section includes descriptions of three projects implemented to encourage student listening outside of school. A useful list of listening websites is provided as a resource for teachers and their students.

Readers will find that the strength of the book is in Lynch's breadth of coverage. Theories of listening and listening processes are linked to classroom approaches and techniques. Although several experts in the field have offered varying definitions of the term "listening" (see Hasan, 2000; Vandergrift, 1999), no definition is provided at the beginning of this book. By making explicit what "listening" means to him, Lynch could have been clearer about what exactly needs to be taught. While the book does not add to existing theories of listening, early chapters explain the background of listening comprehension in lay terms and can equip novice and experienced teachers with a better understanding of the subject. Throughout the book, Lynch describes aspects of conventional listening classes and offers alternatives in areas of approach, materials, and testing.

In the end, uncertainty remains about how best to teach second language listening. Lynch's work raises awareness of common pitfalls and provides sensible direction for teachers and course designers, but definitive results are not ensured. Those with little experience can quickly grasp some of the relevant issues related to the teaching of listening, while the book also provides more knowledgeable educators a perspective into current ideas about listening pedagogy.

## References

- Hasan, A. (2000). Learners' perceptions of listening comprehension problems. *Language, Culture, and Curriculum*, 13, 137-153.
- Vandergrift, R. (1999). Facilitating second language listening comprehension: Acquiring successful strategies. *ELT Journal*, 53, 168-176.

***Second Language Learning and Identity: Cracking Metaphors in Ideological and Poetic Discourse in the Third Space.* Mika Yoshimoto. Amherst, NY: Cambria Press, 2008. xv + 312 pp.**

Reviewed by

Yuan Yuan

Kobe University

In the mosaic of social and human sciences, the issues of *identity* and *second language learning* have been thoroughly researched in two independent disciplines. However, it is not until recently that the significance of exploring the interconnection between second language learning and identity has been realized. Within this expanding discipline, researchers have been struggling to theorize the fluid, highly abstract characteristics of identity within the ongoing process of second language learning. *Second Language Learning and Identity: Cracking Metaphors in Ideological and Poetic Discourse in the Third Space* is an important contribution to the emerging literature on identity and language learning. Mika Yoshimoto's fresh, innovative, analytic approach, namely a combination of auto-ethnography and poetic discourse metaphors such as *haiku*, offers invaluable insights into English language learning experiences of Japanese women in Japan and in English-speaking contexts. She makes complex theoretical constructs such as identity and ideological discourse come alive for readers. She also challenges the reader to re-think language learning and language use from a broader and more political perspective.

The book is divided into three major parts: an introduction (chapters 1-3), analysis (chapters 4-6), and discussion (chapter 7). Chapter 1 offers a sensitive picture of the setting, touching on the author's motivation and desire for exploring the identity issue. Through a condensed autobiography delivered in the forms of *haiku* and narration, readers can trace Yoshimoto's own torture of being a female Japanese as well as an EFL scholar. In chapter 2, after casting a critical eye on a sizable body of literature on language and identity, Yoshimoto situates the study as data driven, interpretative, empirical research. In her qualitative research, she aims at understanding these three questions: "1) What does it mean for a Japanese woman to study English? 2) What stories of Japanese students emerge in in-between spaces? and 3) How do Japanese metaphors relate to Japanese women's identity construction?" (p. 57). Chapter 3, only seven pages in length, provides a brief

introduction to the other three participants in the analysis, Rie, Yoko, and Aya, and to the collected data.

Chapter 4, "Haiku/Emerging Self," contains two parallel parts: an explanation of haiku and the theme of emerging self. This chapter is somewhat confusing because the haiku part had already been mentioned in chapter 1, and the narration of self could have been combined with the three participants' stories in chapter 5.

Chapter 5 delves into the tapestry of the four participants' stories in learning English as a second language. Yoshimoto's use of haiku to theorize the themes that emerged is an innovative approach. The ambiguous poetic characteristic of haiku enables Yoshimoto to illustrate the implicit connection between the participants and the social context they live in, to maintain the fluid, highly abstract characteristics of identity, and to evoke the complexity of the relationship between identity construction and language learning. As an example, she subtitles Rie's story with "ゴム鞠は 変貌自在 強さ秘め" [translated as "Morphing rubber ball, moves anywhere gracefully, a secret mission"] (p. 78). Rie is compared to a moving, morphing rubber ball because of her elasticity in being able to adjust to different environments. However, as the metaphor conveys multiple connotations, it cannot always be interpreted in the same way. As Dey (1993) puts it, "metaphors can raise inappropriate as well as appropriate connotations, and so contribute to confusion rather than clarity" (p. 246). Though the poetry lines were enjoyable to read, I question how this vague sense of metaphor fits the rigorous requirements of an academic work.

Chapter 6 scrutinizes the impact of Japanese metaphors on Japanese women's identity construction. The participants' narrations work as compelling voices revealing how they have been influenced and constrained by the use of gender identity-related proverbs. Through carefully examining the social-ideological meaning of these proverbs, Yoshimoto asks that we seriously consider the impact and consequences of the language we use in our social communication.

Chapter 7, the final chapter, speculates on the central question of what it means for a Japanese woman to study English. The focus shifts from being a female EFL learner in Japan to being a Japanese ESL learner in an English-speaking country. Yoshimoto leads us from gender identity to linguistic identity. In her analysis, she shows how the participants suffer from, negotiate with, and finally triumph over assumptions about the English identity of nonnative speakers compared to native English speakers.

I recommend this book for three reasons. First, the sheer volume of empirical material that Yoshimoto has collected gives the book an aura of authority providing the reader with an understanding of the process of learning English for female Japanese learners in both Japanese and English-speaking contexts. Second, Yoshimoto's attempt to integrate the Japanese poetic writing style of haiku into a so-called western academic writing style provides a means of capturing the complicated relationship between identity and second language learning. Haiku is used beyond poetic style; it creates a space for Yoshimoto to appeal to the participants' desires, to publicize their thoughts in a comparatively implicit and secure way. Most importantly, this book encourages us to rethink and to react to the impact of the ideology conveyed through language use in our daily communication. For example, most women in Japan will feel proud by being praised "女らしい" (feminine), without critically thinking about the dominant ideology of "woman should be feminine." Yoshimoto notes, "I myself, and now Yoko, too, have found that when we go abroad, the Japanese ideal of beauty is not equivalent to the ideal of beauty in other places" (p. 101).

After reading the book, I felt as if, as a female EFL learner living in Japan, I had had a long conversation with Yoshimoto. I give my heartfelt thanks to her for sharing her struggles. I was inspired by her understanding of language learning, her courage to challenge the existing powerful stereotypes, her passion for life, and her compassion for humanity. Finally, when I closed the book, I imagined a butterfly flying toward the sky, the mellow sound of a flute echoing in my head. This book is like the deep red color of the kimono on the cover, which gives me a sense of maturity and peace.

Then I made my own haiku:

蚕まゆ、せみの抜け殻、蝶の旅

Silkworm's cocoon, cicada's cast-off skin, journey to become a butterfly.

## Reference

- Dey, I. (1993). *Qualitative data analysis: A user-friendly guide for social scientists*. London: Routledge.

# Information for Contributors

All submissions must conform to *JALT Journal* Editorial Policy and Guidelines.

## Editorial Policy

*JALT Journal*, the refereed research journal of the Japan Association for Language Teaching (*Zenkoku Gogaku Kyoyiku Gakkai*), invites empirical and theoretical research articles and research reports on second and foreign language teaching and learning in Japanese and Asian contexts. Submissions from Asian and other international contexts are accepted if applicable to language teaching in Japan. Areas of particular interest include but are not limited to the following:

1. Curriculum design and teaching methods
2. Classroom-centered research
3. Cross-cultural studies
4. Testing and evaluation
5. Teacher training
6. Language learning and acquisition
7. Overviews of research and practice in related fields

The editors encourage submissions in five categories: (a) full-length articles, (b) short research reports (*Research Forum*), (c) essays on language education framed in theory and supported by argumentation which may include either primary or secondary data (*Perspectives*), (d) comments on previously published *JALT Journal* articles (*Point to Point*), and (e) book and media reviews (*Reviews*). Articles should be written for a general audience of language educators; therefore, statistical techniques and specialized terms must be clearly explained.

## Guidelines

### Style

The *JALT Journal* follows the *Publication Manual of the American Psychological Association*, 6th edition (available from APA Order Department, P.O. Box 2710, Hyattsville, MD 20784, USA; by email: <order@apa.org>; from the website: <www.apa.org/books.ordering.html>). Consult recent copies of *JALT Journal* or *TESOL Quarterly* for examples of documentation and references. A downloadable copy of the *JALT Journal* style sheet is also available on our website at <www.jalt-publications.org/jj/>.

### Format

Full-length articles must not be more than 20 pages in length (6,000 words), including references, notes, tables, and figures. *Research Forum* submissions should not be more than 10 pages in length. *Perspectives* submissions should be not more than 15 pages in length. *Point to Point* comments on previously published articles should not be more than 675 words in length, and *Reviews* should generally range from 500 to 1000 words. All submissions must be word processed in A4 or 8.5 x 11" format with line spacing set at 1.5 lines. **For refereed submissions, names and identifying references should appear only on the cover sheet.** Authors are responsible for the accuracy of references and reference citations.

### Submission Procedure

Please submit the following materials, except for reviews, as an email attachment in MS Word format to the appropriate editor indicated below:

1. Cover sheet with the title and author name(s).
2. One (1) copy of the manuscript, with no reference to the author. Do not use running heads.
3. Contact information sheet, including one author's full address and, where available, a fax number.
4. Abstract (no more than 150 words).
5. Japanese translation of the title and abstract, if possible (no more than 400ji).
6. Biographical sketch(es) (no more than 50 words each).

**Submissions will be acknowledged within 1 month of their receipt.** All manuscripts are first reviewed by an editor to ensure they comply with *JALT Journal* Guidelines. Those considered for publication are subject to blind review by at least two readers, with special attention given to (1) compliance with *JALT Journal* Editorial Policy, (2) the significance and originality of the submission, and (3) the use of appropriate research design and methodology. Evaluation is usually completed

within 3 months. Each contributing author of published articles and Book Reviews will receive one complimentary copy of the *Journal* and a PDF of the article (Book Reviews are compiled together as one PDF). *JALT Journal* does not provide off-prints. Contributing authors have the option of ordering further copies of *JALT Journal* (contact JALT Central Office for price details.).

### **Restrictions**

Papers submitted to *JALT Journal* must not have been previously published, nor should they be under consideration for publication elsewhere. *JALT Journal* has First World Publication Rights, as defined by International Copyright Conventions, for all manuscripts published. If accepted, the editors reserve the right to edit all copy for length, style, and clarity without prior notification to authors.

### **Full-Length Articles, Research Forum, Perspectives, and Point to Point Submissions**

Please send submissions in these categories or general inquiries to:

[jj-editor@jalt-publications.org](mailto:jj-editor@jalt-publications.org)

Darren Lingley, *JALT Journal* Editor

### **Japanese-Language Manuscripts**

*JALT Journal* welcomes Japanese-language manuscripts on second/foreign language teaching and learning as well as Japanese-language reviews of publications. Submissions must conform to the Editorial Policy and Guidelines given above. Authors must provide a detailed abstract in English, 500 to 750 words in length, for full-length manuscripts and a 100-word abstract for reviews. Refer to the Japanese-Language Guidelines for details. Please send Japanese-language manuscripts to:

[jj-editorj@jalt-publications.org](mailto:jj-editorj@jalt-publications.org)

Ken Urano, *JALT Journal* Japanese-Language Editor

### **Reviews**

The editors invite reviews of books and other relevant publications in the field of language education. A list of publications that have been sent to JALT for review is published bimonthly in *The Language Teacher*. Review authors receive one copy of the *Journal*. Please send submissions, queries, or requests for books, materials, and review guidelines to:

[jj-reviews@jalt-publications.org](mailto:jj-reviews@jalt-publications.org)

Bill Perry, *JALT Journal* Reviews Editor

### **Address for Inquiries about Subscriptions, Ordering JALT Journal, or Advertising**

JALT Central Office

Urban Edge Building 5F

1-37-9 Taito, Taito-ku, Tokyo 110-0016, Japan

Tel.: 03-3837-1630; Fax: 03-3837-1631

(From overseas: Tel.: 81-3-3837-1630; Fax: 81-3-3837-1631)

Email: [jco@jalt.org](mailto:jco@jalt.org) URL: [www.jalt.org](http://www.jalt.org)

## 日本語論文投稿要領

JALT Journalでは日本語で執筆された論文、研究報告、実践報告、書評等を募集しています。文体:一般的な学術論文のスタイルを用い、章立ての仕方や参考文献のデータの書き方などは、*Publication Manual of the American Psychological Association* (6th Ed.)の定める方式に合わせて下さい。不明の場合は、JALT Journalの英語論文を参考にするか、日本語編集者までお問い合わせ下さい。なお、JALT Journalの読者は現場の教師が主なので、特殊な専門用語や統計的手法は、わかりやすく定義するか説明を加えるなどして下さい。原稿: 長さは、参考文献リストも含め18,000字(書評の場合は 1,500字)以内です。A4の用紙に横書きで、1行40字、1ページ30行で印刷して下さい。手書きの原稿は受け付けません。

### 提出するもの:

以下の原稿を電子メールの添付書類、あるいは郵送でお送りください。

- 執筆者の名前と所属機関名を書いた表紙
- MS-Word で保存した本文(執筆者は無記名のこと)
- 執筆者連絡先(住所、電話番号、ファックス、e-mail アドレス)
- 400字以内の和文要旨
- 英文のタイトルと、500~750語の英文要旨(書評の場合は100語程度の英文要旨)
- 100字以内の執筆者略歴
- 審査を経て掲載の認められた草稿は、図表などを全て写植版にしたものにして提出すること

**査読:**編集委員会で投稿要領に合っているかどうかを確認したあと、少なくとも二人の査読者が査読を行います。査読者には執筆者の名前は知らされません。査読の過程では特に、原稿がJALT Journalの目的に合っているか、言語教育にとって意味があるか、独創性はあるか、研究計画や方法論は適切か等が判定されます。査読は通常二か月以内に終了しますが、特に投稿の多い場合などは審査にそれ以上の時間がかかることがあります。

**注意:**JALT Journalに投稿する原稿は、すでに出版されているものや他の学術雑誌に投稿中のものは避けて下さい。JALT Journalは、そこに掲載されるすべての論文に関して国際著作権協定による世界初出版権を持ちます。なお、お送りいただいた原稿は返却しませんので、控を保存して下さい。

投稿原稿送り先またはお問い合わせ:

〒062-8605 北海学園大学 経営学部  
JALT Journal日本語編集者 浦野 研  
電話:(011)841-1161(代) Fax:(011)824-7729  
jj-editorj@jalt-publications.org

### **JALT Journal** 第32巻 第2号

2010年10月20日 印刷  
2010年11月 1日 発行  
編集人 グレン・リングリィ  
発行人 カロリン・ロイド  
発行所 全国語学教育学会事務局  
〒110-0016 東京都台東区台東1-37-9 アーバンエッジビル5F  
TEL (03)3837-1630; FAX (03)3837-1631  
印刷所 コーシンシャ株式会社  
〒530-0043 大阪市北区天満1-18-4天満ファーストビル301 TEL (06)6351-8795