jalt journal

The research journal of the Japan Association for Language Teaching

Volume 45 • No. 1 • May 2023

JALT

全国語学教育学会

Japan Association for Language Teaching

¥1,900 ISSN 0287-2420

JALT Journal

Volume 45 • No. 1

May 2023

Contents

May 2023
Volume 45 • No. 1

JALT Journal Information

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 offers 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. JALT has 32 JALT chapters and 30 special interest groups (SIGs) and is a founder of PAC (Pan-Asian Consortium), which is an association of language teacher organizations in Pacific Asia. PAC holds annual 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 a branch of IATEFL (International Association of Teachers of English as a Foreign Language). JALT is also affiliated with many other international and domestic organizations.

JALT publishes *JALT Journal*, a semiannual research journal; *The Language Teacher*, a bimonthly periodical containing articles, teaching activities, reviews, and announcements about professional concerns; and the annual *JALT Postconference Publication*.

The JALT International Conference on Language Teaching and Learning and Educational Materials Exposition attracts some 2,000 participants annually and offers more than 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 copies of JALT publications, free or discounted admission to JALT-sponsored events, and optional membership in one chapter and one SIG. For an annual fee of ¥2,000 per SIG, JALT members can join as many additional SIGs as they desire. For information about JALT membership, contact the JALT Central Office or visit the JALT website.

JALT National Officers (2023)

President: . . . . . *Clare Kaneko*
Vice President: . . . . . *Kenn Gale*
Auditor: . . . . . *Robert Chartrand*
Director of Treasury: . . . . . *Michael Mielke*
Director of Records: . . . . . *Samantha Kawakami*
Director of Program: . . . . . *Wayne Malcolm*
Director of Membership: . . . . . *Julie Kimura*
Director of Public Relations: . . . . . *William Pellowe*

Chapters

Akita, East Shikoku, Fukui, Fukuoka, Gifu, Gunma, Hiroshima, Hokkaido, Ibaraki, Iwate-Aomori, Kitakyushu, Kobe, Kyoto, Matsuyama, Nagoya, NanKyu, Nara, Niigata, Oita, Okayama, Okinawa, Osaka, Saitama, Sendai, Shinshu, Shizuoka, Tokyo, Tottori, Toyohashi, West Tokyo, Yamagata, Yokohama

Special Interest Groups

Bilingualism; College and University Educators; Computer-Assisted Language Learning; Critical Thinking; Extensive Reading; CEFR and Language Portfolio; Gender Awareness in Language Education; Global Issues in Language Education; Intercultural Communication in Language Education; Japanese as a Second Language; Learner Development; Lifelong Language Learning; Literature in Language Teaching; Materials Writers; Mind, Brain, and Education; Mixed, Augmented, and Virtual Realities; Other Language Educators; Performance in Education; Pragmatics; School Owners; Study Abroad; Task-Based Learning; Teacher Development; Teachers Helping Teachers; Teaching Younger Learners; Testing and Evaluation; Vocabulary

JALT Central Office

Level 20, Marunouchi Trust Tower–Main,
1-8-3 Marunouchi, Chiyoda-ku, Tokyo 100-0005 Japan
Tel.: (+81) 3-5288-5443; Email: jco@jalt.org;
Website: <https://jalt.org>

In This Issue

Articles

This issue contains four full-length research articles in English, one *Research Forum* article in English, and two *Expositions* articles in English. In the first full-length research article, **Andrew Obermeier** presents a vocabulary study focusing on flashcard learning of multiword expressions (MWEs). He identifies differences in how the implicit knowledge development of figurative and literal expressions are differently affected by learning MWEs using flashcards. The study shines light on the potential value of strategic flashcard learning for frequent MWEs. Secondly, **Randy Appel** and **Lewis Murray** analyze a learner corpus of 420 argumentative essays written in English by learners from three L1 backgrounds (Chinese, Korean, Japanese). They identify intra-group tendencies and intergroup production differences using a contrastive interlanguage approach. The results show tendencies regarding functional categories and individual lexical bundles for each L1 group. Thirdly, **Yukiko Ishikawa** investigates the Self-Regulated Learning (SRL) development of eight first-year college students. The results of this year-long study are used to discuss the social cognitive model of SRL in the Japanese context with data at each of the four levels of SRL development: observation, emulation, self-control, and self-regulation. This study provides teachers with insights which can potentially facilitate the creation of a learning environment in which learners set meaningful and achievable goals, and identify and use strategies for self-study. Fourthly, **Sarah Louise Mason**, **Alice Chik**, and **Peter Roger** discuss how researchers working at Japanese universities across the country perceive their research trajectories, the purpose of their research, and how they understand their research engagement, realities affected by increased research production requirements and accountability measures.

The *Research Forum* article is by **Aki Tsunemoto**, **Pakize Uludag**, **Kim McDonough**, and **Talia Isaacs**, and focuses on the relationship between holistic judgments of second language (L2) speech fluency (i.e., perceived fluency) and temporal measures of fluency (i.e., utterance fluency) using an English read-aloud task with Japanese secondary school students. The results showed that articulation rate and clause-internal pauses were significant predictors for perceived fluency.

The first *Expositions* paper is by **Thomas S. C. Farrell**, and discusses important, practical aspects of reflective practice for TESOL teachers. It also provides two different frameworks developed by the author to help teachers

reflect on their practice. The second *Expositions* paper, by **Christine P. Casanave**, provides a unique and direct look into journal writing as a productive and creative process of socialization and of self-reflection for both teachers and students in their attempts at understanding and shaping their academic, research-based, personal, and professional lives.

Reviews

In this issue, readers have access to six reviews on titles with theoretical content and practical applications to serve the interests of researchers and language instructors. **Kathryn Akasaka** opens with a review of *Task-Based Language Teaching: Theory and Practice* from the Cambridge Applied Linguistics series. **Tim Greer** follows with an examination of an edited volume covering the cross-disciplinary and practical applications for Conversation Analysis in second language (L2) classrooms, content-based language instruction, teacher education, and assessment. Next, **Martin Hawkes** takes up the fifth edition of Lightbown and Spada's *How Languages are Learned*, which he contrasts with the earlier fourth edition that was also published in Japanese. **Paul Hullah** then outlines the scope, cognitive perspectives, emotions, and innovations of pedagogical stylistics addressed in chapters based on empirical and theoretical work. **Martyn McGettigan** looks at the monograph authored by Daniel O. Jackson on his research into the underexplored form of reflection: language teacher noticing. Finally, the edited work on humor competency training in ELT from Japan-based John Rucynski Jr. and Caleb Prichard is reviewed by **Cathrine-Mette Mork**.

From the Editors

This is my first issue serving as the Editor of *JALT Journal*. In my time as Associate Editor of the journal, I learned much from, and would like to thank, the past-outgoing Editor, **Eric Hauser**, who got me interested in working on the *JALT Journal* team. After working with the outgoing Editor, **Gregory Paul Glasgow**, for two years, I look forward to working with him in his capacity as Consulting Editor. I have no doubts that I will continue to find solace in Gregory's professional advice and warm friendship. I am also very grateful to **Jeremie Bouchard**, who is the Associate Editor of *JALT Journal*. Jeremie's early transition to the onboarding process made my transition to the Editor role a smooth one.

—Dennis Koyama, *JALT Journal* Editor

We, the *JALT Journal* team, extend our sincere thanks to outgoing Assistant Reviews Editor, **John Nevara**, whose professional support with book reviews has been invaluable. John, thank you for your contributions and insights. We also thank **Theron Muller**, JALT Publications Chair, for his unflagging support of the JALT organization and the *JALT Journal* team. We welcome **Charles Mueller** as our new Associate Japanese-Language Editor. He, Japanese-Language Editor **Kiwamu Kasahara**, and Associate Japanese-Language Editor **Rintaro Sato** welcome Japanese-language manuscript submissions, to ensure that *JJ* can fulfill its mission as a Scopus-approved, bilingual academic publication. We also send our appreciation and gratitude to the *JALT Journal* Editorial Board, our other reviewers, **Cameron Flinn** (*JJ*'s new Production Editor), our proofreaders, and to the authors who submit manuscripts, without whom it would not be possible to publish the journal.

With the beginning of academic year 2023, and all that it entails for so many teachers and language learners across Japan, we would like to make a few announcements and remind our readership of a few important points.

JALT Journal remains committed to publishing high-quality research relevant for language learning and teaching in the Japanese context. We invite readers to read our updated "Aims and Scope" section (formerly "Editorial Policy") in the backmatter, and to consider submitting their research for publication in *JALT Journal.* The previous *JJ* issue was our first special issue and focused on the far-ranging and critically-rich topic of race and native-speakerism in language education, and we invite our readership to consider submitting special issue proposals. Specific details on the submission process for special issue proposals are available on the *JJ* website and are printed at the end of the current issue.

We would also like to highlight that the journal's *Point-to-Point* section is another opportunity for prospective authors to engage in scholarly debate by commenting on an article published in *JJ*. These 1000-word papers are an important part of the journal's contribution to the free exchange of scholarly ideas in our field. The original authors are also invited to follow up with a response to the discussion of their work. We hope you will consider submitting a *Point-to-Point* piece based on an article published in *JALT Journal*.

Finally, we have created a new position on the *JALT Journal* team, English-language Assistant Editor. Experience conducting research and/or writing and publishing academically are important prerequisites for the position. If you are interested in this new position, the position's requirements for application are posted on the *JJ* website. If you have any questions about the position or are interested in volunteering with us, please contact Dennis Koyama at jaltpubs.jj.ed@jalt.org, or Jeremie Bouchard at jaltpubs.jj.ed2@jalt.org.

—Dennis Koyama, *JALT Journal* Editor

—Jeremie Bouchard, *JALT Journal* Associate Editor

Articles

Learning Multiword Expressions with Flashcards: Deliberate Learning and L2 Implicit Knowledge Gains

Andrew Obermeier
Kyoto University of Education

This research investigated two aspects of second language learning: how implicit knowledge develops through explicit learning and how this is affected by multiword expression compositionality. More specifically, the experiment investigated how flashcard learning affected the implicit knowledge development of literal and figurative expressions. As these two types are composed differently, it was hypothesized that their implicit knowledge development would likewise differ. A lexical decision task was conducted in a masked repetition priming experiment to measure implicit knowledge gains, and response time data were analyzed in a linear mixed-effects model with participants and items set as random effects. Results showed that flashcard learning affected the implicit knowledge development of figurative and literal expressions differently.

Keywords: explicit learning; flashcards; implicit knowledge; interface; multiword expressions

本研究では、第二言語学習の2つの側面である、複単語表現の構成性と、明示的学習を通じて暗示的知識がどのように発達するかについて調査した。具体的には、フラッシュカードによる学習が、文字通りの表現と比喩表現の暗示的知識の発達にどのような影響を与えるかを調査した。この2つの表現は構成が異なるため、暗示的知識の発達も同様に異なるという仮説を立てた。暗示的知識の獲得を測定するために、マスク下の反復プライミング法を用いた実験で、語彙

https://doi.org/10.37546/JALTJJ45.1-1

JALT Journal, Vol. 45, No. 1, May 2023

性判断課題を実施し、応答時間データを、参加者と項目をランダム効果として設定した線形混合効果モデルで分析した。その結果、フラッシュカードによる学習は、比喩表現と文字通りの表現の暗示的知識の発達に異なる影響を与えることが示された。

キーワード: フラッシュカード、明示的学習、暗示的知識、複単語表現

Corpus linguists have found that *multiword expressions* (MWEs) make up about 59% of spoken and 52% of written English (Erman & Warren, 2000), so an essential issue for second language learners and teachers is understanding how they are acquired. Although various terms are used to refer to them (e.g., *formulaic sequences*, *chunks*, *collocations, idioms, conventional expressions*), this paper uses MWEs as an umbrella term covering all types of expressions (Siyanova-Chanturia, 2017). This research focused on two broad MWE categories: *figurative* and *literal expressions* (or *figuratives* and *literals*). The opaque meanings of figuratives (e.g., *kick the bucket, once in a blue moon*) make them more challenging to learn and process than literals (e.g., *all the time, get the idea*), which are transparent. As literals and figuratives are composed differently, the investigation focused on whether learning them is likewise different.

Another important issue is the intersection of explicit learning and implicit knowledge development because a high priority for language teachers is to foster these two processes for students. Explicit learning activities are conscious processes such as interpreting textbook explanations, doing worksheet exercises, practicing with drills, and rote memorizing. Explicit knowledge can be applied to monitoring language correctness or incorrectness and is often the focus of tests. Implicit knowledge develops unconsciously as the interlanguage system becomes fine-tuned through use, by which learned language can be accessed more fluently. Second language learners must learn explicitly and develop implicit knowledge to become proficient.

This study reports on a masked, repetition, priming experiment that compared the effects of learning literals and figuratives using flashcards. As this is an explicit learning method, and as masked repetition priming measures implicit knowledge development, the investigation addresses the interface regarding these two MWE types.

Multiword Expressions and Their Compositionality

A central issue to research on the processing of MWEs is that they vary widely regarding their *compositionality*, the degree to which the individual

words that comprise them make up the meaning of the whole expression. The composition of MWEs varies along a continuum. Although some are transparent (i.e., *stay away*), others are less transparent but easy to process (i.e., *on the road*), and others are opaque (i.e., *once in a blue moon*). Grant and Bauer (2004) established major compositional categories showing how literal and figurative expressions generate meaning differently. The meanings of individual words in figuratives differ from those of the whole metaphorical expressions (e.g., *when pigs fly, walk on air*). Conversely, in literal expressions, each word directly contributes to the overall meaning (e.g., *get the idea, know better*).

Although great variation in the metaphorical makeup of figurative expressions exists (see Goatly, 2011), in this experiment, MWEs were allocated to two broad categories: either literal or figurative, depending on their opacity. In other words, although *get the idea* and *all the time* have figurative elements, they are nonetheless highly transparent, so they were classified as literals. Contrastingly, a few expressions such as *kick the bucket* are so opaque that Grant and Bauer (2004) classified them as *core idioms*, arguing that no discernable etymological metaphorical connection can be made. Nevertheless, such terms were classified as *figuratives* because learners could make metaphorical connections to remember them. Understanding literals involves naturally processing the words. This process is more straightforward than understanding figuratives, which involves deriving meaning from metaphors as well as rejecting the literal interpretation of each constituent word.

Collocation dictionaries (e.g., Kjellmer, 1994; Sinclair, 1995) contain thousands of entries and serve as valuable references for seeing examples of their use, but one reason these are not very practical guides for second language learners is that they do not address this issue of compositionality. Compositionality raises problems for language learners because even when they know the correct figurative meanings, they strongly favor literal word interpretations, (e.g., Cieślicka, 2006, 2012). To fill this gap, Martinez and Schmitt (2012) made the PHRASal Expressions List, composed of MWEs that are frequent, meaningful, and difficult for language learners to interpret. Martinez and Schmitt (2012) also provided frequency levels for the 505 MWEs on their list to facilitate prioritization for learning along with the first five thousand most frequent individual words on the British National Corpus. Some MWEs on the PHRASE List are difficult for learners due to their opacity (i.e., *end up*), and others cause problems because they are easily misinterpreted (i.e., although *at all* is very clear in its positive sense

as in *at all times,* it is much less so in its negative sense, as in *Do you exercise at all?*).

Because highly opaque figurative expressions must be remembered as wholes, and transparent literal expressions can be understood when processed word by word, it may follow that MWE compositionality affects whether they are holistically processed. Research on MWEs shows that they may be retrieved holistically rather than being created from scratch by applying grammar (e.g., Sinclair, 1991; Tremblay et al., 2011; Wray, 2002), but the nature of this holistic processing is complex (Siyanova-Chanturia & Martinez, 2015). Holistic processing for figuratives entails both automatically interpreting the whole MWE's meaning to form link and processing the word sequence; whereas, the holistic processing of literals only entails recognizing the word sequence and processing it faster. By comparing the effects of deliberately learning literal and figurative MWEs, this research aims to shed light on whether holistic processing relates to compositionality.

Multiword Expression Flashcard Learning

Deliberate paired-associate vocabulary learning with flashcards involves repeatedly retrieving targets from meanings or meaning from targets. This systematic and repeated retrieval method is a well-established way for language learners to connect first language meanings with L2 vocabulary. Learners can remember vast numbers of paired associates in a short time. For example, Thorndike (1908) showed that 1,200 words studied for 30 hours showed remarkable persistence in memory. Digital flashcard applications now enable language teachers and learners to systematize a database of words to memorize conveniently. Nakata (2011) extensively reviewed free online flashcard applications, considering pedagogically essential features such as presentation mode variety, adaptive sequencing, and timing settings for spaced review. Retrieval using flashcards is more effective than word lists because learners can remove target items they have mastered. Also, cards can easily be shuffled, giving them another advantage over static lists of items in which the order is unchangeable. In static lists, the sequence of the list is also remembered, providing false memory support for the individual items, thus hindering proper lexical knowledge development.

Once learners have a solid base of single-word knowledge, MWE learning is another important goal. Learning MWEs as wholes with flashcards may be an effective learning strategy. Learners can expand their collocation knowledge by practicing with MWE-to-meaning pairs. Given that literal expres-

sions and figurative expressions generate meaning differently, practicing them with flashcards will facilitate learning in different ways. Each word matches its meaning for literal expressions, so flashcard practice will help with fluency development. For figurative expressions, each word must be relearned in its metaphorical context, so practicing with flashcards will both strengthen the meaning-to-form connection and foster processing fluency.

Explicit and Implicit Second Language Learning

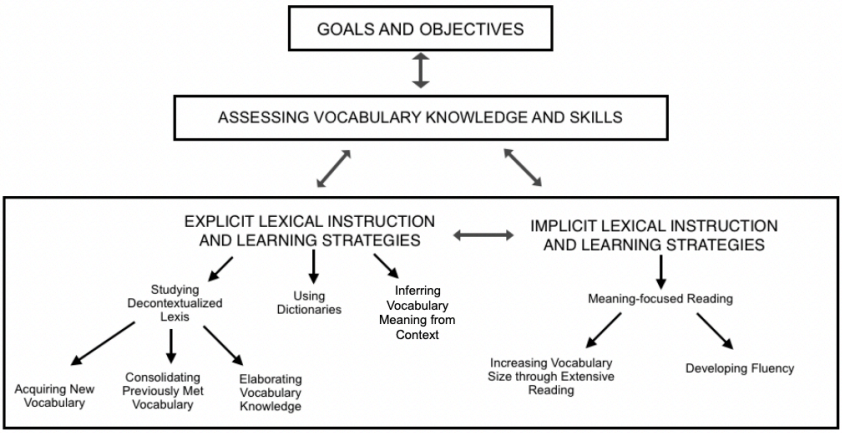
Regarding the explicit/implicit interface, deliberate MWE flashcard learning is commonly classified as an explicit learning strategy that develops explicit knowledge. The current study is unique because it investigates whether deliberate MWE flashcard learning also develops implicit knowledge, which is more commonly associated with incidental learning. The interface has long been a central theme of second language acquisition research that reverberates strongly for language teachers, and Nick Ellis's (2005) review bridged connections to language learning with fields such as psycholinguistics, psychology, neurobiology, and cognitive science. He explained that explicit and implicit neurological processes are physiologically distinct but interact as learners develop their proficiency. Hulstijn (2005) defined and distinguished the interface parameters: implicit and explicit memory, implicit and explicit knowledge, implicit and explicit learning, inductive and deductive learning, and incidental and intentional learning. Rod Ellis (2005) operationalized the explicit/implicit distinction in terms of *awareness, accessibility*, and *use*. He explained that learners are not *aware* of implicit knowledge but are *aware* of explicit knowledge; they *access* implicit knowledge automatically, but *access* to explicit knowledge requires controlled processing; they *use* implicit knowledge in fluent performance, but explicit knowledge is *used* during introspective processing when learners encounter difficulties, plan to write, or make an utterance. Although children tend to learn implicitly, second language acquisition requires teenage and adult learners to develop explicit and implicit knowledge in tandem.

Frameworks for foreign language teaching, lesson planning, course design, and curriculum development often balance explicit and implicit learning. In the Four Strands framework (Nation, 2007), three of the strands develop implicit knowledge (meaning-focused input, meaning-focused output, and fluency development), and one strand develops explicit knowledge (language-focused learning). Textbooks are also designed to balance these two types of learning. Likewise, as Figure 1 shows, Hunt and Beglar (2005) explained how EFL reading program designers set goals, clarified

objectives, and assessed knowledge gains in a curriculum structure built on a balance of explicit and implicit learning strategies.

Figure 1

Explicit and Implicit Learning Strategies in an EFL Reading Curriculum



Note: From "A framework for developing EFL reading vocabulary," by Hunt, A., & Beglar, D., 2005, *Reading in a Foreign Language*, *17*(1), p. 26.

However, although explicit and implicit teaching and learning methods can be balanced in course design, lesson planning, and teaching, implicit knowledge development is rarely the focus of formal assessment. Similarly, explicit learning gains are often investigated in second language acquisition research, but implicit knowledge gains are seldom the focus. This imbalance occurs because implicit knowledge gains are difficult to measure using traditional methods such as pen and paper tests.

Another concern with most interface research is that it has been chiefly focused on grammar acquisition (e.g., DeKeyser, 1997; Norris & Ortega, 2000; Rebuschat & Williams, 2012; Sorace, 2011; Suzuki & DeKeyser, 2017). In contrast, very little research on the implications of the explicit/implicit interface concerning lexical knowledge exists. Sonbul and Schmitt (2013) propose that this neglect of vocabulary interface research might be due to the traditional dictionary metaphor, which regards the mental lexicon as little more than a list of forms and meanings to associate through simple rote learning. Nation's (2013) framework of vocabulary knowledge has helped to overcome the mental dictionary metaphor by showing that know-

ing words entails sophisticated knowledge aspects concerning form, meaning, and use. Nevertheless, Godfroid (2020) explained that this framework concerns explicit language knowledge that can be assessed offline rather than in real-time communicative situations. She transformed it to focus on automaticity, with criteria related to implicit knowledge development. Her framework explains ways to measure the automaticity of form, meaning, and use with real-time methods such as priming experiments, lexical decision tasks, self-paced reading, and eye-tracking. Godfroid (2020) shows how Nation's (2013) criteria may be adapted to consider implicit knowledge development by shifting the focus to real-time processing of form, meaning, and use. Table 1 shows a further adaptation of this framework that focuses on implicit MWE knowledge development criteria. It shows how this experiment measured response times for orthographical and lexical recognition, a narrow slice of the broader lexical knowledge spectrum.

Table 1

Real-Time Lexical Knowledge Aspects Learned with Multiword Expression Flashcards

Knowledge Aspect		**Receptive (R) and Productive (P) Criteria**	
Form	Spoken	R: Does the MWE have auditory representation in memory?	
		P: How rapidly can the MWE be spoken?	○
	Written	R: Does the MWE have an orthographic representation in memory?	○■
		P: How rapidly can the MWE be written or typed?	
	Word parts	R: What word parts are recognizable?	
		P: What word parts can be added or removed?	

Knowledge Aspect		Receptive (R) and Productive (P) Criteria	
Meaning	Form and meaning	R: How rapidly can the MWE's meaning be accessed?	○
		P: How rapidly can MWE be produced to express its meaning?	○
	Concept and referents	R: How rapidly can concepts and referents of the MWE be accessed?	○
		P: How rapidly can the MWE be produced to express a concept?	○
	Associations	R: Has the MWE been integrated into existing semantic networks?	○
		P: How rapidly can associates of the MWE be produced instead?	○
Use	Grammatical functions	R: Is the learner sensitive to the grammar involved with this MWE?	
		P: Can the learner use this MWE in actual conversation?	
	Collocations	R: Are the words of this MWE rapidly recognized?	○■
		P: Can this expression be rapidly produced?	○
	Constraints on use	R: Is the learner aware of constraints on how the MWE is used?	
		P: Can the learner use this MWE correctly?	

Note. ○ = aspects of implicit MWE knowledge developed by practicing with MWE flashcards; ■ = implicit knowledge aspects tested by the current experiment.

Priming to Test for Implicit Knowledge Development

Priming happens when exposure to one stimulus influences a response to a subsequent stimulus without conscious guidance or intention. For example, in semantic priming, the word *table* will be recognized more quickly

when it follows *chair* than *dog* because *table* and *chair* often occur together, and thus neurons associated with these words will fire together. Other types of priming experiments focus on orthography, syntax, or perception. Reber (2013) explained that repetition priming is the most common method for investigating implicit knowledge, which he defined as a form of general plasticity and neural network adaptation. When the brain receives input and internally processes it, it stores the physical structure used. Such structures improve functionality and unconsciously facilitate future cognition.

Priming experiments in second language acquisition research aim to operationalize and measure this facilitation. Standard priming paradigms focus on form processing, grammatical sequencing, meaning interpretation, and lexical associations. When a word, MWE, or construction is learned so well that it primes a related target, it means the language learner has strong, well-integrated knowledge that can be accessed automatically. This automaticity signifies the quality of the knowledge, and evidence of priming illuminates how fluently the knowledge is processed.

Priming Research on Implicit MWE Knowledge Development

At the time of writing, research concerned with implicit MWE knowledge development resulting from flashcard learning was not found. However, Sonbul and Schmitt (2013) conducted a priming experiment to measure implicit knowledge development of technical medical MWEs (*cloud baby, iron lung*) resulting from three different learning conditions. In their enriched condition, participants encountered each MWE three times in a text they read. In the enhanced condition, the MWEs were in the same text but highlighted in red, which made the three encounters more explicit. In the decontextualized condition, learners were presented with the MWEs on PowerPoint slides and told to study them carefully. To test for implicit knowledge development, they conducted a lexical decision task experiment to see if the first words of the MWEs primed the processing of their final words. They did not find significant priming effects and proposed that their experimental learning treatment period was too brief and did not allow for recycling and review, which are needed to develop implicit knowledge. However, their explicit measures showed that all three learning conditions led to significant long-term recall and recognition. Their experiment demonstrates how readily explicit knowledge gains can be measured but how difficult it is to measure implicit knowledge gains.

In a replication and extension of Sonbul and Schmitt (2013), Toomer and Elgort (2019) tested the incidental reading conditions (reading only,

bolding, and bolding plus glossing) with more participants and more time on task. The results of the primed lexical decision task only showed initial evidence of implicit knowledge development when the collocations were presented without enhancement. Their main finding was that repeated encounters with collocations in reading promoted the development of collocational knowledge. Bolding led to the development of explicit knowledge, and the absence of typographic enhancement promoted the development of implicit knowledge.

However, Toomer and Elgort (2019) did not replicate Sonbul and Schmitt's (2013) decontextualized explicit condition, which was most relevant to this current study. Furthermore, in this current experiment, the learners were each given their own sets of flashcards so that they could remove the MWEs they had learned and reshuffle them to enhance memorization. Elgort (2011) conducted encouraging research regarding implicit knowledge development from flashcard learning for single words (pseudowords). In Experiment 2, she conducted a masked repetition priming experiment displaying a mask (#######) for 522 ms, followed immediately by a pseudoword prime (e.g., "forfert") for 56 ms, and then a target ("FORFERT") for 522 ms. The participants made lexical decisions regarding the targets they had just seen while looking at the blank screen. They were instructed to treat the newly learned pseudowords as English words and answer YES for the lexical decision. This experiment showed that identity primes had a facilitation effect, 52 ms faster than the controls. These results indicated that learning the pseudowords with flashcards resulted in acquiring orthographic representations in implicit knowledge. That is, the quality of the knowledge of the newly learned pseudowords was strong enough to prime the targets that followed. In this current experiment, a masked repetition priming lexical decision task very similar to Elgort's (2011) Experiment 2 was employed to investigate changes in the quality of subconscious representations of the MWEs that participants learned with flashcards.

Obermeier (2022) measured semantic association gains in a self-paced reading experiment that likewise compared the effects of flashcard learning on literal and figurative MWEs. That experiment primarily focused on investigating the semantic components in Table 1. Like the current research, results in that experiment were analyzed in a repeated measures linear mixed-effects model with participants and items as crossed random effects. No statistically significant interaction for semantic association gains were found, as measured in an innovative priming paradigm wherein semantically related words that followed the MWEs in sentences were compared. Although the interaction

was not significant, the semantic associates of literals were processed faster than those of figuratives in all three conditions: pre-test, learned post-test, and not-learned post-test. In a separate analysis of the data, a statistically significant interaction showed that deliberate learning resulted in substantial formulaic sequencing gains for literals but no such gains for figuratives. Obermeier (2022) concluded that the learning treatment was too brief to result in the strong semantic acquisition of the figuratives because of their high learning burden. This current research aims to complement findings from that self-paced reading experiment by investigating the effects of flashcard learning on orthographic MWE representations.

Methodology

This investigation focused on implicit knowledge development of literal and figurative MWEs, operationalized by response times in a masked repetition priming lexical decision task. The first research question was: *Does multiword expression flashcard learning develop implicit multiword expression knowledge?* As flashcard learning entails highly focused repetition and retrieval of meaning and form, it was hypothesized that implicit knowledge gains for both MWE types would be statistically significant. The second research question was: *Does implicit knowledge develop differently for the flashcard learning of literal and figurative expressions?* Because learning figuratives is more difficult than learning literals, it was expected that figuratives would be processed more slowly on the pretest. Flashcard learning should result in greater gains for figuratives when the meaning/form connection is established.

Participants

The study's participants (N = 43) were 21 male and 22 female students at a small national teacher training university in Japan. All had studied English for 4 to 8 hours a week for six years in junior high and high school in reading, writing, speaking, listening, and grammar courses. Their ages ranged from 19 to 22. They were enrolled in their first or second year of studies in the English Education Department, training to become elementary, junior high, or high school English teachers. Participants were in two intact classes, 26 in one class and 28 in the other (a convenience sample of 54). Teacher-training students often need to be absent from class for practicum training. For this reason, 11 participants missed one or more classes during the experiment and were excluded from the data analysis.

Soon after beginning their first year of studies, all students took the Global Test of English Communication (GTEC), designed for Japanese university and high school students. Their average total score was 623 (*SD* = 71.89), which, according to the GTEC instructional materials, classified them as Advanced Learners, the second-highest category of the test. Mean reading scores were 241 (*SD* = 29.32), earning them a level of assessment at which "reading a newspaper article with the occasional support of a dictionary is possible." The accompanying materials also state that the approximate TOEIC equivalent is 600, the approximate paper-based TOEFL equivalent is 480, and the approximate Internet-based TOEFL (iBT) is 60. Thirteen of the participants had studied English abroad for four weeks or more. The participants' motivation to learn English was high because they intended to eventually teach it professionally.

Before the experiment, the researcher explained the following three points verbally in English and then in writing in Japanese: (a) their participation in the study was optional; (b) their participation or lack of participation would have no effect on their grade; (c) no personal information would ever be shared. After they finished the experiment, they were debriefed on the purposes of the investigation and preliminary findings. Participants were also given a small gift as a token of appreciation and acknowledgment of their efforts.

Procedures

The experiment was conducted once weekly over five weeks. The primary experimental condition, flashcard learning, was counterbalanced across the two groups of participants. In Table 2, the schedule of the experiment is outlined.

Table 2
Schedule of the Experiment

Session	Minutes	Activity
Week 1	10	Introduction to the experiment
Week 2	30	Masked priming lexical decision pretest
Week 3	40	MWE flashcard learning
Week 4	30	Masked priming lexical decision posttest
Week 5	10	Debriefing

Learning Materials

The experimental materials and instruments were made using a list of 48 MWEs, 24 figuratives, and 24 literals. The target items were selected by two native speakers, who discussed each MWE and categorized it as literal or figurative according to how directly the constituent words matched the overall meaning. A third native speaker confirmed the literal/figurative categorizations. Next, the researcher matched the MWEs with Japanese meanings, and these paired associates were shown to four English learners who were not participants in the experiment to confirm whether the form-to-meaning connections made sense. For example, *next door* was matched to the Japanese meaning 隣の and confirmed. The literal and figurative expressions were counterbalanced across Study Lists A and B to create critical comparisons among the conditions. Therefore, each participant learned 24 MWEs: 12 literal and 12 figurative. All experimental contrasts were made on the items within participants. If participants in one group learned an MWE, the other group did not. Some examples of the paired-associates and the counterbalancing structure for the study lists are shown in Table 3.

Table 3

Examples from the List of Figuratives, Literals, and Japanese Meanings

MWE Composition	Study List	MWE	Japanese
Literal	A	above all	最も
		stay away	避ける
		take place	起こる
	B	deal with	扱う
		feel like	欲しい
		take it easy	のんびりする
Figurative	A	set out	始まる
		sinking ship	絶望
		play hardball	真剣
	B	high handed	攻撃的
		can of worms	複雑
		make waves	迷惑

Participants learned the MWE/Japanese pairs in the experimental treatment in which the English target was typed on one side of a piece of paper, and the Japanese meaning was typed on the other. They were also given a guidance sheet explaining the following instructions (written in Japanese) about the flashcard learning strategy. Before they started studying, the following were explained orally: (1) Practice with 8 MWEs at a time; (2) Recall the Japanese meanings from the MWEs; (3) When recalling the English MWE from the Japanese meaning, say it aloud; (4) When you feel you have learned an MWE well, remove the card; (5) When you remember the first 8 MWEs, add 8 more and study all 16 together; (6) After you remember these 16 MWEs, add the final 8 and study all 24 of them. Participants were given 20 minutes to study independently. Time announcements were made when 10, 5, and 1 minute(s) remained.

The Masked Repetition Priming Lexical Decision Task

The masked priming lexical decision task was conducted in a computer-assisted language learning classroom containing 48 Hewlett Packard Compaq® dc7700 desktop computers with 2.13 GHz Intel Core Duo® processors, displayed on 21.5-inch Iodata® liquid crystal display monitors. It was created using E-prime®, software for developing psychological experiments (Schneider et al., 2002). The pre-test and post-test each took approximately 30 minutes for participants to complete. Before beginning the actual task, participants did 20 practice trials to become familiar with the procedure.

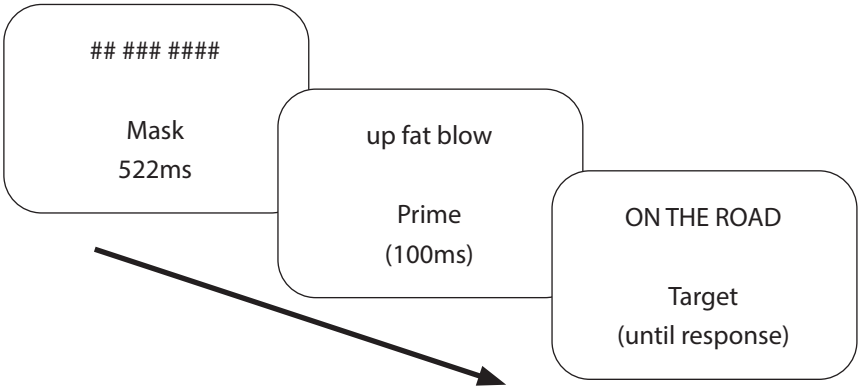
The trial format is presented in Figure 2. Each mask, prime, and target word had the same number of characters as the corresponding word on the next slide. For example, in Figure 2, the mask's ##, the prime's *up,* and the target's *on* have two characters. Likewise, ###, *fat,* and *the* each have three characters, and so on. The mask was presented for 522 ms (slightly over half a second), followed by a prime that was presented for 100 ms (one-tenth of a second). This very short prime presentation time was crucial: it was brief enough to prevent conscious processing yet long enough to stimulate subconscious processing. After the experiment was finished, when asked what they saw, participants said the targets seemed to be slightly blurry at first but then came into focus, confirming the subconscious presentation paradigm.

The target was presented until a response was received. The hypothesis was that if the identity prime (in this example, "on the road", not shown in the figure) had been acquired in the flashcard learning treatment, it would facilitate the processing of the target ON THE ROAD faster than the unrelated

prime "up fat blow". Participants pressed different buttons to make lexical decisions on the targets, answering *YES* if all the words were English or *NO* if one or more words were not English. Figure 2 is an example of a trial in the unrelated priming condition wherein the correct response to the lexical decision was YES. In masked repetition priming, identity primes consistently facilitate the processing of the targets they precede. This procedure has had robust effects and has helped to understand subconscious lexical recognition processes (e.g., Adelman et al., 2014; Forster & Veres, 1998; Grainger, 1998). If an MWE is established in the mental lexicon, an identity prime will subconsciously pre-activate its lexical representation, and the target will be processed more fluently.

Figure 2

Example Trial for the Masked Repetition Priming Lexical Decision Task.



In each trial of the lexical decision task, participants decided whether all the words of the MWE target were English or not. The 144 targets were balanced half and half between 72 intact MWEs and 72 non-word MWEs. For each participant, 24 of the 72 intact targets had been learned, 24 had not, and another 24 were fillers (added to lower the percentage of critical trials and further prevent strategic processing). Although the nonword MWE trials were essential distractors for the lexical decision task, these were excluded from the analysis. In this way, 24 learned and 24 not-learned targets were the critical trials for the experiment and the focus of the analysis.

As a rule of thumb, Brysbaert and Stevens (2018) recommend that an adequately powered reaction time experiment has at least 1,600 observations

per condition. In this experiment, 43 participants, 48 critical stimuli, and 2 test sessions yielded 4128 observations. Regarding research question 1, for the three conditions tested (Learning, Priming, and MWE Composition), each condition had 1376 observations, so the experiment had 86% of the observations needed to meet that criterion. Further power analyses by simulation were conducted using the *simR* package in R (Green & MacLeod, 2016). Based on 200 simulations, the *powerSim* function revealed that both Learning and Priming conditions had 100% of the statistical power necessary, but MWE Composition had only 50% of the power necessary. In the *powerSim* analysis for Research Question 2, in which 1972 literal observations and 1937 figurative observations were analyzed separately in a simpler model, Learning and Priming conditions both had 100% of the statistical power needed.

An example of each trial type (excluding the filler trials) is shown in Table 4. Every critical trial was tested under one level of all three two-leveled conditions: (a) Learning: Learned versus Not-Learned; (b) MWE Composition: Literal versus Figurative; (c) Priming: Identity vs. Unrelated. These contrasts were created in the trial list, in which each MWE was tested once per participant in one or the other level of each condition. That is, all participants experienced the same conditions on different MWEs. The experiment was a series of comparisons between conditions on items.

Table 4

Item Types for the Masked Repetition Lexical Decision Task

Lexical Decision	Primes		
	List A	List B	Targets
YES	dog eat dog (identity)	off the cat (unrelated)	DOG EAT DOG (intact figurative MWE)
	the fly door (unrelated)	all the time (identity)	ALL THE TIME (intact literal MWE)
NO	abobe all (identity)	sheep the (unrelated)	ABOBE ALL (nonword MWE)
	teh od nemes (unrelated)	han of wobes (identity)	HAN OF WOBES (nonword MWE)

Results

Data were collected for 4060 observations, but an initial phase of outlier trimming removed invalid trials. Baayen (2008) explains that extremely fast response times (RTs) signify non-engaged, automatic button-pushing, and extremely slow responses signify confusion or distraction. Accordingly, 80 observations (1.97% of the data) were removed with response times below 200 ms or above 4000 ms. Mean response times (RTs) for all conditions are shown in Table 5, and some comparisons of interest are as follows. For All Trials, mean RTs for the Learned trials (1303.57 ms) were 180.03 ms faster than the Not-learned trials (1483.60 ms). Regarding Priming, for Learned trials, MWE targets in the Identity Priming condition (1229.83 ms) were processed 146.88 ms faster than Learned MWEs in the Unrelated condition (1376.71 ms). Regarding MWE Composition, Literal expressions that were Learned (1264.20 ms) had 79.05 ms faster RTs than Figurative expressions that were Learned (1343.25 ms). Although these mean differences help to describe general trends in the data, more sophisticated modeling is required to interpret the effects of the Learning, Priming, and MWE Composition conditions and their interactions.

The analysis of the crossed linear mixed effects model was conducted using the lmer package in the R environment for open-source statistical software (Bates et al., 2015). The analysis followed the top-down model building strategy specified by West, Welch, and Gateki (2015). The first step was to confirm whether a random effects structure should be included. To test this hypothesis, a "loaded mean" structure containing both fixed and random effects was compared with a model containing only fixed effects (West et al. 2015, p. 66). The ANOVA comparison assessing whether the added random effect variances were zero was rejected with a p-value less than 0.0001, which indicated that the model including the random effects should be included for all subsequent stages of model building.

The random-effects specification was improved through the inclusion and exclusion of different structures. Subjective model comparisons included comparison of Akaike's Information Criterion and the Bayesian Information Criterion. Objective comparisons were assessed using likelihood ratio tests of models using results in the lmer output until the best model was identified (Baayen, 2008; Baayen et al., 2008). The inclusion of random intercepts (Participant and Item) and random slopes (Trial Order and Learning) were judged to best capture the overall random effects structure.

In the initial model, the Learning, Priming, and MWE Composition conditions were all statistically significant fixed-effect predictors. The constant

Table 5

Response Times for the Masked Priming Repetition Lexical Decision Task (Milliseconds)

	Learning Condition		
	Pretest	Not-learned	Learned
M	1611.86	1483.60	1303.57
SEM	16.91	21.64	17.93
n	1947	978	984
SD	746.57	676.93	562.48
95% CI Lower	1578.68	1441.12	1268.38
95% CI Upper	1645.05	1526.06	1338.76

	Priming Condition					
	Identity			Unrelated		
	Pretest	Not-learned	Learned	Pretest	Not-learned	Learned
M	1571.02	1433.63	1229.83	1651.68	1532.97	1376.71
SEM	22.10	31.08	30.95	21.82	30.89	30.83
n	961	486	490	986	492	494
SD	763.69	655.05	537.39	727.68	695.03	577.61
95% CI Lower	1527.68	1372.69	1169.14	1608.89	1472.40	1316.27
95% CI Upper	1614.35	1494.56	1290.51	1694.46	1593.53	1437.15

	MWE Composition					
	Literal			Figurative		
	Pretest	Not-learned	Learned	Pretest	Not-learned	Learned
M	1500.47	1382.88	1264.20	1725.69	1586.41	1343.25
SEM	21.35	27.83	26.32	24.71	35.91	26.75
n	984	494	494	963	484	490
SD	698.26	690.81	562.62	776.91	741.40	560.12
95% CI Lower	1461.13	1328.07	1212.64	1678.63	1526.51	1288.53
95% CI Upper	1546.08	1441.28	1314.59	1776.21	1656.22	1394.07

variance, linearity, independence, and normality assumptions were assessed using the *mcp* function in R's *LMERConvenienceFunctions* package. In this initial model, the distribution of the residuals had a severe negative skew and a very long positive tail, so 71 positive and negative outliers (1.78% of the data) were trimmed, resulting in a bell-shaped distribution of the residuals that resembled the normal distribution.

After confirming differences between the levels of the independent variables, the next step was to investigate the interactions of interest, as specified per the experimental hypotheses. The interaction between Priming (Identity versus Unrelated) and Learning (Pretest, Learned, or Not-learned) tested how flashcard learning affected the RTs of the different prime types. The interaction between MWE Composition (Literal versus Figurative) and Learning tested how flashcard learning affected their RTs differently. The notation for the specification of the final model was as follows:

Fixed Effects:

RT ~ Priming*Learning + MWEComposition*Learning

Random Effects:

(1 + TrialOrder + Learning | Participant) + (1 | Target)

The results of the random and fixed effects for the model are shown in Table 6. The intercept represents the reference levels of the independent variables: Unrelated Priming, Pre-test Learning Condition, and Figurative MWE Composition. The other estimates are in comparison to the intercept level. For simple effects, pairwise effect size calculations were made following Brysbaert and Stevens (2014).

The focus of the investigation for Research Question 1 was the statistically significant interaction explained in Table 6 between the Learned condition and Priming (β = -0.04; t = -2.42; $p < .05$) and how it contrasts with the non-significant result for interaction between the Not-learned condition and Priming. This difference indicates that the MWEs were learned well enough to produce priming effects through flashcard learning. For the simple effects, the estimate associated with the Learned (Posttest) condition (β = -0.14; t = -5.11; d = -0.34) was larger than that for the Not-learned condition (β = -0.07; t = -2.93; d = -0.17). These results, as well as the small but statistically significant Learned x Priming interaction (β =0.038; t = -2.42, p > .05) were evidence of priming effects associated with flashcard learning. As the reference level was Pre-test, the small effect for the Not-learned condition indicates testing effects.

Table 6
Linear Mixed-Effects Model for the Masked Repetition Priming Lexical Decision Task

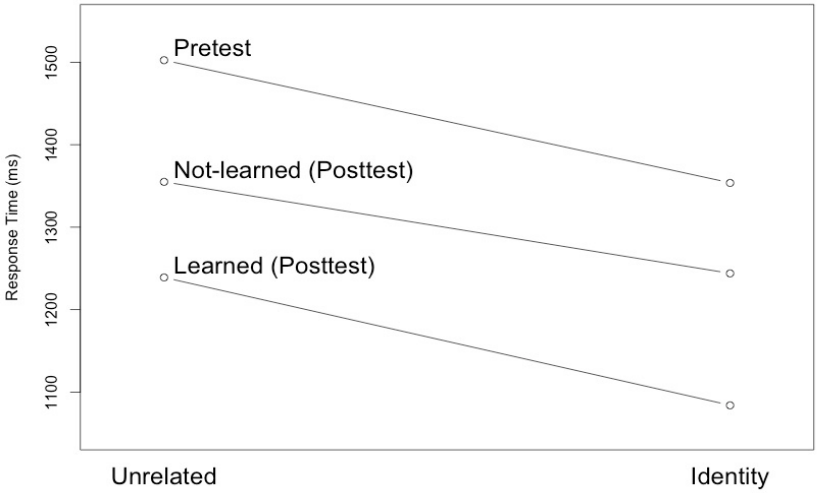
Random Effects					
	Variance	*SD*			
Target (Intercept)	0.01	0.14			
Participant (Intercept)	0.05	0.23			
Trial order (Slope)	0.01	0.96			
Learned (Slope)	0.03	0.18			
Not-learned (Slope)	0.02	0.14			
Residual	0.05	0.22			
Fixed Effects					
		β	*SE*	*t* value	*d*
Intercept		-0.67	0.44	-15.26*	
Priming (identity)		-0.73	0.10	-7.23*	-0.18
Learned (Posttest)		-0.14	0.28	-5.11*	-0.34
Not-learned (Posttest)		-0.07	0.25	-2.93*	-0.17
MWE Composition (Literal)		-0.11	0.41	-2.61*	-0.31
Learned x Priming		-0.04	0.02	-2.42*	
Not-learned x Priming		-0.01	0.02	0.42	
Learned x MWE Composition		0.04	0.02	2.17*	
Not-learned x MWE Composition		0.03	0.02	1.59	

* $p < .05$

Figure 3 shows faster RTs for Identity priming for all three learning conditions. It also shows incremental facilitation overall as exposure increases because the Learned trials are faster than the Not-learned trials, and both are faster than the Pretest baseline.

Figure 3

Learning and Priming Conditions for all Multiword Expressions



Research Question 2 concerned whether flashcard learning affects Literal and Figurative expressions differently. Table 6 above shows the statistically significant interaction between MWE Composition and Learning ($\beta = 0.04$; $t = 2.17$, $p < .05$), indicating that flashcard learning had different effects on the Figurative and Literal targets. To better understand the effects of learning conditions, priming conditions, and MWE Composition, separate investigations were conducted on the literals and figuratives by specifying the following model for each:

Fixed Effects:

RT ~ Priming*Learning

Random Effects:

(1 + Trial Order + Learning | Participant) + (1 | Target)

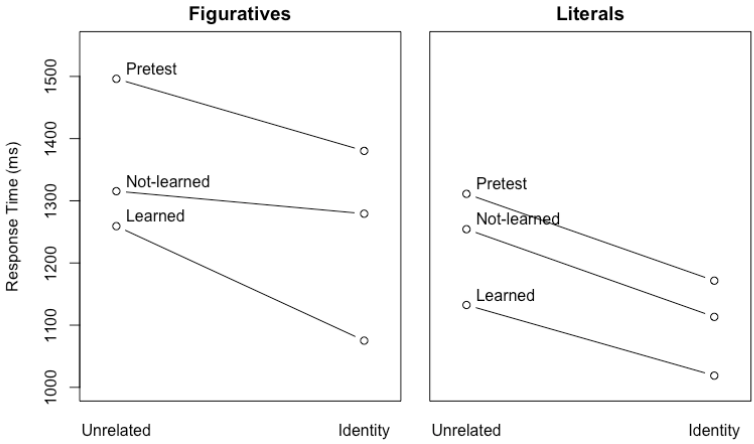
As in the previous analysis, the model structure was confirmed step by step, and the interaction between Priming and Learning was tested for significance.

The interaction between Priming and Learning was statistically significant for Figuratives but not for Literals. Figure 4 below shows the different effects of learning on the two MWE types. For Literals, the RTs for targets in Identity and Unrelated priming conditions decrease in equal progression

for both the Learned and Not-learned conditions, showing that the priming effects between Unrelated and Identity primes were the same at Pre-test and Post-test in both the Learned and Not-learned conditions. Furthermore, the overall pre-test to post-test response time changes for the Literals (from around 1350 ms at the Unrelated Pretest to 1000 ms at Identity Learned Postest) shows that these were processed more consistently and faster than Figuratives (which changed from around 1500 ms to 1100 ms over the same conditions). Most importantly, comparing the Figurative Not-learned and Learned line pairs with the corresponding Literal line pairs shows the dramatic difference in effects that deliberate learning had on these two different types. Learned Figuratives had greater gains in Identity Priming effects than Learned Literals.

Figure 4

Interactions between Priming and Learning Conditions for the Literal and Figurative Expressions



The statistical results confirmed the effects shown in Figure 4. For Figurative expressions, the interaction between Learning and Priming was statistically significant. For the Learned Figuratives, Identity priming was significantly faster than Unrelated priming (β = -3.14; t = -3.037; p = .002). Thus, the line is steeply sloped. For the Not-learned Figuratives, this difference was not significant (as shown by the nearly horizontal line). Identity and Unrelated priming effects were constantly incremental for the Learned

and Not-learned Literals, as shown by the nearly parallel three lines on the right side of Figure 4.

Discussion and Conclusion

For the full data set with all the MWEs, the interaction between Priming and Learning conditions was statistically significant for the Learned targets but not for those in the Not-learned condition. These different effects confirmed Research Question 1, showing that flashcard learning resulted in strong Identity priming effects, evidence of facilitated subconscious orthographic processing for MWEs overall. Separate analyses of Literals and Figuratives were conducted to investigate Research Question 2. For Figurative expressions, the statistically significant interaction between Learning and Priming conditions showed priming effects for Learned targets but not for Not-learned targets. However, no such priming effects were found for the Literal expressions. Together, these results showed that learning with flashcards through repetition and retrieval facilitated the development of implicit orthographic knowledge for all MWEs, but the effects of learning were more substantial for Figurative expressions, which have a heavier learning burden.

Although this research had some valuable findings, its limitations must also be mentioned. First, although prior learning was accounted for with a pretest, this is not standard in priming research because of the strong tendency to produce testing effects. A better way to control for prior MWE knowledge would be to use highly specialized unknown targets like the medical MWEs that Sonbul and Schmitt (2013) used. The second limitation was the convenience sample. Severe participant attrition resulted in insufficient statistical power to test the full model for Research Question 2, so a second separate analysis was required with a simpler model without the MWE Composition predictor. Another constraint resulting from the convenience sample (taken during scheduled class time) was limited time on task. A third limitation is also concerned with time on task. As the participants were guided to remove flashcards once they remembered them, they may have removed literal flashcards sooner than figurative ones, and this imbalance of study time may have influenced the results.

Keeping these limitations in mind, it nevertheless seems fair to argue that the pedagogical implication of these findings is that flashcard learning benefits the learning of figuratives but not literals. For both MWE types, automatic orthography and word sequence recognition must be developed. However, literal and figurative semantic compositions entail different learn-

ing processes. For literals, the direct meaning-to-form connection for each word is also automatized each time it is encountered, meaning their integration into the mental lexicon is straightforward. Contrastingly, when learning a figurative expression incidentally, the learner must reject the direct semantic interpretation of each word, and this is not possible until the whole figurative expression is recognized (Cieślicka, 2006; 2012). Thus, processing figurative expressions entails the extra steps of rejecting individual word meanings and then learning the metaphor of the whole expression. These two additional steps seem to hinder the development of automaticity for figuratives.

Researchers have explored explicit learning methods for deeply processing figurative expressions such as focusing on etymology (Boers, Eyckmans, & Stengers, 2007), cognitive semantics (Boers, 2011), and pictorial elucidation (Boers et al., 2009). Such methods may entail rich and thoughtful processing that fosters durable associations, but they may also require learning superfluous explicit knowledge that cannot be applied in real-time communicative situations. In deliberate flashcard learning, such deep processing is not the aim. Instead, the strategy aims to automate the association of the metaphorical meaning to the whole expression through repetition and retrieval.

Strategic flashcard learning of the PHRASE List (Martinez & Schmitt, 2012) is undoubtedly an effort wisely spent, as these frequent MWEs will likely be encountered in natural English use. Abundant, thematic flashcard learning focused on specialized MWEs found in accompanying texts would balance explicit and implicit learning strategies as supported by research cited herein (Hunt & Beglar, 2005; Sonbul & Schmitt, 2013; Toomer & Elgort, 2019). Flashcard learning of figuratives entails bypassing misleading (yet normal) individual word processing and automatizing the connection between the whole expression and its metaphorical meaning. Whether learned incidentally or learned deliberately as a whole, each encounter with a literal expression entails processing facilitation. In sum, findings from this study call for strategic flashcard learning of frequent MWEs with opaque meanings accompanied by massive exposure that will provide incidental learning opportunities.

Andrew Obermeier is an associate professor at Kyoto University of Education. His research interests include vocabulary acquisition and positive psychology in language learning.

References

Adelman, J. S., Johnson, R. L., McCormick, S. F., McKague, M., Kinoshita, S., Bowers, J. S., Perry, J. R., Lupker, S. J., Forster, K. I., Cortese, M. J., Scaltritti, M., Aschenbrenner, A. J., Coane, J. H., White, L., Yap, M. J., Davis, C., Kim, J., & Davis, C. J. (2014). A behavioral database for masked form priming. *Behavior Research Methods, 46*(4), 1052–1067. https://doi.org/10.3758/s13428-013-0442-y

Baayen, H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.

Baayen, H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition, 1*(9), 1–20. https://doi.org/10.5334/joc.10

Boers, F. (2011). Cognitive semantic ways of teaching figurative phrases: An assessment. *Review of Cognitive Linguistics, 9*(1), 227–261.

Boers, F., Piquer Píriz, A. M., Stengers, H., & Eyckmans, J. (2009). Does pictorial elucidation foster recollection of idioms? *Language Teaching Research, 13*(4), 367–382. https://doi.org/10.1177/1362168809341505

Boers, F., Eyckmans, J., & Stengers, H. (2007). Presenting figurative idioms with a touch of etymology: More than mere mnemonics? *Language Teaching Research, 11*(1), 43–62. https://doi.org/10.1177/1362168806072460

Cieślicka, A. (2006). Literal salience in online processing of idiomatic expressions by second language learners. *Second Language Research, 22*(2), 115–144. https://doi.org/10.1191/0267658306sr263oa

Cieślicka, A. (2012). Do nonnative language speakers chew the fat and spill the beans with different brain hemispheres? Investigating idiom decomposability with the divided visual field paradigm. *Journal of Psycholinguistic Research, 42*, 475–503. https://doi.org/10.1007/s10936-012-9232-4

DeKeyser, R. M. (1997). Beyond explicit rule learning. *Studies in Second Language Acquisition, 19*(2), 195–221.

Elgort, I. (2011). Deliberate learning and vocabulary acquisition in a second language. *Language Learning, 61*(2), 367–413. https://doi.org/10.1111/lang.2011.61.issue-2

Ellis, N. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition, 27*(2), 305–352.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition, 27*(2), 141–172.

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text & Talk, 20*(1), 29–62.

Forster, K. I., & Veres, C. (1998). The prime lexicality effect: Form-priming as a function of prime awareness, lexical status, and discrimination difficulty. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(2), 498.

Goatly, A. (2011). *The language of metaphors* (2nd ed.). Routledge.

Godfroid, A. (2019). Sensitive measures of vocabulary knowledge and processing: Expanding Nation's framework. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (1st ed., pp. 433–453). Routledge.

Grainger, J. (1998). Masked priming by translation equivalents in proficient bilinguals. *Language and Cognitive Processes, 13*(6), 601–623. https://doi.org/10.1080/016909698386393

Grant, L., & Bauer, L. (2004). Criteria for re-defining idioms: Are we barking up the wrong tree? *Applied Linguistics, 25*(1), 38–61.

Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution, 7*(4), 493–498. https://doi.org/10.1111/2041-210x.12504

Hulstijn, J. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning: Introduction. *Studies in Second Language Acquisition, 27*(2), 129–140.

Hunt, A., & Beglar, D. (2005). A framework for developing EFL reading vocabulary. *Reading in a Foreign Language, 17*(1), 23–59.

Kjellmer, G. (1994). *A dictionary of English collocations: Based on the Brown corpus*. Clarendon Press.

Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics, 33*(3), 299–320. https://doi.org/10.1093/applin/ams010

Nakata, T. (2011). Computer-assisted second language vocabulary learning in a paired-associate paradigm: A critical investigation of flashcard software. *Computer Assisted Language Learning, 24*(1), 17–38.

Nation, P. (2007). The four strands. *Innovation in Language Learning & Teaching, 1*(1), 2–13. https://doi.org/10.2167/illt039.0

Nation, P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*(3), 417–528.

Obermeier, A. (2022). The effects of learning multiword expressions deliberately: Implications regarding implicit knowledge development. *JACET Kansai Journal, 24*, 96–115.

Reber, P. J. (2013). The neural basis of implicit learning and memory: A review of neuropsychological and neuroimaging research. *Neuropsychologia, 51*(10), 2026–2042. https://doi.org/10.1016/j.neuropsychologia.2013.06.019

Rebuschat, P., & Williams, J. (2012). Implicit and explicit knowledge in second language acquisition. *Applied Psycholinguistics, 33*(4), 829–856. https://doi.org/10.1017/S0142716411000580

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime reference guide*. Psychology Software Tools.

Sinclair, J. M. (1991). *Corpus, concordance, collocation*. Oxford University Press.

Sinclair, J. M. (1995). *COBUILD English collocations*. Harper Collins.

Siyanova-Chanturia, A. (2017). Researching the teaching and learning of multi-word expressions. *Language Teaching Research, 21*(3), 289–297. http://doi.org/10.1177/1362168817706842

Siyanova-Chanturia, A., & Martinez, R. (2015). The idiom principle revisited. *Applied Linguistics, 36*(5), 549–569. https://doi.org/10.1093/applin/amt054

Sonbul, S., & Schmitt, N. (2013). Explicit and implicit lexical knowledge: Acquisition of collocations under different input conditions. *Language Learning, 63*(1), 121–159. https://doi.org/10.1111/j.1467-9922.2012.00730.x

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism, 1*(1), 1–33. https://doi.org/10.1075/lab.1.1.01sor

Suzuki, Y., & DeKeyser, R. (2017). The interface of explicit and implicit knowledge in a second language: Insights from individual differences in cognitive aptitudes. *Language Learning, 67*(4), 747–790. https://doi.org/10.1111/lang.12241

Thorndike, E. L. (1908). Memory for paired associates. *Psychological Review, 15*(2), 122.

Toomer, M., & Elgort, I. (2019). The development of implicit and explicit knowledge of collocations: A conceptual replication and extension of Sonbul and Schmitt (2013). *Language Learning, 69*(2), 405–439. https://doi.org/10.1111/lang.12335

Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning, 61*(2), 569–613. https://doi.org/10.1111/lang.2011.61.issue-2

West, B., Welch, K., & Gateki, A. (2015). *Linear mixed models: A practical guide using statistical software* (2nd ed.). CRC Press.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.

A Contrastive Interlanguage Analysis of Lexical Bundles in English as a Foreign Language Writing: L1 Chinese, Japanese, and Korean

Randy Appel
Waseda University

Lewis Murray
Kanazawa University

Correct and register-appropriate use of frequently recurrent word sequences (e.g., lexical bundles) plays an important role in proficient linguistic output. However, second language (L2) writers' use of these multiword items is still insufficiently understood, particularly in relation to the influence of first language (L1) background. This exploratory study analyzed a learner corpus of 420 argumentative essays to determine how lexical bundles were used by L2 English academic writers from 3 L1 backgrounds (Chinese, Korean, Japanese) to identify intra-group tendencies and intergroup production differences. A contrastive interlanguage approach identified unique tendencies related to functional categories and individual lexical bundles for each L1 group. Findings include relative overuse of text-oriented bundles by L1 Chinese writers, overuse of participant-oriented bundles by L1 Japanese writers, and a general tendency to underuse of lexical bundle types and tokens by L1 Korean writers of L2 English. Methodological and pedagogical implications of these findings are discussed.

Lexical bundles(単語連鎖)のように高頻度で使用される語の連続を、適切に、正しいレジスターで用いることは、熟練した言語の産出に不可欠である。しかし、第二言語学習者によるその使用の実態は、特に母語の影響との関連では十分に理解されていない。本研究では、420の学術的文章からなるコーパスの分析を通して、母語背景(中国語、韓国語、日本語)をもつ英語の第二言語学習者によるlexical bundlesの使用を調べ、母語集団内の傾向や集団間でのその産出の違いを明らかにする。対照中間言語分析によって、談話機能や個別のlexical bundlesに関して、

https://doi.org/10.37546/JALTJJ45.1-2

JALT Journal, Vol. 45, No. 1, May 2023

各母語集団独自の傾向があることがわかった。研究の成果として、中国語母語の書き手がテキスト志向のbundlesを、日本語母語の書き手が参与者志向のbundlesを、それぞれ比較的多用する傾向があることや、一般に韓国語母語の書き手にはlexical bundlesのタイプやトークンの使用頻度が低い傾向がみられることを示す。これらの方法論的・教授法的含意についても論じる。

Keywords: formulaic sequences; L1 differences; multiword structures; second language writing

Research into the use of Formulaic Sequences (FSs), defined as multiword structures believed to be stored and produced as single units (Wray, 2002) has grown exponentially over the past several decades. These studies cover a range of foci, yet one of the main findings is that FSs play a vital role in fluent and proficient language use (e.g., Chen, 2019; Wray, 2002). Despite their importance, contrastive research into how these structures are used by second language (L2) English writers of differing first language (L1) backgrounds is limited. In particular, studies examining use of lexical bundles (LBs), a frequency-based approach to the identification of FSs, by writers of varying L1 backgrounds are rare. Furthermore, studies that have examined LBs in this way have often contained methodological issues preventing distinctions between L1 specific and universal tendencies.

As a result, further studies examining more closely controlled and comparable corpora are needed to better understand how LBs are employed by L2 English writers of varying L1 backgrounds. Findings may lead to improved identification and distinction of L1-related and universal production tendencies that could be used to inform pedagogic interventions aimed at English as a Foreign Language (EFL) and English as a Second Language (ESL) users. The current exploratory study therefore analyzed 420 EFL essays by writers of three distinct L1 backgrounds: Chinese, Japanese, and Korean, with English proficiency, writing conditions, and topic controlled across groups.

Literature Review

Formulaic Sequences & Lexical Bundles

FSs are increasingly recognized as a crucial aspect of fluent and proficient language use, due in large part to the widespread use of corpus informed research that has helped drive growth in this area. Thus far, scholars have shown that FSs are prevalent in L1 speech and writing (e.g., Schmitt, 2004), aid in perceptions of fluency (e.g., Wray, 2002), and ease processing and production burdens associated with unplanned linguistic output (e.g., Kuiper, 1996).

Whereas definitions of FSs can vary depending on the goals of the researcher, LBs carry a more stable definition that results in greater interstudy comparability. Introduced by Biber et al. (1999), LBs are simply defined as multi-word strings (often four words in length) that meet minimum frequency and range criteria. This quantitative focus means LBs often cross semantic and syntactic boundaries and may not hold the same psycholinguistic status as wholly stored and produced FSs. However, like FSs, LBs contribute to perceptions of linguistic proficiency (e.g., Shin, 2019) and distinguish L1 from L2 users (e.g., Lu & Deng, 2019).

The LB approach has grown from a relatively niche method to one widely used to analyze L1 and L2 discourse across a range of genres and registers. Examples include Hyland (2008), who used a 3.5-million-word corpus of academic texts to reveal production tendencies that help distinguish scholarly disciplines (e.g., engineering, biology, business, applied linguistics). Similarly, Durrant (2017) analyzed the British Academic Written English (BAWE) corpus identifying distinct LB production patterns that differentiate hard and soft sciences.

Lexical Bundles in L2 English Writing

As with many forms of corpus-driven/based research, early LB studies commonly focused on L1 English discourse. However, this focus has gradually shifted to examine structures by L2 English writers of varying proficiencies (e.g., Appel & Wood, 2016; Chen, 2019). In general, these studies aim to identify production patterns that distinguish high- and low-level writing, with findings informing teaching interventions aimed at improving proficiency. For example, Chen (2019) used a large collection of essays from the International Corpus Network of Asian Learners of English (ICNALE) to reveal that EFL writers at higher proficiency levels used a wider range of LBs in their texts. Similarly, Appel and Wood (2016) examined data from a frequently used standardized English proficiency test to highlight how lower-level writers favor LBs indicating personal stance.

Although previous studies generally grouped L2 English learners from varying L1 backgrounds together in hopes of identifying more widely applicable findings, the identification of production patterns that distinguish L2 users on the basis of their L1 has been growing in popularity. This area of LB research follows a more general trend in corpus informed studies of ESL/EFL writing that aims to identify L1 specific and universal production tendencies in L2 output (e.g., Jarvis & Pavlenko, 2008), often focusing on L1 Chinese learners of L2 English (e.g., Bychkovska & Lee, 2017; Chen &

Baker, 2010). For example, Bychkovska and Lee (2017) compared post-secondary English texts produced by L1 Chinese and L1 English writers to reveal that L1 Chinese students made use of more LB types and tokens than L1 English writers. These findings were attributed to a higher number of conversational LBs in L1 Chinese writing and heavy dependence on direct translation equivalents.

Although a focus on L1 Chinese EFL writing continues, additional L2 English users have also been investigated. Allen (2011), for instance, used a corpus of EFL writing to show that L1 Japanese writers of English tended to overuse LBs that had translation equivalents in their native language. For example, with reference to stance bundles, the author notes the high frequency of *it can be said*. Comparing L1 Korean EFL writers and native English users, Shin (2019) discovered that L1 Korean students displayed a greater tendency for stance and discourse organizing LBs.

Limitations in Previous Research

The above-mentioned studies add valuable knowledge regarding how LBs are used by various populations of L2 English writers, yet several inherent limitations persist. Notably, most of this research has used one-to-one contrasts involving a single group of L2 English writers in comparison to a reference corpus of L1 English (e.g., Allen, 2011; Chen & Baker, 2010; Shin, 2019). As this approach does not include additional L1 groups for comparison purposes, conclusions regarding whether identified production patterns are L1-related, or common to all L2 English writers, are impossible. Furthermore, L1/L2 comparisons often involve target language proficiency differences that may result in the misattribution of findings. Thus, moving away from L1/L2 comparisons in favor of contrasts targeting the interlanguage of multiple L1 groups may be more valuable (Ortega, 2011).

Several studies have begun to involve multiple populations of L2 English writers in their research (e.g., Appel & Murray, 2020; Karabacak & Qin, 2013; Paquot, 2017). Unfortunately, these studies have often failed to adequately control for proficiency and writing conditions and/or used extremely small sample sizes. For instance, in Karabacak and Qin (2013), only 17 samples from each of the L1 groups (Turkish, Chinese, and English) were used. Thus, it is difficult to make generalizable statements that could apply more broadly to each population of writers. Although Paquot (2017) analyzed a much larger collection of writing and used an innovative approach to highlight potential L1 related production tendencies, reliance on the International Corpus of Learner English (ICLE) could be

seen as a limitation, as the writing comes from post-secondary institutions with varying academic standards, writing conditions, and target language proficiencies. An examination of a small collection of writing from the ICLE found a range of B2 to C2 on the Common European Framework of Reference for languages (Granger & Thewissen, 2005) suggesting results from studies using the ICLE should be taken with caution. Similarly, the three corpora of ESL writing analyzed by Appel and Murray (2020) were comprised only of 'passing grade' papers, but the authors acknowledged that this 'pass/fail' distinction may have been overly broad in terms of controlling for proficiency, thus negatively impacting findings.

The Current Study

With limitations of previous research in mind, the present study aimed to use a more closely comparable collection of L2 English samples to perform a contrastive interlanguage analysis of L1 Chinese, Japanese, and Korean EFL writing. These L1s were chosen as they represent three of the most common groups of L2 English users from East Asian countries studying in English medium universities. Thus, a better understanding of these students' writing could offer benefits in terms of more targeted instruction that better addresses the needs of each group. The main research question providing focus to this study is provided below:

RQ. How do L1 Chinese, Japanese, and Korean EFL students make use of LBs in their academic English writing?

As an exploratory study, we focused on identification of potential L1 related production tendencies through the analysis of LBs. Although the discussion proposes factors that may explain our findings, it is beyond the scope of this paper to more definitively identify specific root causes. It is hoped that future research will build on the present study by incorporating analyses of L1 corpora, translation equivalents, pedagogic materials, pedagogic approaches, and L1 congruence to better understand the role each of these factors may play in the highlighted results.

Method

Corpora

Data were assembled from version 2.3 of the Written Essay Module of the ICNALE. The ICNALE is composed of essays and speeches by post-secondary students from 10 countries using standardized data collection procedures

that include common topics, writing conditions, access to materials, and allotted time. This corpus was specifically designed to facilitate studies focused on contrastive interlanguage analyses. In total, 5,600 essays from 2,800 writers are included in the ICNALE; however, there is substantial variance in terms of number of samples and assessed proficiency among each of the L1 sub-corpora. For example, although the Japanese and Chinese sections of ICNALE both contain 400 samples, only 50 of these have been assessed to be A2 in the L1 Chinese section and 154 of these are assessed at this level in the Japanese section.

For the current study, 140 essays from each of the three previously mentioned L1 groups (Chinese, Japanese, Korean) were gathered from the B1 level[1], as this proficiency band contained a relatively large pool of data from which to draw. All essays were written to address the same writing prompt: *It is important for college students to have a part-time job* (agree or disagree) as a way of controlling for any potential topic influence. As can be seen in Table 1, Chinese writers produced the longest average essays and showed the greatest standard deviation, yet all three groups were comparable in terms of total corpus size and mean essay length.

Table 1

Descriptive Statistics of the Corpora

	Chinese (n = 140)	Japanese (n = 140)	Korean (n = 140)
Total running words	34,575	31,892	31,988
Mean (range)	245 (195-338)	226 (176-302)	227 (189-326)
Standard deviation	34	23	27

Extraction Criteria

Range & Frequency

Range and frequency are the main identifying criteria for LBs, yet values vary from study to study. Minimum range is a means of avoiding idiosyncratic language use from a minority of texts/users which could skew results by misrepresenting general tendencies. Previous studies have often set range as either a raw number (e.g., Chen & Baker, 2010; Shin, 2019) or percentage (e.g., Appel & Murray, 2020; Hyland, 2008) of the total number of texts. As raw numbers can be influenced by the total number of samples in each

corpus (i.e., achieving a 5-text minimum in a corpus of 50 essays may be more difficult than achieving this same number in a corpus of 500 essays), a percentage threshold for range was used in the current study. As the main function of range is the elimination of idiosyncratic tendencies, 10% (14 texts from each corpus) was used to achieve this goal.

In instances where large corpora or corpora of substantially different sizes are analyzed, a normalized frequency (typically a value per million words) is used. Conversely, in studies examining smaller corpora, or those with more comparable word counts, a raw frequency is more commonly applied. Because the three corpora in the current study are of a comparable size, and all essays were relatively short (approximately 230 words, on average), the previously established minimum range figure of 14 was also applied as the frequency criterion. Thus, any bundle appearing in at least 10% of texts from any L1 group (14 occurrences) would fulfil both the range and frequency criteria.

Length

For sequence length, 4-word bundles are common as this often produces a manageable set of items for analysis (e.g., Chen & Baker, 2010), shorter sequences are contained within their boundaries (Cortes, 2004) and they offer relatively clear functional roles (Hyland, 2008). However, an exclusive focus on 4-word structures has been criticized in recent years as leading to potential misidentification because longer and shorter sequences are hidden from analysis (e.g., Adel & Erman, 2012; Appel & Trofimovich, 2017).

In the current study, we began by extracting all 3-word sequences, with instances of contracted forms treated as two words (e.g., *don't*, *won't*). However, substantial overlap suggested the presence of longer repeated structures. Therefore, target length was expanded to include any sequence meeting the aforementioned frequency and range, regardless of length. In doing so, it was possible to identify highly frequent LBs of up to 14 words in length. To eliminate the presence of partially overlapping sequences, all extracted items were reviewed, and shorter structures embedded in longer sequences were eliminated before beginning the analyses. For example, *a part time* was identified in the L1 Chinese corpus as a 3-word LB. The bundle *a part time job* was also identified, superseding the 3-word bundle, which in turn was superseded by the 5-word bundle *a part time job in*. As this 5-word bundle met the previously established extraction criteria, this 5-word sequence was retained but the shorter LBs it contained (*a part time*, *a part time job*) were eliminated from subsequent analyses.

Prompt-Related Bundles

Contrary to many earlier studies that used writing from a range of topics and/or genres, the current study examined argumentative essays addressing one common prompt. Thus, although previous scholars have often removed prompt/topic related LBs from their analyses, this was unnecessary in the present research. However, although bundles containing the topic-related words 'part-time', 'college' and 'students' were included throughout the analysis, they are not included in the analysis of individual items in Tables 4, 5, and 6 as a way of providing focus and allowing greater emphasis on the discussion of more interpretable findings.

Analysis

Extraction and analysis of LBs followed three steps. First, LBs meeting the identification criterion were extracted from each corpus. These were then reviewed to eliminate overlapping sequences (i.e., shorter sequences that were constituent in longer LBs). Second, cleaned lists of 3-14-word sequences were functionally classified using Hyland's (2008) categorization This taxonomy was used as it builds on Biber et al.'s (2004) classifications, has been argued to be better suited to academic texts, and follows our previous research (Appel & Murray, 2020) which can lead to greater comparability. This classification system includes three major functional categories. Research-oriented bundles aid the explanation of real-world occurrences, often through direct reference to concrete objects and abstract concepts (e.g., ***The part-time job that*** *most students have*); text-oriented bundles aid in the organization of discourse by helping to guide the reader, often with signposting language (e.g., ***First of all****, …*); and, participant-oriented bundles are writer/reader-focused and often serve to provide statements that make clear the writer's personal opinions (e.g., ***I think that*** *this should be stopped*). Each major functional category also contains distinct subcategories (see Hyland, 2008). Following functional classifications, loglikelihood statistics were used to highlight significant intergroup production differences related to functional (sub)categories and individual LBs. In all cases, findings are only reported as significant if these contrasts yielded $p < .01$.

We begin by providing a brief overview of general findings regarding frequency of various sequence lengths in the three corpora. This is followed by analysis of the most commonly used LBs by each L1 group. Finally, functional category and individual item comparisons are used to highlight significant intergroup production differences.

Results

General Findings

L1 Japanese were the most frequent users of LBs as a whole, with a total of 51 types and 1,138 tokens (Table 2). L1 Chinese writers closely followed L1 Japanese writers in terms of total LB occurrences and L1 Korean writers were the least frequent users. Most striking is the high number of longer LBs (greater than 5-words) in the L1 Japanese corpus, which suggests significant overlap and a high level of intra-group similarity. In contrast, L1 Korean writers seem to possess the greatest intra-group variance as both type and token counts across nearly all sequence lengths were comparatively low.

Table 2
3-word to 14-word Bundles by L1 Group

Bundle Length	Chinese	Japanese	Korean
3-word	25 (566)	29 (665)	20 (465)
4-word	11 (236)	11 (238)	11 (274)
5-word	8 (215)	1 (33)	4 (111)
6-word	3 (66)	5 (92)	1 (21)
7-word	-	2 (32)	-
8-word	-	1 (15)	-
12-word	1 (24)	-	1 (23)
13-word	-	1 (44)	-
14-word	-	1 (19)	-
Total	48 (1,107)	51 (1,138)	37 (894)

Note. Type counts are listed outside of parentheses with token counts listed within parentheses. All 9-, 10- and 11-word bundles were constituent in longer bundles.

Functional Analysis

To ensure consistency in functional assignment (Table 3), classifications were performed independently by the two authors of this study before reconvening to discuss discrepancies. Interrater reliability for these initial classifications was 91% (agreement on 124 of 136 total LBs). Full agreement was achieved through joint discussion.

In the following analyses, the terms *overuse* and *underuse* are used to refer to instances of statistically distinct production by one L1 group in comparison to the other two L1s. These terms are used in a relative manner and should not be seen as an indication of improper use. Where pedagogic implications are given, this is explicitly stated.

Table 3
Lexical Bundles by Functional Category

	Chinese	Japanese	Korean
Research-oriented	850 (77%)	697 (61%)	778 (87%)
Location	55 (5%)	27 (2%)	33 (4%)
Procedure	276 (25%)	132 (12%)**	222 (25%)
Quantification	138 (12%)**	66 (6%)	48 (5%)
Description	266 (24%)	245 (22%)	189 (21%)**
Topic	115 (10%)**	227 (20%)	286 (32%)
Text-oriented	102 (9%)**	52 (5%)**	16 (2%)**
Transition	102 (9%)**	52 (5%)**	16 (2%)**
Participant-oriented	155 (14%)**	389 (34%)**	100 (11%)**
Stance	155 (14%)**	389 (34%)**	100 (11%)**
Total	1,107 (100%)	1,138 (100%)	894 (100%)**

Note. $^{**}p < .01$; Only sub-categories with LB occurrences are listed. Type counts are listed outside of parentheses with percent of token counts listed within parentheses.

Two major functional categories (*text-oriented*, *participant-oriented*) displayed significant L1 related production differences (Table 3). Furthermore, loglikelihood statistics indicated that each L1 held a unique tendency related to the *text-oriented* category, with L1 Chinese writers the most frequent users. The *participant-oriented* category also displayed similar results, with each L1 making use of these items in a statistically unique manner. However, in this case, L1 Japanese writers were the most frequent users.

Within the *research-oriented* category, the *procedure*, *quantification*, *description*, and *topic* subcategories showed significant intergroup production differences. For the *procedure* subcategory, L1 Japanese writers were the least frequent users; L1 Chinese overused *quantification* yet underused

the *topic* subcategory; *description* was underused by L1 Korean writers. In terms of total use across all categories, L1 Korean writers were found to underuse LBs (tokens).

Individual Lexical Bundle Analysis

L1 Chinese

L1 Chinese EFL writing contained the highest number of unique L1-related production tendencies (19) for individual LBs. These items were roughly split between overused (10) and underused (9) items. In relation to significant functional category differences identified in Table 3, only one item (*all kinds of*) could be identified as contributing to the relative overuse of the quantification subcategory (i.e., LBs used to describe amounts). However, 4 LBs from the text-oriented category (*all in all*, *at the same time*, *last but not*, *what's more*) helped to explain the relative overuse of this particular functional category.

Table 4
Individual Item Overuse/Underuse by L1 Chinese Writers

Category	Bundle	Chinese	Japanese	Korean
Research-oriented: Location	*in the society*	17	4	2
Research-oriented: Purpose	*be able to*	1	21	16
Research-oriented: Purpose	*for us to*	34	12	0
Research-oriented: Purpose	*I want to*	4	28	35
Research-oriented: Purpose	*they (have/ want) to*	2/4	15/14	19/17
Research-oriented: Quantification	*a lot of (money)*	51 (2)	116 (27)	77 (19)
Research-oriented: Quantification	*all kinds of*	14	1	1
Research-oriented: Description	*importance of money*	0	17	8
Text-oriented: Transition	*all in all*	15	0	0
Text-oriented: Transition	*at the same time*	17	3	2
Text-oriented: Transition	*but it is*	3	13	16

Category	Bundle	Chinese	Japanese	Korean
Text-oriented: Transition	*last but not*	16	0	0
Text-oriented: Transition	*what's more*	24	0	0
Participant-oriented: Stance	*as far as I*	15	0	1
Participant-oriented: Stance	*I think that*	8	96	30
Participant-oriented: Stance	*in my opinion*	34	4	8
Participant-oriented: Stance	*we all know*	15	0	0

L1 Japanese

Significant differences for individual LBs among L1 Japanese writers covered all three major functional categories, yet were primarily related to the *participant-stance* subcategory. With all items from this category indicating relative overuse, findings in Table 5 help to explain the previously identified functional overuse by L1 Japanese EFL writers. Overused items are primarily used to indicate writers' personal opinions, with several making explicit mention of the writer by way of 'I'. In terms of relative underuse of the *research-procedure* subcategory (used to detail processes) highlighted above, two LBs (*for us to*, *to get a*) help to explain this.

Table 5

Individual Item Overuse/Underuse by L1 Japanese Writers

Category	Bundle	Chinese	Japanese	Korean
Research-oriented: Purpose	*for us to*	34	12	0
Research-oriented: Purpose	*to get a*	12	2	22
Research-oriented: Quantification	*a lot of (things)*	51 (1)	116 (20)	77 (3)
Research-oriented: Description	*we can't*	10	25	2
Research- oriented: Topic	*the statement that*	1	17	0
Research- oriented: Topic	*with this statement*	1	14	0
Text-oriented: Transition	*and so on*	13	52	13

Category	Bundle	Chinese	Japanese	Korean
Participant-oriented: Stance	*I agree with the statement*	1	33	0
Participant-oriented: Stance	*I agree with this*	1	25	1
Participant-oriented: Stance	*I think it is*	19	53	15
Participant-oriented: Stance	*I think that (it is important for)*	8 (2)	96 (16)	30 (3)
Participant-oriented: Stance	*it is important to*	8	24	8
Participant-oriented: Stance	*so I think it*	5	17	2
Participant-oriented: Stance	*we have to*	8	24	3
Participant-oriented: Stance	*why I think*	0	16	2

L1 Korean

For L1 Korean writers, only three unique production tendencies were uncovered, all from the research-oriented functional category. Given the general underuse of LBs among L1 Korean writers, it is unsurprising that two of the three unique production tendencies in Table 6 were related to relative underuse. For the overused item (*is very expensive*), concordance lines revealed that this LB was exclusively used to bring focus to the high cost of tuition students face when attending post-secondary institutions.

Table 6
Individual Item Overuse/Underuse by L1 Korean Writers

Category	Bundle	Chinese	Japanese	Korean
Research-oriented: Purpose	*for us to*	34	12	0
Research-oriented: Quantification	*is very expensive*	1	1	14
Research-oriented: Quantification	*the most important*	22	14	2

Discussion

In contrast to most previous research, the current study avoided L1/L2 comparisons in favor of a contrastive interlanguage approach involving multiple L1 groups to better identify potential L1-related tendencies. This decision was largely driven by the growing recognition that L1/L2 comparisons suffer from a comparative fallacy that implies L2 users are somehow deficient in their language use (Hunston, 2002; Larsen-Freeman, 2014) and that using L1 discourse as a baseline hinders accurate descriptions of the L2 variety being analyzed (Bley-Vroman, 1983). Results of these inter-language (i.e., L2) comparisons revealed numerous differences in LB use by each L1 that suggest varying approaches to academic English writing.

L1 Chinese

L1 Chinese writers fell between Japanese and Korean in terms of total LB tokens. This contrasts with Appel and Murray (2020), which identified substantially more LBs in the L1 Chinese ESL corpus than in either the L1 Arabic or French corpora used for comparison. One potential reason for this discrepancy is the choice of L1 groups in each study. For example, Appel and Murray suggested that the relatively high number of LBs in L1 Chinese ESL writing may have resulted from emphasis given to collectivist thinking in Chinese culture. Given the more similar cultural basis of the three groups of L1 writers in the current study, this distinction may have become less apparent (see Hofstede et al., 2010).

Although the inclusion of more culturally comparable writer groups may have reduced the uniqueness of L1 Chinese production in this regard, individual LB patterns did indicate a uniquely collectivist approach among these writers. For example, aside from two overused bundles containing the first-person singular (*as far as I, in my opinion*), overuse of first-person plural (*for us to*, *we all know*) and underuse of first-person singular (*I want to, I think that*) were found. These findings, particularly overuse of first-person plural, may suggest a preference for LBs expressing 'in group' membership perhaps indicating how these writers view their position within a collectivist society. This argumentative approach positions writer and reader as part of the same collective, with those who disagree inherently viewed as outsiders, evidenced in the examples listed below where the writers seem to pursue inherent agreement with their position:

> *<W_CHN_PTJO_003>* ***We all know*** *that food made by ourselves tastes more delicious.*

> *<W_CHN_PTJO_005> As **we all know**, students have a lot of free time to manage.*
>
> *<W_CHN_PTJO_014> As **we all know**, college tuition is not a small sum.*

Further support for this position is found in Liardet (2018), where *we all know* was also identified as commonly used by Chinese EFL writers. Liardet situates this pattern as a subjective contracting evaluation that separates 'in group' agreement from 'out group' dissent. That L1 Chinese EFL writers are the lone group in this study from a communist society, wherein greater encouragement may be given to group membership, is likely an influential factor in the writing produced and LBs identified. However, as additional factors may also be at play, not all tendencies should be seen as a result of cultural influences, and factors such as pedagogic materials and instructional approach need to be analyzed in future research on this topic (see below).

The high number of text-transition bundles in the L1 Chinese corpus is in line with Leedham and Cai (2013), who examined production patterns for linking adverbials in L1 Chinese EFL writing. Of the six 3-4-word items they found to be overused when compared to L1 English writers, three were also overused relative to other L2 English users in the present study (*at the same time*, *last but not [least]*, and *what's more*), suggesting they may be regularly repeated patterns among L1 Chinese EFL writers.

Leedham and Cai (2013) ascribe this to teaching materials and an emphasis on rote learning within the Chinese education system. They also note that pedagogic approaches in mainland China typically fail to distinguish register, resulting in frequent use of less academic phrases, such as *what's more*, a bundle repeatedly identified as a distinctive feature of L1 Chinese L2 English writing (e.g., Appel & Szeib, 2018; Lee & Chen, 2009; Leedham & Cai, 2013). Given that the current study also found overuse for this item, pedagogic interventions may be necessary to reduce usage, with greater attention given to the importance of genre and register differences.

L1 Japanese

L1 Japanese writers had the highest overall number of LBs, with the largely similar nature of their writing suggesting collectivism, yet also frequent use of singular first-person pronouns, signifying individualism. Nam (2016) also identified relative overuse of first-person pronouns among L1 Japanese participants when comparing L1 Korean and Japanese EFL writers. This dichotomy between group and self may be accounted for somewhat by

Hofstede et al. (2010), who position Japan at a midway between collectivism and individualism. Again, however, more research, including analyses of pedagogic materials, will be needed to better understand these findings. For example, Northbrook and Conklin (2018) uncovered a high frequency of LBs featuring pronouns in textbooks used in Japanese junior high schools, arguing that frequent exposure could lead to subsequent language use. Thus, the high number of overused bundles incorporating first-person pronouns identified in the current study, as well as Nam, could be a lasting influence of pedagogic materials.

L1 Japanese were also the most frequent users of longer bundles, many of which incorporated portions of the essay prompt, with the full prompt occurring 72 times (compared with 24 and 23 occurrences in the L1 Chinese and L1 Korean corpora, respectively). This contrasts with the findings of Appel and Murray (2020), whose analysis of three L1 groups (Arabic, Chinese, and French) revealed L1 Chinese as the most frequent users of longer LBs, including those drawn from the essay prompts. Again, this difference may be a result of including more socially similar L1 groups in the present study. Granted partial use of the essay prompt is not in itself bad practice–Wray and Pegg (2009) note that it is also common practice among L1 writers–verbatim copying of the entire prompt appears to be a feature unique to the L1 Japanese context (at least when compared to L1 Chinese and Korean writers of English in this study).

Two longer prompt related phrases from the L1 Japanese group's top 10 bundles further highlight this feature: *I agree with the statement* was identified as overused by L1 Japanese writers (33 occurrences), with only one instance in the L1 Chinese corpus and entirely absent from the L1 Korean corpus, and *I agree with this* (25 times in the L1 Japanese corpus) occurred only one time in the L1 Chinese and Korean corpora. Although it is difficult to pinpoint the root cause of these tendencies, together with the highly frequent use of the various portions of the essay prompt, previous language teaching pedagogy may again be at play. The findings suggest that L1 Japanese students are being told to clearly signal their opinion with reference to the essay question, and are seemingly being provided phrases for doing so. However, further research involving L1 Japanese learners and the language teaching they commonly receive would be needed to confirm this.

L1 Korean

L1 Korean EFL writers were the least frequent users of LBs overall, apparent primarily through lower use of text- and participant-oriented bundles.

Comparatively low frequencies were also apparent in individual LBs, indicating less overlap, and greater language diversity.

With formulaic language a prevalent feature of academic writing (Hyland, 2008), lower frequency may signal reduced adherence to genre/register norms. However, because all writers were assessed to be at a similar proficiency level, further studies are needed to more closely examine this issue. Other potential explanations include a greater willingness among L1 Korean EFL writers to express themselves in non-standard ways and potentially higher lexical diversity.

L1 Korean EFL writers were especially infrequent users of participant-stance bundles. Jaworska et al. (2015) examined the use of stance expressions (labelled in the present study as 'participant-stance oriented') by L1 English and L1 German groups in argumentative essays in German, ascribing the L1 English group's greater use of impersonal and cautious language when expressing stance not to L1/L2 status, but to the transference of L1 rhetorical conventions. From this perspective, certain cultures place greater responsibility on writers to make text organization explicit, whereas others (including Chinese, Japanese, and Korean) place onus for understanding with the reader (see Leńko-Szymańska, 2008).

If, as claimed, all three groups in this study follow the same convention regarding text organization, the differences identified here in terms of the L1 Korean EFL writers' relatively low use of both transition and participant stance bundles may again be related to pedagogy. Leńko-Szymańska (2008) examined linking expressions among various L1 groups, finding significantly different levels of use between L1 groups classified as belonging to the same writing tradition, yet none between groups from different writing traditions, ascribing these differences in part to home country pedagogy. Here again, however, these conclusions are tentative and further research, which includes corpora of pedagogic materials used in each home country, is needed to more closely analyze each potential source of production differences.

Implications

Although not all identified production differences suggest a need for pedagogic intervention, L1-specific targeted instruction may be beneficial in improving the appropriateness of each group's academic English. Furthermore, the commonalities that were discovered suggest a combination of targeted (i.e., L1-specific) and general (i.e., universal to all L2 English users) instruction may prove beneficial.

In the case of L1 Chinese EFL writers, overuse of LBs less appropriate to the academic written register (e.g., *last but not [least]*, *what's more*) may require pedagogic interventions. Furthermore, as noted above, this group's use of stance bundles tended towards signaling a more subjective contracting evaluation (i.e., one suggesting in-group agreement and rejecting outgroup dissent), which may be less appropriate in academic writing. Thus, focused instruction in register-appropriate transition phrases and stance expressions should be incorporated into the pedagogic approach for these students. Again, however, our findings were based on a relatively limited data set and future research assessing how well these findings apply to the target population at large is needed.

L1 Japanese EFL learners tended to overuse a narrow range of stance bundles and high number of bundles with first-person pronouns. Phrases including *I think* feature prominently and have been found with high frequency in previous studies of L1 Japanese corpora (e.g., Kobayashi, 2009; McCrostie, 2008), yet also among other L1 groups (e.g., Petch-Tyson, 1998; Ringbom, 1998). Though the present study found overuse in comparison to the other two L1 groups, more research may be needed to establish if it is indeed an L1-specific phenomenon or a more widespread issue among L2 learners. Kobayashi (2009) does however attribute the high frequency among L1 Japanese learners to L1 transfer. Contrastingly, Fordyce (2014) states that the 'more difficult' stance phrases typically make use of modal verbs, a structure he argues is problematic for L1 Japanese learners as it does not exist in the L1. Thus, there may be a need for focused instruction and guided practice making use of these sequences. L1 Japanese EFL writers' heavy reliance on personal pronouns may require similar treatment; if this is indeed a feature L1 Japanese learners have acquired through home country pedagogic materials, explicit instruction to reverse this tendency will be needed and an effort to revise these materials may be necessary.

L1 Korean EFL writing indicated general underuse of LB types and tokens, suggesting greater intra-group variance and potential deviance from standard academic written norms. This underuse was especially common with bundles used to signal transition and stance, both of which one might expect to be clearly marked within an argumentative essay. Combined, these factors indicate a potential lack of awareness regarding meeting register expectations through appropriate LB use. To highlight the importance of transitions, students could complete activities comparing texts lacking sufficient transition signals with those making effective use. To avoid the issue of register inappropriate use of transitions (e.g., *what's more*) identi-

fied in the L1 Chinese group, attention should be paid to including academic expressions. Fill-in-the blank activities, where students select register and context appropriate transitions, could also be used.

A second area with potential pedagogic implications relates to the influence of teaching materials within the L1 context. For L1 Chinese, this was seen in the overuse of text-transition bundles, similar to those identified by Leedham and Cai (2013), who suggest that sample texts and vocabulary lists typically provided to students in China may be partly to blame. Similarly, overuse of LBs featuring pronouns identified among L1 Japanese writers may result from teaching materials (Northbrook & Conklin, 2018). Thus, the influence of pedagogic materials may be a factor that merits closer examination in future studies. If such pedagogic factors are found to cause L1 specific over/underuse, remedying the materials at fault would be more beneficial than attempting to subsequently counter the symptoms created.

Finally, in terms of methodological implications, an aspect addressed by Appel and Murray (2020) and further developed here (i.e., the value of not limiting extracted sequence length to the common 4-word length, but instead including all bundles which fulfill the identification criteria applied), deserves mention due to the impact this can have on results. Appel and Murray went beyond the typically applied 4-word bundle limit to include all items from three to seven words in length, though the current study opted to remove the upper word limit altogether. Although restricting sequence length may help maintain a focus on the most frequent patterns, if, as here, the purpose is to examine and compare actual language use across corpora, employing identification criteria which allow the full range of these differences to surface seems critical. Doing so allows for a more complete picture by looking beyond the 4-word structures commonly sought which could often be more accurately viewed as single, extended items. For example, in the present study, this approach was used to help identify frequent and verbatim use of the essay prompt by L1 Japanese EFL writers, a factor which may otherwise have gone unnoticed.

Limitations and Future Research

Limitations in the current study arise from two main areas which future researchers should attempt to address. First, although efforts were made to control for proficiency across L1 groups, the essays included in the ICNALE use automated measures to classify writers into each proficiency band. Thus, more strictly controlled proficiency measures may be needed to more adequately control for the influence of proficiency differences on findings.

The second main limitation concerns the lack of understanding regarding root causes of identified production differences. As an exploratory study, we aimed to identify L1 differences to make tentative suggestions regarding why such tendencies were exhibited. Given the implications of such features as overuse of register-inappropriate language in terms of how a writer is assessed, further studies that look more closely at potential root causes for these tendencies are clearly needed in order to redress them. In-depth examinations of pedagogic materials used in each country, collections of L1 writing from each group, and potential translation equivalents, could all prove beneficial.

Conclusion

Findings from this study contribute to the growing body of research suggesting that particular L1 groups produce L2 language in significantly distinct ways. Production differences were found to occur at various levels of analysis, including overall LB type/token counts, functional category, and individual item, indicating areas where each L1 group may benefit from focused pedagogic interventions. Various factors were proposed as influencing these distinct production tendencies, including cultural elements, L1 transfer, and language teaching pedagogy. However, the influence of each factor on production remains unclear, meaning further research is necessary to bring greater clarity to the issues at hand.

Notes

1. The ICNALE uses scores from high-stakes proficiency (e.g., TOEFL, TOEIC) and vocabulary size tests to map L2 writers' proficiency ratings onto CEFR levels.

Randy Appel is an associate professor in the Global Education Center at Waseda University. His research interests include corpus approaches to linguistic inquiry, L2 English academic writing, and the use of formulaic language by L1 and L2 users.

Lewis Murray is an associate professor in the Foreign Language Education department at Kanazawa University. His research interests relate primarily to L2 learners' acquisition of formulaic language in relation to academic writing.

References

Adel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes, 31*(2), 81–92. https://doi.org/10.1016/j.esp.2011.08.004

Allen, D. (2011). Lexical bundles in learner writing: An analysis of formulaic language in the ALESS learner corpus. *Komaba Journal of English Education, 1*, 105–127. https://park.itc.u-tokyo.ac.jp/eigo/KJEE/001/105-127.pdf

Appel, R., & Murray, L. (2020). L1 differences in L2 English academic writing: A lexical bundles analysis. *Journal of English for Academic Purposes, 46*, 1–15. https://doi.org/10.1016/j.jeap.2020.100873

Appel, R., & Szeib, A. (2018). Linking adverbials in L2 English academic writing: L1 related differences. *System, 78*, 115–129. https://doi.org/10.1016/j.system.2018.08.008

Appel, R., & Trofimovich, P. (2017). Transitional probability predicts native and non-native use of formulaic sequences. *International Journal of Applied Linguistics, 27*(1), 24–43. https://doi.org/10.1111/ijal.12100

Appel, R., & Wood, D. (2016). Recurrent word combinations in EAP test-taker writing: Differences between high- and low-proficiency levels. *Language Assessment Quarterly, 13*(1), 55–71. https://doi.org/10.1080/15434303.2015.1126718

Biber, B., Conrad, S., & Cortes, V. (2004). If you look at …: Lexical bundles in university teaching and textbooks. *Applied Linguistics, 25*(3), 371–405. https://doi.org/10.1093/applin/25.3.37

Biber, B., Conrad, S., & Reppen, R. (1999). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press. https://doi.org/10.1017/CBO9780511804489

Bley-Vroman, R. (1983). The comparative fallacy in interlanguage studies: The case of systematicity, *Language Learning*, *33*(1), 1–17. https://doi.org/10.1111/j.1467-1770.1983.tb00983

Bychkovska, T., & Lee, J. (2017). At the same time: Lexical bundles in L1 and L2 university student argumentative writing. *Journal of English for Academic Purposes, 30,* 38–52. https://doi.org/10.1016/j.jeap.2017.10.008

Chen, A. C.-H. (2019). Assessing phraseological development in word sequences of variable lengths in second language texts using directional association measures. *Language Learning, 69*(2), 1–38. https://doi.org/10.1111/lang.12340

Chen, Y., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning and Technology, 14*(2), 30–49. https://doi.org/10125/44213

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes, 23*(4), 397–423. https://doi.org/10.1016/j.esp.2003.12.001

Durrant, P. (2017). Lexical bundles and disciplinary variation in university students' writing: Mapping the territories. *Applied Linguistics, 38*(2), 165–193. https://doi.org/10.1093/applin/amv011

Fordyce, K. (2014). The differential effects of explicit and implicit instruction on EFL learners' use of epistemic stance. *Applied Linguistics, 35*(1), 6–28. https://doi.org/10.1093/applin/ams076

Granger, S., & Thewissen, J. (2005). *Towards a reconciliation of "can do" and "can't do" approach to language assessment.* [Paper presentation]. The Second Annual Conference of the European Association of Language Testing and Assessment, Voss, Norway.

Hofstede, G., Hofstede, G., & Minkov, M. (2010). *Cultures and organizations: Software of the mind* (Rev. 3rd ed.). McGraw-Hill.

Hunston, S. (2002). Pattern grammar, language teaching, and linguistic variation: Applications of a corpus-driven grammar. In R. Reppen, S. M. Fitzmaurice, & D. Biber (Eds.), *Using corpora to explore linguistic variation* (pp. 167–183). John Benjamins. https://doi.org/10.1075/scl.9.11hun

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes, 27*, 4–21. https://doi.org/10.1016/j.esp.2007.06.001

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. Routledge. https://doi.org/10.4324/9780203935927

Jaworska, S., Krummes, C., & Ensslin, A. (2015). Formulaic sequences in native and non-native argumentative writing in German. *International Journal of Corpus Linguistics, 20*(4), 500–525. https://doi.org/10.1075/ijcl.20.4.04jaw

Karabacak, E., & Qin, J. (2013). Comparison of lexical bundles used by Turkish, Chinese, and American university students. *Procedia - Social and Behavioral Sciences, 70,* 622–628. https://doi.org/10.1016/j.sbspro.2013.01.101

Kobayashi, Y. (2009). Extracting characteristics of English essays written by Japanese EFL learners based on the NS/NNS text categorization model. *Proceedings of Jinmonkon Symposium Series of Information Processing Society of Japan, 16,* 261–268.

Kuiper, K. (1996). *Smooth talkers: The linguistic performance of auctioneers and sportscasters*. Lawrence Erlbaum.

Larsen-Freeman, D. (2014). Another step to be taken: Rethinking the endpoint of the interlanguage continuum. In Z. H. Han & E. Tarone (Eds.), *Interlanguage: Forty years later* (pp. 203–220). John Benjamins. https://doi.org/10.1075/lllt.39.11ch9

Lee, D., & Chen, S. (2009). Making a bigger deal of the smaller words: Function words and other key items in research writing by Chinese learners. *Journal of Second Language Writing, 18*(4), 149–165. https://doi.org/10.1016/j.jslw.2009.07.003

Leedham, M., & Cai, G. (2013). Besides…on the other hand: Using a corpus approach to explore the influence of teaching materials on Chinese students' use of linking adverbials. *Journal of Second Language Writing, 22*(4), 374–389. https://doi.org/10.1016/j.jslw.2013.07.002

Leńko-Szymańska, A. (2008). Non-native or non-expert? The use of connectors in native and foreign language learners' texts. *Aile: Acquisition et Interaction en Langue Étrangère, 27*, 91–108. https://doi.org/10.4000/aile.4213

Liardet, C. (2018). 'As we all know': Examining Chinese EFL learners' use of interpersonal grammatical metaphor in academic writing. *English for Specific Purposes, 50*, 64–80. https://doi.org/10.1016/j.esp.2017.11.005

Lu, X., & Deng, J. (2019). With the rapid development: A contrastive analysis of lexical bundles in dissertation abstracts by Chinese and L1 English doctoral students. *Journal of English for Academic Purposes, 39*, 21–36. https://doi.org/10.1016/j.jeap.2019.03.008

McCrostie, J. (2008). Writer visibility in EFL learner academic writing: A corpus-based study. *ICAME Journal, 32,* 97–114. http://icame.uib.no/ij32/ij32_97_114.pdf

Nam, D. (2016). Lexical bundle structures in argumentative writing of Korean and Japanese learners of English. *Studies in Foreign Language Education, 30*(3), 1–24. https://doi.org/10.16933/sfle.2016.30.3.1

Northbrook, J., & Conklin, K. (2018). "What are you talking about?" An analysis of lexical bundles in Japanese junior high school textbooks. *International Journal of Corpus Linguistics, 23*(3), 311–334. https://doi.org/10.1075/ijcl.16024.nor

Ortega, L. (2011, June). *New trends in SLA research: Theories, methods, ethics*. Hsinchu, Taiwan: Invited lecture at the National Tsing Hua University.

Paquot, M. (2017). L1 frequency in foreign language acquisition. Recurrent word combinations in French and Spanish EFL learner writing. *Second Language Research, 33*(1), 12–32. https://doi.org/10.1177/0267658315620265

Petch-Tyson, S. (1998). Writer/reader visibility in EFL written discourse. In S. Granger (Ed.), *Learner English on computer* (pp. 107–118). Routledge. https://doi.org/10.4324/9781315841342-8

Ringbom, H. (1998). Vocabulary frequencies in advanced learner English: A cross-linguistic approach. In S. Granger (Ed.), *Learner English on computer* (pp. 41–52). Routledge. https://doi.org/10.4324/9781315841342-3

Schmitt, N. (Ed.). (2004). *Formulaic sequences: Acquisition, processing and use.* John Benjamins. https://doi.org/10.1075/lllt.9

Shin, Y. (2019). Do native writers always have a head start over nonnative writers? The use of lexical bundles in college students' essays. *Journal of English for Academic Purposes, 40*, 1–14. https://doi.org/10.1016/j.jeap.2019.04.004

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press. https://doi.org/10.1017/CBO9780511519772

Wray, A., & Pegg, C. (2009). The effect of memorized learning on the writing scores of Chinese IELTS test takers. *IELTS Report, 9*, 191–216. https://www.ielts.org/-/media/research-reports/ielts_rr_volume09_report5.ashx

Self-Regulated Learning and First-Year College Success: A Longitudinal Case Study in Japan

Yukiko Ishikawa
Soka Women's College

The 1st year of college poses the challenge of adjusting to a new environment. Self-regulated learning (SRL) is crucial for a successful transition to college. This study investigated 1st-year students' SRL development and SRL skills' impact on time management for self-study. 8 participants at a 2-year women's college in Japan were interviewed 4 times in their 1st year and once at the end of their 2nd year. The SRL developmental trajectories of the participants in both higher and lower proficiency groups revealed that some of them followed but not all of them completed the cyclic phases of SRL and the levels of the SRL developmental process as Zimmerman's (2000, 2004) social cognitive model suggests. SRL skills also impacted students' time spent on language learning. The factor that impeded SRL development was difficulty with emerging needs; this included setting realistic goals, balancing study and social obligations, and controlling emotional distractions.

大学の初年次は新しい環境に適応する上で新入生にとっては大きなチャレンジである。大学へのスムーズな移行には、自己調整学習が重要となる。この研究では、新入生の自己調整学習スキルの発達過程と、そのスキルが自己学習へ及ぼす影響を調査した。日本の女子短期大学で学ぶ8名の学生に、1年次に4回、また2年次の終わりに1回のインタビューを行った。習熟レベル上位と下位のグループがたどった自己調整学習の発達過程を分析した結果、ジマーマン(2000, 2004)の社会的認知モデルが提唱した通り3段階の学習フェーズと4段階の発達過程をたどった学生と、そうではない学生がいたことがわかった。また、自己調整学習スキルは英語の自己学習時間にも影響することが明らかとなった。現実的な目標設定の難しさ、勉強とそれ以外の社会的な活動との両立、心理的なストレスをコントロールする難しさなどが自己調整学習を阻む初年次特有の要因として示唆された。

https://doi.org/10.37546/JALTJJ45.1-3

JALT Journal, Vol. 45, No. 1, May 2023

Keywords: low proficiency learners; transition to college

The transition to college may pose numerous challenges to new college students (Upcraft et al., 2005). A successful transition for first-year students includes not only developing intellectual abilities, but also factors such as developing support networks, exploring their identity, and choosing a career path (Upcraft et al., 2005). Researchers and practitioners have attempted to identify what predicts success in college, examining prior subject knowledge, entrance exam scores, aptitude test results, personal traits, and other factors (Harvey et al., 2006; Krsmanovic et al., 2020). Other researchers have focused on self-regulatory skills as critical skills to survive the first year (Thibodeaux et al., 2017; Wolters & Brady, 2021). Developing and using effective self-regulatory skills is essential to succeed in college. This paper focuses on the development and use of self-regulated learning (SRL) skills in the first year in college.

Literature Review

Transition to College

The transition to college in the U.S. and U.K. contexts has been widely discussed because student retention has been a major issue. There is an increasing need to support a diverse population of students, including minority and first-generation students, in the first year of college to avoid attrition (Conefrey, 2018; Harvey et al., 2006). Attrition is not a major problem for Japanese universities. However, Japanese colleges and universities are facing the situation of having to recruit from a smaller pool of candidates in an attempt to maintain their enrollment levels. In 2021, 46.4% of private Japanese universities reported that they had not filled all of their places for new enrollment (Kyodo, 2021). Although Japan's population of 18-year-olds has been decreasing, the number of universities in Japan has increased due to government deregulation of university establishment standards (Brasor, 2017; Harada, 2015). According to Harada (2015), "by around 2000, Japan had already entered an age of 'universal' access to higher education—meaning that everyone can go to college as long as they are not picky about the school or faculty" (para. 2). To maintain enrollment levels, universities are admitting more high school graduates of lower academic ability; teachers have noticed a decline in academic skills of university students (Matsutani, 2012). More academic support seems to be necessary for these less prepared incoming students. The situation is more severe for two-year colleges in Japan. More students are entering four-year universities. Since the 1990s, the number of

two-year colleges has declined to under 60% and the number of students to 25% (Doi, 2017). Although little is reported on their academic ability, Matsumoto et al. (2011) reported that two-year college students are more likely to be depressed possibly due to the factors unique to two-year colleges: they need to graduate in two years, thus there are more classes every day and accordingly more assignments, and they need to start job hunting in less than half a year after entrance to college. Many students choose to transfer to a four-year university and the first-year GPA is often one of the selection criteria. Students in a two-year college are likely to be more susceptible to stress and may need additional support in adjusting to college.

First year education is provided to support students with smooth transition to college. According to Yamada (2012), first year education began to draw attention around 2000 in Japan. Common topics covered in first year education courses include study skills, orientation to the university, introduction to majors, information literacy, and career planning (Yamada, 2019). Yamada (2012) pointed out that although first year education programs have been implemented in the United States since the late 1970s with methods underpinned by research, Japanese first year education programs still lack empirical grounding, claiming that practice has been taking precedence over research in Japan. A large number of research projects on first year education have been reported, but many are school based and on a small scale. On February 24, 2022, a search on the CiNii (Scholarly and Academic Information Navigator) database for "初年次教育 [first year education]" yielded 1952 results. They are mostly reports of school or classroom practices and cover a wide range of topics, including career planning, academic writing, online learning, and placement testing.

Studies suggest that several aspects determine a smooth transition to college. Some have pointed out that factors such as a sense of belonging and friendship may be involved in a successful transition (Bowman et al., 2019; Pittman & Richmond, 2008; Wilcox et al., 2005). Wolters and Brady (2021) claimed that time management is especially critical for first-year students because more autonomy is required in the use of time compared to that in secondary school. Thibodeaux et al. (2017) investigated first-year students' use of time and concluded that first-year students might not be good at self-regulatory skills and need more support with their time management skills in particular. Some researchers have examined the effectiveness of high-impact practices (e.g., writing-intensive courses, first-semester seminars) for first-year students. They have shown that their first-year courses including such practices are effective in improving first-generation students' self-

efficacy and self-regulated learning (Conefrey, 2018) and are effective also in an online learning environment (Stephen & Rockinson-Szapkiw, 2021). These studies added insights into helping students navigate the transition to college; however, little has been investigated in the context of Japan, with even less in a foreign language-learning context in Japan. The impact of the self-regulatory skills of first-year students on their adjustment to college in Japan needs more investigation.

Self-Regulated Learning (SRL)

In the research field of educational psychology, Zimmerman and Martinez-Pons (1986) observed high school students and examined high-achieving students' utilization of a greater number of learning strategies compared to other students. The strategies included goal-setting and planning, organizing and transforming, environmental structuring, seeking social assistance, and performing self-evaluations. The researchers called these SRL strategies. Zimmerman (2000) developed a social cognitive model of SRL and defined it as "self-generated thoughts, feelings, and actions that are planned and cyclically adapted to the attainment of personal goals" (p. 14). According to Zimmerman (2000), self-regulation involves the interaction of personal, behavioral, and environmental processes, and it comprises three cyclical phases. The forethought phase concerns actions that happen before learning, such as task analysis, goal setting, and strategic planning. The performance phase involves actions that occur during learning, such as self-instruction, attention focusing, and self-monitoring. In the self-reflection phase after learning, self-judgment and self-reactions occur. The actions in these three phases take place cyclically. Zimmerman (2000, 2013) further explained that SRL develops in four levels. First, at the observational level, a learner carefully watches a model learn or perform. Next, at the emulation level, a learner imitates the model's general pattern or style. Third, at the self-control level, a learner can use a self-regulatory skill in structured settings without the presence of models. Finally, at the self-regulation level, a learner can perform skills and adapt to changing conditions in naturalistic settings. Learners develop self-regulatory skills best when they learn progressively from Levels 1 to 4.

SRL in the language-learning context has been studied recently. Wang and Bai (2017) developed the Questionnaire of English Self-Regulated Learning Strategies (QESRLS) and examined Chinese secondary school students in an EFL context. They reported that students' use of self-regulated learning strategies was weakly related to their English proficiency midterm exam

scores (r = .19. p < .01) and final exam scores (r =.25, p < .01). Kim et al. (2015) found that the QESRLS was valid for Korean students and the results showed that students with higher self-efficacy used more SRL strategies in language learning. The Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich et al., 1991) has been administered widely to investigate the relation between SRL and language learning. In Garrido-Vargas's (2012) study of Hispanic students in the United States, the results suggested that SRL was related to the academic achievement of these second language learners (reading score r = .41, p = .03; writing score r = .49, p = .007). Fukuda (2017) also found that SRL was related to language proficiency among Japanese university students. Three factors of learning strategies (Metacognitive strategies r = .307, p < .01, Effort regulation r = .332, p < .01, and Coping with problems r =. 270, p < .01) were significantly correlated with TOEIC scores. The results also showed significant differences in SRL between low- and high-proficiency learners in the motivational and learning strategy factors. Fukuda (2019) interviewed these students and elaborated on the characteristics that these low- and high-proficiency learners showed. These studies have added to the understanding of SRL and its relation to language learning. However, many of the studies are correlational studies and more research using qualitative methods is needed to capture the dynamic nature of SRL in language-learning contexts.

Research Questions

Previous studies have investigated SRL in language-learning contexts, but many studies are cross-sectional, and few have focused on the context of the first year in a two-year college in Japan. This study aims to demonstrate the development process of SRL skills among first-year students in a language-learning context at a two-year women's college in Japan based on longitudinal observation through interviews. This study focuses on the following research questions:

RQ1. What are the trajectories of SRL skill development for first-year college students?

RQ2. How do SRL skills impact students' time management for language self-study?

Method

The research reported was part of a doctoral project that investigated the SRL trajectories of students' first year in college. The participants of this

study were eight first-year students at a two-year women's college in Japan. The students were recruited in 2015 from a cohort of 250 students from English and business majors. One of the compulsory English courses for all first-year students met twice a week in eight course sections grouped according to students' TOEIC scores upon entrance to college. Four students from the top two groups, two from each major, were randomly selected and invited to participate in the study by email. They were categorized as "higher proficiency" and had TOEIC scores ranging from 300 to 400. Four students from the 6th and 7th groups from the top, two from each major, were also randomly selected and invited to participate in the study. They were categorized as "lower proficiency," with TOEIC scores below 300. Students from the 8th group were not chosen to avoid any issues associated with positionality as the author taught this class. Although it is rare that the researcher becomes a complete participant or a complete observer, it is important that the researcher be attentive to the effect of power that they might have on the research (Merriam & Tisdell, 2016). All of the eight students who were invited agreed to participate in the study and completed five interviews. Participants 1-4 were in the higher proficiency group and Participants 5-8 were in the lower proficiency group. Participants 1, 3, 5, and 6 were business majors and Participants 2, 4, 7, and 8 were English majors. The research proposal was submitted to and approved by the college's Institutional Review Board and informed consent was obtained from each participant before their first interview.

Five semi-structured interviews took place over two years: at the middle and the end of the first semester, after the summer break, at the end of the first year, and at the end of the second year. Each interview was conducted in Japanese and lasted approximately one hour. The interview questions covered the participants' learning history, experiences adjusting to college, study skills they used for their coursework and English self-study, and reflections on each semester (see Appendix for interview outline). Interviews were recorded and data were logged immediately after each interview. The interview recordings were transcribed and coded by the researcher using eclectic coding and hypothesis coding (Saldaña, 2013). For the first cycle coding, I used eclectic coding, a combination of two or more compatible coding methods where the researcher's "first-impression" (Saldaña, 2013, p. 188) responses serve as codes. For hypothesis coding, codes were developed based on Zimmerman's (2000, 2004) SRL model. First, single-case analysis was conducted to examine each participant's development and use of SRL skills over the first year. Then each participant was shown a brief description of the analysis and asked if any data were mistaken or if they wished

to have any data deleted from the analysis. After participant checking was finished, cross-case analysis was conducted to investigate similarities and differences within and across each proficiency group.

Findings

SRL Trajectories of Higher Proficiency Group

Among the four students in the higher proficiency group, the developmental trajectories of Participants 1 and 2 were consistent with the trajectory patterns of the highly self-regulated learners described by Zimmerman (2000). The trajectories of Participants 1 and 2 showed clear cyclical patterns of the three phases of SRL. For example, both participants had clear goals at the point of Interview 1, in the middle of their first semester. Participant 1 wanted to study abroad and find a job immediately after graduating from college. Participant 2 wanted to transfer to the college's affiliated university. At an early point, they had learned they needed high GPAs and higher TOEIC scores by the end of the first year to achieve their goals. They searched online for the TOEIC test center schedule and registered to take the TOEIC test every two to three months. They were able to organize their goals by setting key subgoals and planning ahead, which are key strategies in the forethought phase of SRL (Zimmerman, 2004). Participant 1 said: "I want to reach the TOEIC score 500. By the time I apply for the study abroad program, I figured I have three chances to take the TOEIC test. So, I immediately registered for the TOEIC test in June" (Interview 1). Participants 1 and 2 also showed effective use of the self-study strategies typically used by self-regulated learners in the performance phase. Participant 1 described her self-study methods:

> It is actually good to study with friends, laughing together and letting out stress. It is better studying that way on campus. Then after going home, I concentrate on reviewing materials on my own. It is like a two-stage structure. I study like this these days. (Participant 1, Interview 2)

This comment represents her use of SRL performance phase strategies such as self-instruction, help-seeking, motivational strategies, and environmental structuring (Zimmerman, 2000). Similarly, Participant 2 exemplified using motivational strategies:

> Recently, I set something to look forward to. I tell myself 'I can work hard because I will enjoy myself afterwards.' Otherwise,

> I cannot study hard….Even while I'm studying, I say 'I will eat chocolate if I finish this task.' I set these small rewards from time to time. Food is my incentive. (Participant 2, Interview 4)

After the performance phase, Participants 1 and 2 showed successful self-reflection behaviors. They reflected on which strategies were effective after each TOEIC test. Participant 1 considered reviewing her textbook just before the test was effective in June, and she found memorizing parts of speech effective in September. Her score continued to improve. She achieved her goals of reaching a TOEIC score of over 500, studying abroad, and getting a job as she had planned at the beginning of the first year. Participant 2 thought the test-taking skills she had learned in class in the first semester were effective for the TOEIC tests she had taken in June and September. Thus she continued using these skills, but her score improved only by 10 points in November. She realized that she should study more specifically for the TOEIC tests and that only doing coursework was not enough. She started studying a TOEIC textbook outside of class. Her TOEIC score improved by 130 points, and she was able to achieve her goal of transferring to a university. The trajectories of Participant 1 and 2 showed successful completion of one cycle of the SRL phases.

Participant 3 showed partial use of SRL skills but did not achieve her original goal. In Interview 1, her goal was to pass a certificate test each semester and thus get a recommendation from the college for a company. She had decided on this goal prior to entering college, because she had heard from some graduates that acquiring certificates would help her get a job and that the college provides students with good support for acquiring certificates (e.g., bookkeeping, business writing). She decided which certificate she would aim for each semester, took a Saturday course at the college for each certificate, and went to the learning support center when she needed help. However, she was not able to get either of the certificates she wanted and did not get a recommendation for a company. Looking back at her first year, she considered one reason for her failure:

> It was definitely a lack of study time. I took both the Saturday course and the certificate course in the curriculum, but I hardly ever studied other than in class. This college provides substantial support for passing these certificates, so I thought I would get them easily. It did not change the fact that I need to study by myself. (Participant 3, Interview 4)

Participant 3 also regretted that she had not studied hard to improve her TOEIC scores, which would be advantageous in job hunting (Interview 4). She had a clear goal and was able to set key subgoals, plan strategically, and seek support when necessary. She was able to reflect on her own learning; however, she could not plan based on the reflection and take action. It seemed that she was not able to develop the skills to self-instruct in the performance phase.

Participant 4 showed no development of SRL skills in the first year. In the middle of the first semester, her goal was rather vague: she wanted to get a job using English. Although she did not plan well and did not specifically reflect on her performance, her TOEIC scores continued to improve. This was due to the influence of her dormitory roommate. Her roommate was hard-working, and Participant 4, who described herself as a competitive person, started studying hard to compete with her. However, dormitory students had to change roommates each semester. With a different roommate in the second semester, Participant 4 suddenly had difficulty in keeping motivated to study. She had anxiety about not performing very well and felt depressed by the end of the second semester.

Zimmerman (2000) suggested that learners develop self-regulatory skills by learning them from Level 1 to 4 (observation, emulation, self-control, and self-regulation). Of the four participants in the higher proficiency group, Participants 1 and 2 seemed to reach the self-regulation level. Both had built on the skills they acquired and adapted to new needs. In contrast, Participants 3 and 4 seemed to be at the levels of emulation and self-control, respectively. Participant 3 may have seen some graduates successfully passing certificate tests and getting a recommendation from the college for a company. She thought she would do the same; however, she did not use skills that were necessary to actually study for and pass the exams. Thus, her SRL skills were most likely at the emulation level. Participant 4, who did not seem to develop any SRL skills, was easily influenced by the people surrounding her. In the first semester, when she had an inspiring roommate, she just followed what she did. Although she mentioned in interviews that she had used some self-instruction skills in high school (e.g., highlighting key words, watching movies in English), she did not adopt them in college, even when she had difficulty motivating herself in the second semester. Therefore, some of her SRL skills were at the self-control level, in which learners display a skill under structured conditions.

SRL Trajectories of Lower Proficiency Group

The participants in the lower proficiency group showed how lack of success in one SRL phase leads to dysfunction in the following phase. The participants in the lower proficiency group had unclear or unrealistic goals in the forethought phase. Throughout her first year, Participant 5 was troubled that she could not feel the purpose of studying and lacked a clear goal. She was sure that she did not want to continue studying and wanted to work after graduating from the two-year college, but she was unsure what kind of job she wanted. At the end of her first year, she said "I hear that GPA is very important even for job hunting. However, I am not intending to transfer to university…I wonder why I study" (Participant 5, Interview 4). She could not concentrate on studying and did not perform well academically. Participants 6 and 7 were also uncertain what they wanted to do after college, so they wanted to decide after taking some courses and studying abroad in the first year. However, this posed a great challenge for them as the college pressures students to choose their career plans early. At two-year colleges, most students who plan to work after graduation have to start the job-hunting process by the end of the first year. For students who wish to transfer to a university, the option of transferring depends on their first-year GPA. If students are unsure of their career plans and do not do well academically from the first semester, it is difficult to attain goals that become clearer in the second semester. Participant 6 described her confusion:

> [I came to this college] because I can study business [as a major] as well as women's studies and philosophy as general studies subjects. If I want to study either of them more, I may think of transferring to the university. I'm still debating…. However, I found out that these subjects [women's studies and philosophy] are basically for the second year students….I also wanted to study abroad. Then I was going to think about transferring after studying abroad. But I found out the timing is not very good…. The decision-making process was not done the way I had thought. It was very shocking. (Participant 6, Interview 1)

As for Participant 8, her original goal was to retake an entrance examination she had previously not passed for the public university she wanted to attend. However, by the end of the second semester, she realized that this goal was too difficult for her and decided to aim for a transfer to the college's affiliated university, which was a more achievable goal. Not having clear

goals leads to difficulty in setting key subgoals and planning strategically in the forethought phase of SRL. Only Participant 8 employed self-reflection and set a goal by adapting to new needs.

Another challenge that three participants (Participants 5, 6, and 7) in the lower proficiency group had was not being able to use effective self-instruction skills in the performance phase of SRL. Both Participants 5 and 6 showed light use of organizing and transforming skills in the beginning (i.e., note-taking). Participants 5 and 6 tried to find note-taking skills that suited them in the first semester. Participant 5 had heard from a senior student about taking notes using the iPad memo function and tried it. However, she stopped and did not mention note-taking in the second semester. Participant 6 also tried organizing lecture notes for her business courses, but she always crammed at the last minute. She barely finished organizing her notes before the first semester final exams. She also did not mention note-taking in the second semester. As for English studies, Participants 5 and 6 both recognized studying for the TOEIC test was important. Participant 5 bought a TOEIC textbook and tried to study outside class in the first semester, but her motivation did not continue as she had problems with friends in the second semester (as described later in this section). Participant 6 also had difficulty in persisting. She tried many self-study methods as taught by a professor but could not continue any of them. She took a Saturday course for the TOEIC reading section but could not stop sleeping during class. She said, "I don't even do homework, so my score cannot be improving" (Interview 3). Participants 7 and 8 majored in English and believed that the coursework would not lead to TOEIC score improvements. Participant 7 thought self-study was more important than learning in class. She did not do class assignments but focused on vocabulary learning outside class. At first, she used an English-English dictionary to look up words. Then she used a smartphone application for this purpose instead, but her usage gradually declined because her part-time job decreased the time she had to study. Participant 7 tended not to take the coursework seriously and found persistence and time management difficult. Unlike Participant 7, Participant 8 recognized the good points of the coursework and managed to balance the coursework and self-study in the second semester. She did dictation exercises for homework every day, saying "Dictation homework takes much effort, but if I listen to the teacher, my score is likely to improve" (Interview 4). However, she did not feel dictation was enough to increase her TOEIC score, so she studied reading comprehension questions for the TOEIC test and vocabulary on her own. These study methods (i.e., rehearsing and memorizing) were not frequently used

among high achievers in Zimmerman and Martinez-Pons' study (1986), but they belong to the SRL strategies of the performance phase.

Another notable characteristic of the participants in the lower proficiency group was how they coped with emotional stress. Participant 5 felt that her roommate was very intelligent and that she could not keep up with her when studying together in the morning. In addition, in the second semester, she found herself having trouble with friends in her school club and could not concentrate on her studies. She could not develop effective strategies to cope with her emotional stress. Participant 5 described her situation:

> After the college festival, I was motivated to study hard, but so many things happened. I had trouble with friends and I felt depressed. Since then, I have not been able to move forward… because of many troubles, I keep thinking during class, so I could not listen to the lectures attentively. (Participant 5, Interview 4)

Participant 6 used to worry about keeping up academically when her dormitory roommate studied until late at night. However, she decided that sleeping was more important to her and stopped competing with her roommate. Participant 6 stated:

> Some friends study on no sleep or study until 2 or 3 o'clock during the night and wake up at 6 o'clock in the morning. But I have never stayed up through the night and I try not to, because both sleeping too much and little sleep cause migraines. But when I hear friends studying like that, I feel really nervous. (Participant 6, Interview 2)

In the first semester, Participant 6 also joined a morning study group. However, she could not keep up with the other students and became sleepy in class, so she stopped attending. She did not make an additional effort to make time to study during the day instead and she regretted this at the end of the first year, saying "I should have made more study plans" (Participant 6, Interview 4). Participant 7 did not worry about friendship as much. However, she seemed to have difficulty in asking for help when she needed support. She was thinking of studying abroad but she hesitated to visit the college office to obtain the necessary information until the end of the first year. Her indecisiveness and lack of persistence seemed to be rooted in her lack of confidence and in her self-efficacy, seemingly preventing her from

taking action, exhibiting how self-efficacy is closely linked with SRL (Pajares, 2008; Zimmerman, 2011). Participant 8 used to be distracted by her noisy dormitory mates and worried if they excluded her. During the second semester, she eventually decided not to worry about her dormitory mates and started focusing on her studies.

In the self-reflection phase of SRL, three of the participants (Participants 5, 6, and 7) of the lower proficiency group did not reflect on themselves well and tended to attribute their results to external factors such as teachers and staff members. Participants 5 and 6 regretted not being proactive as well as being sleepy during class but blamed their teachers for their sleepiness. Participant 5 said, "The lecture pace is so fast…he [the professor] explains everything so fast, so I cannot even take notes" (Interview 1). Similarly, Participant 6 also claimed her professors had poor teaching skills, saying, "They do not have a license to teach, so…I should not say this, but some teachers teach badly. I do not understand [their explanations], to be honest…then I eventually feel sleepy in class" (Interview 1). Participant 7 also did not reflect on her own behavior and tended to blame others for her results. She faulted the office for not informing her of the job hunting orientation schedule in advance, because she already had scheduled her part-time work hours. In contrast, Participant 8 tended to put responsibilities on others as well in the beginning, she started self-reflecting more and changed her behavior based on her reflections later in the first year. Participant 8 gradually started to exhibit the self-analysis skills that self-regulated learners use in the reflection phase of SRL (Zimmerman, 2000).

Of the four levels of SRL development (Zimmerman, 2000, 2013), three of the participants (Participants 5, 6, and 7) in the lower proficiency group seemed to be able to do some observation and emulation, but only Participant 8 progressed to the self-regulation level. Participant 5 seemed to observe other students and tried to emulate some skills. However, she did not reach the level of self-control, in which learners utilize the skills under structured conditions (Zimmerman, 2000, 2013). Participant 6 also observed other students, but she did not continue studying like them. Participant 7 did not seem to observe or emulate some skills of others. In contrast, Participant 8 was able to independently find self-instruction strategies based on self-reflection as well as strategies to cope with emotional stress.

SRL Skills and Self-Study

The trajectories of the eight participants showed that their SRL skills were related to allotting time for independent language learning. All of the

participants found it challenging to balance their coursework and self-study, particularly the business majors. At this college, improving TOEIC test scores was important for students because the scores were used in the selection processes for transfers to the affiliated university and enrollment for study abroad programs. Additionally, a high TOEIC test score would be advantageous in job hunting. This greatly affected the participants' goal setting and planning about their studies. However, the participants who reached the self-regulation level were successful in balancing their course studies and language learning outside class. Participant 1 (higher proficiency), who majored in business, structured her self-study method, and learned to set aside time to study English. Similarly, Participant 2 (higher proficiency) and Participant 8 (lower proficiency), who both majored in English, put priority on their coursework and also managed their self-studies. However, Participant 3 (higher proficiency) and Participants 5 and 6 (lower proficiency), all business majors, had difficulty in structuring their environments to study both business and English outside of their coursework. Participant 7 (lower proficiency, English major) did not set aside enough time to self-study and had difficulty in persisting. Participant 4 (higher proficiency), who majored in English, was only motivated to study when she saw her friend studying.

Among the eight participants, five (Participants 4, 5, 6, 7, and 8) lived in a dormitory and experienced living away from family for the first time. They had to manage chores, dormitory responsibilities (e.g., cleaning public spaces, keeping the curfew) as well as interpersonal issues. The other three participants (Participants 1, 2, and 3) lived with family but had to adjust to long commutes of 1 to 2 hours that took time and energy from them. Thus, their new circumstances posed additional physical and mental challenges. Some participants tried to balance their studies and extracurricular activities. A few participants (Participants 1 and 2) managed to engage in club and student government activities and still set aside time for study. However, for many of the participants (Participants 3, 5, and 7), club activities and part-time jobs took most of their time and energy outside of class.

SRL skills played an important role to manage competing demands. Use of effective SRL skills enabled participants to make the time for self-study and other responsibilities. Some participants had difficulty in making the balance between these needs due to lack of SRL skills.

Discussion

The trajectories of the eight participants in their first year of college suggest that various emerging demands affected their SRL development.

Zimmerman (2008) explained that goal setting is important in SRL as it is relevant to all three of its learning phases. Goals set in the forethought phase affect strategy implementation and self-observation, and self-reflection leads to goal setting in the next cycle. Setting goals in the forethought phase also affects self-motivation beliefs, which in turn affect the performance phase. This explanation corresponds to the findings of this study.

The participants who struggled to self-regulate their learning tended to have vague goals at the beginning of the first year. Having unclear goals made the transition to college particularly challenging in the context of this study, a two-year college in Japan. Although students at this college could delay deciding on their plans until the end of the first year, deciding earlier is better as the selection is based on GPAs, TOEIC test scores, and certificates achieved in the first year. Although the participants in this study were under pressure to make choices about their future early, some needed time to understand what they wanted (e.g., Participants 6 and 7). The pressure led these participants to be unable to focus on how they would spend their first year.

Academic demands that college places on students are one of the biggest emerging stresses. Students often struggle to adjust to lecture-style courses and the number of assignments. The participants in this study had to deal with these demands and find time for self-study. To gain advantages in transferring to a university and job hunting, they had to improve their TOEIC scores and pass exams for certificates in subjects such as bookkeeping and business writing. The participants needed to carefully plan to structure their environments to organize their self-studies so they could achieve these key subgoals.

Many first-year students have to deal with other emerging demands in adjusting to college. All of the participants in this study struggled to adjust to new living environments and new social obligations. Unless they had effective planning and environmental structuring strategies, it was difficult for them to manage all of these demands.

Interpersonal issues sometimes became a source of emotional distractions. Participants mentioned various struggles in dealing with their dormitory mates. As Frenzel and Stephens (2013) pointed out, negative emotions can lead to less effective use of strategies. Among the participants in this study, comparing their performance to that of others often caused negative emotions (e.g., Participant 5 and 6). Such comparisons can deemphasize self-observation as it reflects using normative criteria in the self-reflection phase rather than comparisons to one's own previous performance (Zim-

merman, 2000). A lack of SRL strategies can cause difficulty in controlling emotions and managing distractions from goals; the effective use of SRL strategies is essential to balance these demands.

What Teachers Can Do

This study suggests that SRL skills are important for a successful transition to college in the Japanese context. Teachers can be aware of the importance of understanding the context of their first-year learners, especially the difficulties that lower proficiency learners may be facing. There are some approaches that teachers can take to help first-year students develop into self-regulated learners. They can explicitly teach motivation and learning strategies in the first semester. Seli and Dembo (2019) wrote a comprehensive textbook about self-regulated learning strategies for first-year students; teachers can adapt the strategies where necessary and teach them in the Japanese context. Teachers can also have first-year students review their time use and emphasize time management to prioritize self-study. Teaching time management skills may be more valuable if it is done at the beginning of the second semester because this is a good time for students to reflect what they would have done differently in the first semester (Thibodeaux et al., 2017). In addition, teachers can make learning materials available before class so that students can preview the materials in the forethought phase. Providing a curriculum with a transparent structure can help students to plan and prepare for the class beforehand (Zhou & Rose, 2021). Fukuda (2019) suggested helping students with setting short-term goals may be important for lower-proficiency learners. By setting short-term, achievable goals, students can gain self-efficacy (Fukuda, 2019).

Conclusion

This case study investigating the SRL development of eight first-year students mainly suggests three points. First, Zimmerman's (2000, 2004) social cognitive model of SRL is applicable to first-year college students in the Japanese context. The SRL developmental trajectories of the participants also fell into one of the four levels of SRL development: observation, emulation, self-control, and self-regulation (Zimmerman, 2000). Those who were able to set key subgoals, based on their long-term goal in the forethought phase, were able to utilize effective self-instruction strategies in the performance phase. Based on self-reflection, they successfully set their next goals and the strategies to achieve them. In contrast, those who had vague goals

in the forethought phase struggled to set key subgoals and strategies, could not find self-instruction skills suited for them, failed to self-reflect, and tended to attribute their failure to others. Second, SRL skills influenced their self-study in English. Both business and English major students struggled to balance their coursework and language learning outside class. However, the participants who were good at setting goals and using strategies showed the use of effective environmental structuring strategies to set aside time for self-study. Third, many emerging demands (e.g., academic demands, the pressure of making career choices, new living environments, and interpersonal issues) affected or interfered with SRL development, and at the same time, SRL skills were essential to manage these demands. SRL is greatly influenced by environmental factors.

Although this longitudinal observation of eight first-year students helped gain a deeper understanding of the transition to college and the SRL development process in the Japanese context, the sample is still limited. More investigation with a broader sample can enhance the understanding of SRL development in Japan. Furthermore, effective intervention should be explored to support students' successful transition to college.

Yukiko Ishikawa is an associate professor in Soka Women's College. Her research interests include self-regulated learning and language advising.

References

Bowman, N. A., Jarratt, L., Jang, N., & Bono, T. J. (2019). The unfolding of student adjustment during the first semester of college. *Research in Higher Education, 60*, 273–292. https://doi.org/10.1007/s11162-018-9535-x

Brasor, P. (2017, April 8). Higher education keeps overreaching. *The Japan Times*. https://www.japantimes.co.jp/news/2017/04/08/national/media-national/higher-education-keeps-overreaching/

Conefrey, T. (2018). Supporting first-generation students' adjustment to college with high-impact practices. *Journal of College Student Retention: Research, Theory & Practice, 23*(1), 139–160. https://doi.org/10.1177%2F1521025118807402

Doi, S. (2017, August 24). Heri yuku tandai gakusei-sū, 90-nendai no 4-bun'no 1 'ao tan' mo boshū teishi e [Junior colleges decreasing, student number a quarter Aoyama Gakuin Women's Junior College ceases admission]. *Asahi Shimbun,* 011.

Frenzel, A. C., & Stephens, E. J. (2013). Emotions. In N. C. Hall & T. Goetz (Eds.), *Emotion, motivation, and self-regulation: A handbook for teachers* (pp. 1–56). Emerald Group.

Fukuda, A. (2017). The Japanese EFL learners' self-regulated language learning and proficiency. *Journal of Pan-Pacific Association of Applied Linguistics, 22*(1), 65–87. https://files.eric.ed.gov/fulltext/EJ1188004.pdf

Fukuda, A. (2019). What less-proficient EFL learners tell us about their language learning: Qualitative analysis of self-regulated learning. *Journal of Pan-Pacific Association of Applied Linguistics, 23*(1), 103–126. https://files.eric.ed.gov/fulltext/EJ1224176.pdf

Garrido-Vargas, M. (2012). *Relationship of self-regulated learning and academic achievement among English language learners*. [Doctoral dissertation, The University of Arizona]. UA Campus Repository. http://hdl.handle.net/10150/242375

Harada, K. (2015, March 16). Universities struggle to cope with shrinking population and globalization. *Nippon.com*. https://www.nippon.com/en/features/h00095/

Harvey, L., Drew, S., & Smith, M. (2006). *The first-year experience: A review of literature for the higher education academy*. Higher Education Academy.

Kim, D., Wang, C., Ahn, H. S., & Bong, M. (2015). English language learners' self-efficacy profiles and relationship with self-regulated learning strategies. *Learning and Individual Differences, 38*, 136–142. https://doi.org/10.1016/j.lindif.2015.01.016

Krsmanovic, M., Cox, T. D., & Johnson, J. D. (2020). Who improves most? The differences in first-year students' learning attitudes and behaviors measured by college success factor index. *Journal of the Scholarship of Teaching and Learning, 20*(2), 1–14. https://doi.org/10.14434/josotl.v20i2.27446

Kyodo. (2021, September 28). Places at 46% of private Japanese universities not filled in 2021. *The Japan Times*. https://www.japantimes.co.jp/news/2021/09/28/national/university-places-unfilled/

Matsumoto, R., Miyazato, S., Uto, H., Terazono, R., Hayata, T., Hirashima, K., Sonoda, M., Tanigawa, S., Mitsuda, T., & Ueoozono, A. (2011). Yokuutsu-kan to karada fuchō-kan kara mita joshi tanki daigakusei no seishin-teki kenkō no genjō to kadai [The present situation and problems of mental health from viewpoint of depressed mood and subjective physical problems on Kagoshima Women's Junior College students]. *Bulletin of Kagoshima Women's Junior College, 46*, 193–203. https://kwjc.repo.nii.ac.jp/?action=pages_view_main&active_action=repository_view_main_item_detail&item_id=306&item_no=1&page_id=13&block_id=17

Matsutani, M. (2012, January 10). Student count, knowledge sliding. *The Japan Times*. https://www.japantimes.co.jp/news/2012/01/10/reference/student-count-knowledge-sliding/#.VBZ3P_l_t8

Merriam, S. B., & Tisdell, E. J. (2016) *Qualitative research: A guide to design and implementation* (4th ed.). Jossey-Bass.

Pajares, F. (2008). Motivational role of self-efficacy beliefs in self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and application* (pp. 111–140). Erlbaum.

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)*. NCRIPTAL (The National Center for Research to Improve Postsecondary Teaching and Learning), School of Education, University of Michigan.

Pittman, L. D., & Richmond, A. (2008). University belonging, friendship quality, and psychological adjustment during the transition to college. *The Journal of Experimental Education, 76*(4), 343–361. https://doi.org/10.3200/JEXE.76.4.343-362

Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed.). Sage.

Seli, H., & Dembo, M. H. (2019). *Motivation and learning strategies for college success: A focus on self-regulated learning* (6th ed.). Routledge. https://doi.org/10.4324/9780429400711

Stephen, J. S., & Rockinson-Szapkiw, A. J. (2021). A high-impact practice for online students: The use of a first-semester seminar course to promote self-regulation, self-direction, online learning self-efficacy. *Smart Learning Environments, 8*, Article 6. https://doi.org/10.1186/s40561-021-00151-0

Thibodeaux, J., Deutsch, A., Kitsantas, A., & Winsler, A. (2017). First-year college students' time use: Relations with self-regulation and GPA. *Journal of Advanced Academics, 28*(1), 5–27. https://doi.org/10.1177/1932202X16676860

Upcraft, M. L., Gardner, J. N., & Barefoot, B. O. (2005). Introduction. In M. L. Upcraft, J. N. Gardner, & B. O. Barefoot (Eds.), *Challenging & supporting the first-year student: A handbook for improving the first year of college* (pp. 1–12). Jossey-Bass.

Wang, C., & Bai, B. (2017). Validating the instruments to measure ESL/EFL learners' self-efficacy beliefs and self-regulated learning strategies. *TESOL Quarterly, 51*(4), 931–947. https://www.jstor.org/stable/44984799

Wilcox, P., Winn, S., & Fyvie-Gauld, M. (2005). 'It was nothing to do with the university, it was just the people': The role of social support in the first-year experience of higher education. *Studies in Higher Education, 30*(6), 707–722. https://doi.org/10.1080/03075070500340036

Wolters, C. A., & Brady, A. C. (2021). College students' time management: A self-regulated learning perspective. *Educational Psychology Review, 33*, 1319–1351. https://doi.org/10.1007/s10648-020-09519-z

Yamada, R. (2012). *Gakushi katei kyōiku no shitsu hoshō ni mukete* [Toward quality assurance of undergraduate education]. Toshindo.

Yamada, R. (2019). *2040-Nen daigaku kyōiku no tenbō* [2040 outlook for university education]. Toshindo.

Zhou, S., & Rose, H. (2021). Self-regulated listening of students at transition from high school to an English medium instruction (EMI) transnational university in China. *System, 103*, Article 102644. https://doi.org/10.1016/j.system.2021.102644

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press.

Zimmerman, B. J. (2004). Sociocultural influence and students' development of academic self-regulation: A social-cognitive perspective. In D. M. McInerney & S. Van Etten (Eds.), *Big theories revisited* (pp. 139–164). Information Age.

Zimmerman, B. J. (2008). Goal setting: A key proactive source of academic self-regulation. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and application* (pp. 267–296). Erlbaum.

Zimmerman, B. J. (2011). Motivational sources and outcomes of self-regulated learning and performance. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 49–64). Routledge.

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist, 48*(3), 135–147. https://doi.org/10.1080/00461520.2013.794676

Zimmerman, B. J., & Martinez-Pons, M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal, 23*(4), 614–628. https://doi.org/10.3102/00028312023004614

Appendix

Interview Outline

First Interview

(Adjusting to college)

1. How is coming to college different from high school? What do you find most challenging in your college life now? How do you deal with the challenge?
2. Which classes do you find most enjoyable and which do you find most difficult?
3. Are English classes here different from high school English classes?
4. What is your personal goal? Are you doing anything to reach that goal?

(Learning history)

1. Experience in high school. Which subject did you like the most? How did you study for the subject? What was the English class like?
2. Experience of the entrance exam. How did you enter this college? Was it a recommendation by your high school? Which did you take, the Japanese or English test? How did you prepare for the entrance exam? Did you go to a cram school?
3. Experience of studying English. Did you go to lessons outside school? How long did you go to these lessons?
4. Experience abroad. Have you been abroad? Was it a study abroad or a short trip with family? Where and how long did you go? Did you use English there?

Second Interview

1. Reflecting on the first semester, which area would you like to change? Did you see any changes in yourself compared to the beginning? In what area did you do well? Who did you talk to when you had a problem?
2. How did you manage your time, doing assignments outside of class? What learning style did you find suitable for yourself? Where did you spend your time most after school? Did you have any strategies to keep your motivation to study?

3. What kind of support would you like to have in college?
4. What is your plan for summer? Based on your experience in the first semester, do you have any goals from now?

Third Interview

1. Please tell me your experience in summer etc. after we met last time. Did you see any changes in yourself since last time? How did the experience change your study habits and motivation to study? How did you deal with the change?
2. What were your study habits like during the summer and the college festival?
3. Did you have any strategies to keep your motivation to study?
4. What are your goals from now?

Fourth Interview

1. Looking back at your first year in college, how did you change over the year? What would you have done differently?
2. What were your turning points during the first year? What was the biggest challenge you faced? What was the biggest difference academically from high school?
3. Which study habits and motivational strategies do you think you acquired in the first year?
4. What advice would you give to incoming students? What kind of support do you think would be helpful for incoming students?
5. What are your goals for the second year?

“We Wouldn’t Have Hired Me Without a Solid Research Background”: Higher Education English Language Teachers’ Research Engagement

Sarah Louise Mason
Macquarie University
Alice Chik
Macquarie University
Peter Roger
Macquarie University

Although English as a foreign language (EFL) teachers in higher education have traditionally focused on teaching, they are increasingly required to engage in research (Reyes-Cruz & Perales-Escudero, 2016). The present article, therefore, explores how such teachers understand their research engagement. The study collected qualitative data from 63 such respondents in Japan. Findings revealed that requirements for research engagement differed widely depending on the institution. In addition, according to the respondents, although early career research experiences were perceived to be largely negative, many teachers overcame this hurdle, with most respondents viewing their research engagement to be pivotal for career development.

Keywords: narrative frame; neoliberalism; research engagement; research practices

https://doi.org/10.37546/JALTJJ45.1-4

JALT Journal, Vol. 45, No. 1, May 2023

Universities have fulfilled two roles for centuries: the creation of knowledge and the education of students. Until recently, teachers of EFL at universities were often only expected to fulfill the second of these roles (Borg, 2013; Reyes-Cruz et al., 2018). However, many governments have extended English language teachers' work to include research. This increased complexity in the work of higher education English language teachers (HEELTs) can be traced to neoliberalist academic culture intimately connected to knowledge-based economies. The essential characteristic of a knowledge economy is that it relies more heavily on its intellectual capabilities than on natural resources or physical labor (Powell & Snellman, 2004). For example, reliance on the export of medical technology rather than the export of coal would be a marker of a knowledge economy. Therefore, it becomes imperative for governments to increase the quantity of knowledge output as frequently as possible by transforming universities to become or remain globally competitive.

In the neoliberal context, the success of a higher education institution is measured in terms of knowledge output, performance data, global competitiveness, and income (Kubota, 2016). Governments have been requiring universities to transform their roles by adopting neoliberalist management practices. For example, institutions achieve extensive accountability through monitoring the manner in which research output is externally funded, controlled by the government, and divorced from teaching (Olssen & Peters, 2005). In concrete terms, in order to maximize cost efficiency, university administrations have implemented flexible workforce policies in which employees can be let go at will. This has resulted in many academics being employed in temporary positions with limited access to secure employment. Meanwhile, as academics face greater pressure to be productive in their research endeavors by numerically reporting their publications and presentations to meet strict evaluation criteria, university administrators closely monitor faculty applications for government and private funding. Therefore, it can be evidenced that HEELTs are placing an increased international emphasis on research output; for example, in China (Bai & Hudson, 2011; Yuan, 2017), Mexico (Reyes-Cruz et al., 2018), and Turkey (Borg, 2007).

The authors' personal and researcher histories underpin the motivation for the study. The first author has taught English as a foreign language (EFL) in Japan for more than 20 years, eight of those in higher education. The second author previously worked in higher education in Hong Kong for many years and has considerable experience supervising novice researchers in both Hong Kong and Australia as they undertake terminal degrees. The

third author has supervised PhD candidates located in Japan and Australia and worked in higher education language departments for many years. All the authors have observed that some HEELTs enthusiastically embrace research, while others do not. These experiences have led them to ponder how and why some HEELTs become eager researchers, and furthermore, how HEELTs themselves perceive their behavior in regard to research. For example, how much research they conduct and the type of research practices they engage in. These observations led them to explore the current situation in Japan, a country with an immense higher education sector and a stated commitment to English language teaching at all levels of education. Our position as knowledge workers, rather than as members of management, may well have influenced our interpretation of the data. Finally, it should be noted that as we are researchers ourselves, we have a personal interest in discovering what factors contribute to the successful conduct of research.

The present study investigates how HEELTs conceptualize *research engagement* in Japan. These educators play an important role in Japanese higher education — in many cases teaching compulsory English language courses required for graduation — and represent a wide variety of disciplinary backgrounds. In defining *research engagement*, we follow Borg (2015), who distinguished between two types: engaging *with* research and engaging *in* research. Engaging *with* research tends to be more passive; for example, it involves reading research publications and attending presentations given by others. Conversely, engaging *in* research involves actively doing one's own research. Our study focuses on engagement in research, which is understood as research productivity, or the writing, presenting, and publishing of research.

The professional roles of academics working in higher education are in a state of flux globally due to neoliberal policy implementation. This precariousness has led to HEELT roles being likely to vary depending on context. Therefore, the present study aims to elicit HEELTs' understandings of their particular institutional contexts and experiences of research engagement, with a view to understand how it is perceived in the Japanese context. We want to know, from their point of view, if, or how, research has been integrated into their careers, how they perceive their institutional contexts, and what their attitudes toward and purposes are for doing research. It is important to know about these perceptions because unless research is something that HEELTs embrace and incorporate into their perception of their jobs, higher education institutions (HEIs) and governments are unlikely to realize their goals. Without a clear understanding of precisely how effectively

their new roles are being carried out, role conceptualization is unlikely to be successful for the individuals concerned and for the research-related goals of the HEIs.

The purpose of this study is to survey a population of HEELTs to obtain a sense of how they relate their professional lives to their research, whether positive or problematic. In short, we seek to investigate HEELTs' perceptions of research as it relates to their professional roles in higher education in Japan. Therefore, our research question is:

RQ. How do HEELTs understand their research engagement in higher education institutions in Japan, and how might these understandings illuminate the institutional forces at play in their research engagement?

Neoliberalism and Research in HEIs

We begin our literature review with a brief overview of neoliberalism and follow by examining the impact of neoliberalism on higher education. Finally, we describe the effect of neoliberalism on HEELTs' research engagement.

Neoliberalism is an economic theory in which trade is undertaken by private individuals for profit, as opposed to being undertaken by the state for social welfare. Neoliberalist economic theory has three main guiding characteristics: privatization, deregulation, and individual or corporate responsibility (Block et.al., 2012). Privatization of entities is encouraged because corporations and individuals are believed to operate more efficiently than governmental bodies. Deregulation of markets is also seen as a way to increase market efficiency through facilitating competition between providers of goods and services. Finally, both companies and individuals are held responsible for their own success or failure in the neoliberalist free market.

With respect to HEIs globally, there are three main domains where neoliberalism's influence can be observed: institutions, academics, and students (Kubota, 2016). HEIs experience the effects of neoliberalism through privatization, marketization, and corporatization; academics are required to pursue research grants and provide evidence of publication and are rewarded if their works are cited frequently, and students must pay higher fees and study curriculum designed "for developing human capital" (Kubota, 2016, p. 488). Working in higher education, HEELTs cannot escape neoliberalism. Published studies consider on a granular level the precise institutional requirements for research from HEELTs, specific ways they reward or penalize teachers for their research output or lack thereof, and ways in which they

facilitate or hinder research. We also summarise what the same research teaches us about the attitudes and motivations, or purposes, for research held by HEELTs.

HEELTs have been required to become active in research as university and government guidelines have become more stringent in EFL programs in many countries (see, e.g., Barkhuizen, 2021, on Colombia; Borg & Liu, 2013, and Xu, 2014, on China; Reyes-Cruz & Perales-Escudero, 2016, on Mexico). Research requirements may concern both research productivity and the acquisition of funding. These requirements are linked to penalties and rewards in a variety of ways, depending on the institutional context. This can be seen, for example, in Reyes-Cruz and Perales-Escudero's (2016) qualitative, interview-based study of 26 HEELTs working at a public university in Mexico. The study reported various rewards for productivity in research, generally monetary in nature, but it also stated that teachers can also be rewarded in terms of status through designation as members of elite research groups. Correspondingly, failure to publish regularly is directly linked to the threat of termination.

Xu (2014) also identified a range of rewards and penalties. In a mixed methods study that investigated the research practices of 104 HEELTs in China, Xu reported a range of rewards for research productivity, including bonus payments, honorary titles, and praise. Conversely, her participants also reported a range of penalties; for example, failure to publish could result in disqualification from receiving "excellent teacher" evaluations or denial of promotion. It is apparent that institutions use various rewards and penalties to shape the behavior of their HEELTs.

In league with the aforementioned, transparent and direct efforts to shape HEELTs' research behavior can also either facilitate or impede engagement. Little research has explored features that successfully facilitate EFL research, although Borg & Liu (2013), in a study of 725 Chinese college English teachers, reported some satisfaction with funding arrangements. Their findings also indicated that previous experience with research and mentorship are perceived to be significant facilitators of research efficacy (Reyes-Cruz & Perales-Escudero, 2016). Nonetheless, most research focuses on obstacles to research engagement, such as a lack of expertise (Allison & Carey, 2007), or a gap between institutional expectations for research and the actual skill level of the HEELTs involved (Borg & Liu, 2013). Heavy workloads, particularly teaching workloads, are a common hindrance because they inevitably reduce the time available for research (Allison & Carey, 2007; Borg, 2007; Reyes-Cruz et al., 2018; Xu, 2014). Local research culture

also plays a significant part. Borg and Liu (2013) reported a highly individualistic research culture that pitted researchers against each other as they competed for promotion. They argued that such a culture leads to secrecy and inevitably prevents the formation of collegial support on Chinese campuses that might otherwise facilitate research engagement. In sum, there is a tendency for previous research in the field to focus on challenges to rather than facilitators of research.

Neoliberalist ideologies that encourage competition and strictly measured output, and the administrative responses of universities, in particular publication requirements concerning the frequency of publication and venue of publication, frame both HEELT attitudes to research and their purposes for engaging in it. Research in the field frequently reports negative attitudes and instrumental purposes. Allison and Carey (2007), in a study of 22 ESL and English for academic purposes (EAP) teachers working at a university language center in Canada, found a high level of ambivalence toward research. The teachers sought recognition but were insecure because they felt that effort spent on research would not be rewarded. Conversely, amongst those who wanted to do research, their purpose was to progress in their careers. Similarly, Borg and Liu (2013) found that their participants were ambivalent or often had negative attitudes to research. They also determined that a primary motivator for research engagement by the teachers was instrumental: they read research and performed research in the lead-up to periodic promotion opportunities. Their findings were echoed by Xu (2014), who described both a lack of interest and a lack of motivation in her respondents. These Chinese studies found that negative attitudes toward research went hand in hand with instrumental purposes for research. Indeed, Borg and Liu (2013), Bai and Hudson (2011), and Xu (2014) all reported a desire for graduation or promotion as the primary purpose for undertaking research.

The Study

Context

Japan is one of many countries seeking to transform its higher education workforce through increased marketization and internationalization of its higher education sector. Japanese educational policy has explicitly sought to raise the standard of research at universities by encouraging increased publication of research by faculty, increased credentialization of faculty (the percentage of faculty with higher degrees), and increased globalization of faculty (hiring from outside of Japan) to increase participation in international academic networks and collaborative opportunities (Brotherhood

et al., 2019; Wadden & Hale, 2019). Therefore, HEELTs in Japan may face increasing pressure to acquire higher degrees and to produce research in their field. The present study, rather than assuming this to be the case, aims to clarify HEELT perceptions. Japanese higher education institutions rely on receiving grants-in-aid for scientific research from the Japanese government. These grants-in-aid, known as *kakenhi,* are the only competitive government funding scheme to support researchers in all academic disciplines in higher education.

Methodology

This study focuses on professional autobiographies collected from HEELTs working in Japan. A narrative approach was selected as it allows researchers "to uncover commonalities that exist across stories" (Polkinghorne, 1995, p. 14) while obtaining rich data typical of qualitative approaches. Data were collected as part of a larger research project on HEELTs in the Japanese context. The aim of the present study was to elicit professional histories in story form from a relatively large number of HEELTs to identify past and present research experiences, contextual factors that might facilitate or hinder researcher engagement, as well as attitudes to research. Respondents were requested to provide demographic information such as their geographical location and place of employment (see Appendix). Narrative data was collected via a narrative frame (Barkhuizen & Wette, 2008; Hiratsuka, 2018; Swenson & Visgatis, 2011).

Most pertinent to the present study are the works by Barkhuizen (2009) and Xu (2014). Both explored HEELT research engagement and utilized narrative frames, which are a tool for collecting brief stories from respondents. In essence, they are templates, story skeletons that the respondents flesh out by inserting their own personal experiences. According to Barkhuizen (2009), a primary advantage of narrative frames is that they allow researchers to collect stories from a relatively large number of respondents, while simultaneously confining the stories to those concerns most central to the research at hand. This overcomes one problem associated with narrative studies in that they are usually only able to investigate the experiences of a limited number of participants collected by means of interview methodologies. Barkhuizen (2009) and Xu (2014) used narrative frames to establish a small corpus of HEELT narratives that could then be mined to reveal issues of concern regarding HEELT research. A goal of the present study was to establish what issues are most salient in researcher engagement in Japan prior to undertaking a larger study involving more detailed and intimate

research through interviews with a small group of participants. Therefore, narrative frames were selected as an appropriate tool for collecting data.

The frame was distributed in November 2019, initially to members of a sub-group of a language teacher professional organization whose target membership was faculty at colleges and universities. It was also distributed to six professional acquaintances of the first researcher who fit the necessary requirements for participation (i.e., they taught English at the tertiary level in Japan). Seventy-four responses were collected, of which 63 were completed in full.

Narrative Frame Data

A narrative frame is a template, written in paragraph form and composed of *sentence starters* or the first part of a sentence which the respondent is required to complete. As with stories generally, narrative frames contain characters, locations, and a temporal aspect. The frame used in the present study consisted of 13 sentence starters that were designed to elicit brief professional autobiographies that focussed on respondents' research activities. Data were collected electronically using Qualtrics survey software. Dialogue boxes enabled participants to write as much or as little as they liked in order to complete the sentences. The frame was designed to allow participants to reflect on their professional lives, in particular, their research activities. The frame was designed in such a way that the respondents were required to read all 13 sentence starters before being able to enter their responses. This was to ensure that they had a sense of the narrative structure of the frame prior to completing it. The instructions and sentence starters provided to the respondents were as follows:

> Following is a set of 13 statement starters. They are designed to encourage you to reflect on your research experiences. When put together, the sentences tell a brief story about your research experience. Please complete the statements. You may extend your answers by giving examples or explaining your statements. You may write as much, or as little, as you wish.

1. I have been teaching English at a tertiary institution for _____ years.
2. And am now a ____ (job title) _____.
3. In the beginning (of my teaching career at tertiary institutions) my research experiences were ___ (adjective).

4. This was because ___.
5. My present institution(s) requires teachers at my level to be involved in research to the following extent:
6. My institution(s) rewards or penalizes teachers' research activity (or lack thereof), for example ___.
7. I engage in my own research ___ (frequency)____.
8. If I were to conduct research my purpose would be to ____.
9. My attitude to my own research is that ___.
10. The biggest challenges to my participating in research are ___.
11. On the other hand, the biggest facilitators to my participating in research are ___.
12. My research (output or quality) would improve if ___.
13. Looking to the future, in the coming 12 months I expect my research output to _____.

Analysis

In line with other research utilizing narrative frames (see Barkhuizen, 2009; Hiratsuka, 2018; Swenson & Visgatis, 2011) the present study used qualitative content analysis (Corbin & Strauss, 2015; Polkinghorne, 1995; Schreier, 2013).

The use of Qualtrics facilitated the process of qualitative content analysis with reports generated by Qualtrics sorting each response into files according to its sentence starter. For example, all responses to the sentence starter:

> In the beginning of my teaching career at tertiary institutions my research experiences were ___(adjective)____.

were automatically collated into a file at the time of report generation. Each file (i.e., each group of sentences beginning with the same starter) was then coded. The narrative frame used in the study provided a pre-existing code frame, with themes linked to the sentence starters used in the frame. The coding took place in a recursive manner: it involved a preliminary reading of the files to determine initial codes, re-reading of sentences to allocate them to codes, and the occasional generation of new codes when necessary. Relationships between codes were noted in memo form at the time of coding. After coding, interrelationships between the categories were reconsidered and at times the number of categories was reduced.

Qualitative content analysis produced three main categories of findings: researcher career trajectories, research cultures, and research perspectives. All respondents gave informed consent, and all have been given pseudonyms in order to preserve their anonymity.

Findings

Respondents

The study included a diverse range of participants in terms of gender, age, location, and nationality. The gender distribution was roughly equal, 52.86% of respondents were male, 47.14% were female, and no respondents chose "other." Both Japanese (18.84%) and non-Japanese (81.16%) were represented from throughout Japan. Master's degrees were held by 72.46% of respondents, with the remainder holding terminal degrees. The majority, 93.65%, worked at universities, and their frames suggest that they worked at a range of university types, including research-intensive and more educationally focused. Also, most respondents, 74.60%, held full-time contracts. The shortest length of time a respondent had been employed in higher education was two years, and the longest was 35 years. A large percentage of respondents, 86.76%, said they had peer-reviewed publications.

Themes

The use of a qualitative content analysis framework aligns with our desire to focus on teachers' lived experiences. The analysis of the participants' responses produced three categories: (a) *research trajectories*, including their opinions as to the success of their early research activities, their present involvement in research, and their future expectations; (b) *research cultures*, for example, institutional requirements, institutional rewards and penalties, and challenges to and facilitators of research; and (c) *present perspectives*, such as purposes for conducting research and attitudes to research. The terms "research trajectories" and "present perspectives" are not found in the literature on the topic of researcher engagement; instead, they arise from our analysis of the data and encapsulate dominant themes. The most commonly occurring themes are reported below. Because the frame enabled respondents to express more than one opinion regarding any category, the references in the categories do not equal 63.

Research Trajectories

Table 1 summarises the data concerning individual researcher career trajectories, such as research production in the early years of working in higher education, reasons for the type of research engagement in the early years, the present level of research engagement, and anticipated research production in the coming year.

Table 1

Research Trajectories: Research Beginnings, Present Involvement, and Future Expectations

Themes	Details of Themes	Frequency $N = 63$
Research in the early years	No start	10
	Limited start	37
	Poor start: emotionally difficult	6
	Good start	9
Reasons for "no, unproductive, or poor" starts	Lack of skills & research training	12
	Lack of relevant work experience & knowledge of higher education institutions	12
	Position was teaching focussed	12
	Lack of time	11
	Lack of interest in researching	3
	Lack of money or resources	2
Reasons for "good" starts	Excited by topic, had a research niche	5
	Had already gained research skills from master's or doctoral program	5
	Stimulating environment	4

Themes	Details of Themes	Frequency *N* = 63
Research frequency at present	Constantly	30
	Sporadically	23
	Rarely	6
	Never	4
Anticipated future output	Rise: increase, improve, etc.	26
	Remain constant: stay the same, remain steady, etc	34
	Fall: output will decrease	3

Perhaps the most salient theme is the difficulty with research in the early stages of employment in higher education. Narratives included references to not researching at all in this career stage, using words such as "nil," and "nothing." Others wrote of limited starts, for example, "very little" or "not sufficient." The emotional toll of trying to conduct research in this career phase was another theme, with words used such as "negative," "frustrating," and "stressful." This is consistent with findings of other studies of academic careers (Boice, 2000) in general and regarding HEELT faculty members (Yuan, 2017). The main reasons given for this difficulty were poor preparedness for a career in higher education, such as lack of research skills and insufficient training, in addition to limited work experience and knowledge of higher education institutions.

Historically, prevalent roles for HEELTs excluded research. Illustrating this point, Melania (female, tenured professor, doctorate) wrote "We were considered teachers, not researchers - we couldn't even use our research funds for conferences - that was my first position in Japan in 1993 at [university] - I think the situation has changed since then." This may indicate a change over time concerning expectations for their faculty members that institutions that formerly did not emphasize research productivity for HEELTs have changed stance and now do so. This finding aligns with other research in EFL contexts (see Barkhuizen, 2021; Borg & Liu, 2013; Reyes-Cruz & Perales-Escudero, 2016) concerning transformation in the roles of HEELTS.

While there were far fewer references to exciting and successful starts,

some indicate that poor starts are by no means universal. Respondents used words such as "helpful," "exciting," and "enlightening." Having an established research niche (5 references) supported some respondents in their early years. This finding aligns with Xu's (2014) finding that a stable research interest promotes research engagement.

In respect of their present research engagement (Table 1), a dominant theme was constant research engagement (30 references). Yet, comments suggest a variety of institutional expectations that downplay the value of research are still extant. One respondent suggested that their institution does not place much value on research. Illustrating this point, Ben (male, associate professor, MA), wrote "A couple of years ago, my institution suddenly made the weight we can put on research on our self-evaluation much smaller. We were suddenly told that we are basically 'just teachers'."

Optimism concerning future research engagement was a strong theme (Table 1), with many respondents indicating an expected increase in research output (26 references). We conclude that, at least concerning the small cohort of respondents in the present study, poor starts to research careers can be overcome.

In sum, many of the respondents had trouble with research in the early stages of their careers in higher education. While these findings may not appear encouraging initially, it appears that once HEELTs are more advanced in their careers it is possible to overcome initial hurdles. This interpretation is supported by the large number whom report being constantly engaged in research at present (30 references), a marked increase in the small number of respondents who reported good starts to research (9 references). However, it should also be noted that more than half of the respondents were not constantly engaged in research at the time of the study. These findings are similar to those of other studies on HEELTs (Borg & Liu, 2013; Reyes-Cruz & Perales-Escudero, 2016; Xu, 2014).

Research Cultures

Table 2 summarises the themes concerning research cultures that shaped the research trajectories reported in Table 1. These research cultures are composed of largely institutional factors that impact the conduct of research, by encouraging, supporting, facilitating, hindering, or preventing it.

Table 2
Research Cultures: Factors that Impact the Conduct of Research

Themes	Details	Frequency *N* = 63
Institutional requirements	None	26
	Publish and present	18
	Vague or the respondent did not know	13
	Apply for government grants	9
	Publications are required only for initial appointment, for promotion or for contract renewal.	9
Research facilitators	Professional networks/relationships with colleagues	22
	Adequate research budgets	19
	Internal motivation	17
	Having enough time	5
	Enrolment in further study	4
	None	4
Research barriers	Lack of time	44
	Inadequate funds or academic resources	11
	No collaborators/isolation	8
Rewards for research production	None	14
	Financial (grants, bonuses, salary increases)	13
	Contract renewal, promotion, gain full-time employment	12
	Don't know	9

Themes	Details	Frequency $N = 63$
	Positive annual evaluations and reputation, social capital	8
	Awards (prizes)	2
Penalties for not engaging in research	Financial (less research funding, no promotion, or employment contract not extended/renewed)	8
	Indirect (through gossip)	3
	Get "spoken to" by department head	1

One noticeable theme concerning institutional requirements (Table 2) was the lack of a requirement to conduct research (26 references). This response was reported by 17 respondents who were employed part-time, for instance: "There are no requirements; research activity is left up to individual instructors" (Ryo, male, part-time lecturer, MA). These part-time respondents all noted that publications were purely required for gatekeeping purposes. This means that they are not required to do any research once they have a job.

Even amongst participants employed full-time, institutional requirements for research varied widely. Part of the explanation for this may relate to the nature of the respondents' employment contracts. For example, Peta (female, lecturer, MA) commented, "Since I am on a lecturer five-year contract, I am not required to be involved in any research per se, although I am encouraged to collaborate in developing materials."

Nonetheless, another theme was a requirement to publish and present. Some institutions were highly prescriptive in research output requirements. For example, one respondent reported that teachers at his university were required to: "Publish in [the] in-house university journal once a year and once every three years in the school's other journal or [provide] proof of publishing outside the university" (Kye, male, lecturer, MA). Earl (male, professor, doctorate) responded "All faculty members are expected and encouraged to research and publish. We don't have an individual quota but as a department (35 full-time faculty) we have to publish 50 papers a year."

These comments directly align with neoliberalist management practices that require academics' research output and citations to be monitored in numerical terms. However, whether requirements for research engagement are always clearly communicated in written form is uncertain. Kanako (female, lecturer, MA), wrote "at least one presentation at an academic conference and at least one published journal article [written] in English" was required; however, when questioned about this via email, she reported being told of this requirement verbally and in private. This seems to indicate a lack of open and readily available guidelines at Kanako's institution. This anecdote should be taken in conjunction with other comments (13 references) where respondents indicated requirements were vague or they did not know.

Some senior respondents noted that as they had risen through the hierarchy at their institutions, the requirement to publish had been reduced, either because it was only ever a requirement for initial appointment or because other work duties had to take priority. One full-time respondent observed that he was not expected to do much research because he had been promoted to professor rank and was therefore required to do a larger amount of committee work instead. Yves (male, associate professor, MA) stated "I'm already an associate professor. We wouldn't have hired me without a solid research background, but the pressure to produce any more is strictly intrinsic." The variety of responses might also reflect differences between the type of institutions where the respondents were employed. There is a striking contrast in the following two comments both written by senior faculty regarding institutional research requirements. Justin, (male, associate professor, MA) puzzled "on a volunteer basis. Although this is odd because the president requires all of us to apply for research grants. The dean of the department considers research activities as optional and not counted as working hours." Justin's comments clearly express the president's perception that all faculty members should be competing for external funds. Yet, the dean of his department appears to cling to an approach more common under the liberal model that sees HEELTs as primarily educators. In sharp contrast, another respondent wrote, "it is a research university, *all* [emphasis added] teachers are expected to be researchers" (Deanna, female, professor, doctorate). Clearly, there is a difference depending on the exact nature of the institution.

Facilitation of research was a somewhat vexed theme. Not all respondents could suggest facilitators to their research engagement; for example, one wrote, "… the biggest facilitators to my participating in research are *not too many to be honest* [emphasis added]" (Harry, male, lecturer, MA). When

respondents were able to pinpoint facilitators, the most commonly reported facilitator was social engagement (22 references), such as professional networks and relationships with colleagues, and this was closely followed by funding (see Table 2). The most reported challenge to conducting research was lack of time (44 references), closely followed by lack of resources (11 references). This finding is consistent with other research in the field that indicates time pressure is felt throughout foreign language departments (Bai & Hudson, 2011; Barkhuizen, 2009; Reyes-Cruz & Perales-Escudero, 2016).

Rewards and penalties were largely conceptualized in financial terms (Table 2), for example, grants, bonuses, salary increases, and job security or improved contractual conditions. Non-monetary rewards were also referenced, for example, positive annual evaluations, reputation, and social capital and prizes. Penalties for lack of output also appeared in the narratives (Table 2). These included direct financial impact through loss of funding for research and indirect financial impact through loss of employment due to failure to have contracts renewed (12 references). However, it appears that institutions in Japan tend to encourage their staff to conduct research by linking research output to initial engagement, contract extension, and promotion, and by subsequently monitoring output through annual evaluations.

To sum up, there are three primary findings regarding research cultures at Japanese institutions from these HEELTs. First, there is considerable variation in research requirements. Variation is dependent on both institutional idiosyncrasies and employment contracts. Second, when asked to identify research facilitators, respondents could easily identify social factors such as networks, but when asked to identify challenges to conducting research, they were more likely to identify time and resources rather than social factors such as isolation. Finally, respondents more often referenced rewards for engaging in research than penalties for not engaging in research.

Present Perspectives

Table 3 summarises the themes concerning participants' perspectives on their current research engagement. It focuses particularly on participants' attitudes to their present engagement in and purposes for conducting research.

Table 3
Participant Perspectives on Their Current Research Engagement

Themes	Details of theme	Frequency $N = 63$
Attitudes to research	Positive: Fun, satisfying, etc.	33
	Negative: Difficult, inadequate time available, etc.	13
	Not a priority compared with administration or teaching	10
	It should enhance teaching	10
	Mainly done for career advancement	4
Purposes for research	Contribute to society	21
	Improve resume, employment chances	17
	Pursue intellectual interests	16
	Improve own teaching quality	11
	Gain respect, prestige, extend one's "reach"	6
	Camaraderie	1

Pleasure in engaging in research was a strong theme present in the data, a finding that distinguishes our study from others in the field (Allison & Carey, 2007; Borg & Liu, 2013). A range of positive emotions were expressed, including joy, satisfaction, and fun (33 references). While many respondents reported a range of positive emotions, not all respondents were positive; one part-time respondent observed, "There is no point in doing it if I'm not going to be rewarded/recognized" (Burt, male, part-time lecturer, MA) and another commented "I don't have enough confidence to share my research and feel I am still roaming in the dark" (Olivia, female, lecturer, MA).

The lower priority of research in comparison to other work demands was also a theme, for example, "[I] treat it as something I do when I have time but teaching and committee work is a higher priority in my context" (Elizabeth, female, professor, MA).

Not surprisingly, there was some overlap between attitudes to research and purposes for engaging in research. A connection between research engagement and enhancing teaching quality is evident in the data, for example, "If I were to conduct research my purpose would be to contribute to teaching contexts" (Yoko, female, lecturer, doctorate). While this theme conflicts with the argument made above regarding the divorce of teaching from research in neoliberalism (Olssen & Peters, 2005), it is highly likely the result of many HEELTs' formal qualifications in education and English language teaching.

Purposes for conducting research (see Table 3) largely fell into two groups, altruistic and personal. Altruistic themes include improving teaching quality (11 references) and contributing to society (21 references). Personal purposes were evident, such as career advancement (17 references) and pursuit of intellectual interests (16 references). Rikana (female, professor, doctorate), for example, wrote that her purpose in conducting research "would be to pursue my intellectual interests and possibly to contribute to the society as a whole." Some HEELTs indicated that their attitude to research is that it is for career advancement (4 references). Interestingly, there were six references to gaining the respect of others or increasing one's prestige in the field. For example, Yves (male, associate professor, MA) commented "LOL! I have many research interests… many different purposes. But essentially… to expand my reach in the global community of educators." Yves' statement can be understood in the overall context of career building and professional roles in higher education. Taken in conjunction with references to career advancement in the attitudes category, and references to improving employment chances in the purposes category, it seems that many of the HEELTs in our sample believe research engagement to be a significant component of their careers, if not their jobs, in higher education. These findings are congruent with other studies that also find highly instrumental purposes for engaging in research among HEELTS (Bai & Hudson, 2011; Barkhuizen, 2021; Borg & Liu, 2013; Xu, 2014).

To sum up the findings, there are stark differences between HEELTs in attitudes to their research and purposes for engaging in it. Furthermore, research is not of equal importance to all HEELTs, and for some already advanced in their careers, it has a less significant role than administration. These findings highlight the gatekeeping role and instrumental function of research in this context.

Discussion

Our findings address the research question, "How do HEELTs understand their research engagement in higher education institutions in Japan, and how might these understandings illuminate the institutional forces at play in their research engagement?" in three ways: first by investigating HEELTs' views of the interrelationship of their careers with research; second, by investigating the ways institutions mold their research engagement; and third, by investigating HEELT attitudes to research along with their purposes for conducting research. We now discuss our findings and their implications, in the context of the neoliberalist management practices put forward earlier in this paper.

Globally, there has been a strong push since the 1990s for increased research production and accountability by universities as they move toward neoliberalist management norms (Olssen & Peters, 2005). The narrative frames analyzed in this article do indicate some change over time toward an increasing requirement for research; however, this is not universal. Despite the observation that Japanese authorities are increasing pressure on academics generally to be research productive, there are still some roles open to HEELTs where this is not required. Eleven respondents employed full-time said that they were not aware of any requirements concerning research production. This suggests that requirements to engage in research may not always be particularly stringent, that not all HEELTs are required to be active in research, and that more traditional views of HEELTs are still viable in some institutions. The lack of a requirement for research was also a theme in responses by all part-time HEELTs.

Our findings provide a rich picture of research behavior over the course of the professional lifespan lived out in numerous locations across Japan. We believe they can be useful to present and future EFL faculty members who are interested in being research productive, by providing them with detailed insight into this context. Nonetheless, the huge diversity of institutional expectations evidenced in the findings leads us to recommend that individual EFL academics should first explore the needs and expectations of their institution, rather than making what may well be false assumptions about a presumed "publish or perish" culture. Beyond being careful to read any written information communicated directly from their institution, we suggest that academics may find it valuable to initiate nuanced discussions about research-related expectations with department heads and other colleagues employed in their own institutions.

The apparent precarity of employment at higher education institutions, with many academics employed on short-term and strictly limited-term contracts, means that these HEELTs may also need to consider future employment at institutions with vastly different expectations from their present place of employment. The findings lead us to conclude that various strategies exist for future-proofing one's employment in the Japanese higher education sector by increasing engagement in research. Our findings show that research engagement was often hindered by a lack of time, skills, and motivation. In the present competitive environment in higher education, it seems unlikely that university managers and administrators will increase time allocation for research. Given that almost all HEELTS indicated time poverty and yet some HEELTs did carve out the necessary time for research, we look to those HEELTs for guidance: Some respondents indicated that obtaining a doctorate equipped them with the necessary research skills and confidence to be effectively research engaged. While a PhD is costly and therefore perhaps beyond the reach of many HEELTs, other respondents indicated that membership of professional organizations, and in particular membership of special interest groups in those organizations, facilitated the development of both research skills and research engagement. We, therefore, recommend that active participation in these types of organizations may assist HEELTs to build research-related capital that will serve them well in their careers.

As with teachers working outside of higher education, many respondents in the present study argued that they found meaning in their research because it directly supported their teaching. This is the same argument that has been put forth by scholars promoting teacher research among general teacher populations. Continued participation in research may not be a mandatory aspect of HEELTs' professional roles at all higher education institutions in Japan; nonetheless, it can provide meaning and job satisfaction for them.

One significant theme is the importance of social connectedness as it relates to productive research engagement. This contrasts with other studies that reported individualistic research cultures (Borg & Liu, 2013). If pleasure and camaraderie are important aspects of a research career, how can this be encouraged? In a neoliberalist managerial environment, academics are held accountable for their own economic success. While we in no way wish to suggest that universities should not provide funds for the research work of part-time and early career researchers, the clear implication is that in the absence of such support HEELTs in these categories should prioritize

self-funding memberships of professional associations, attendance at their conferences, symposium dinners, and the like, in order to facilitate friendship building, networking, and collegiality.

Conclusion

In this study, we asked the question "How do HEELTs understand their research engagement in higher education institutions in Japan, and how might these understandings illuminate the institutional forces at play in their research engagement?" In concert with other studies in the field, our findings indicate that many HEELTs understand the role of research in higher education in Japan in instrumental terms. They perceive its gatekeeping function, and a notable percentage of the present study reported negative experiences and unsupportive research cultures. In view of the neoliberalist ideology surrounding many institutions, it is somewhat surprising that many respondents did not report pressure to be engaged in research. For example, "I'm already an associate professor. We wouldn't have hired me without a solid research background, but the pressure to produce any more is strictly intrinsic." This comment does reflect that in many instances, research output continues to have primarily a gatekeeping function. It is difficult to read comments such as this and not feel somewhat disheartened. It implies that at least some institutions are not seeking to promote continued research engagement, and the loss to society if the HEELTs' intrinsic motivation does not continue is twofold. First, if academics choose not to do research, they are also choosing not to take advantage of the established pedagogical washback from researching in one's field (Barkhuizen, 2021; Borg, 2010). Second, society loses because academics that are not engaged in research are no longer contributing to the knowledge economy or to increasing humankind's knowledge in general.

A limitation of the present study concerns the sample distribution (18.84% Japanese – 81.16% non-Japanese), which would undoubtedly have impacted the results. We believe this resulted from the original nationality distribution of the special interest group in which the study was advertised. The study, therefore, presents the experiences of non-Japanese respondents more so than those of Japanese respondents.

Our findings indicate that while many HEELTs have instrumental purposes for research, others' purposes are connected to a desire to enhance their teaching. In this instance, continued participation in research activities reflects intrinsic interest, rather than a professional identity mandated by their higher education institution. This possibility requires further research.

Sarah Mason is a lecturer at Takushoku University in Tokyo, Japan. Her research interests include course feedback and identity in language teaching and research. Sarah is a doctoral candidate at Macquarie University.

Alice Chik is an associate professor in the School of Education, Macquarie University, Australia. Her research interests include language learning and multilingual literacies in digital environments.

Peter Roger is an associate professor in the Department of Linguistics at Macquarie University, Australia. His research interests include motivation, identity in language learning, healthcare communication, and interpreter-mediated interactions.

References

Allison, D., & Carey, J. (2007). What do university language teachers say about language teaching research? *TESL Canada Journal, 24*(2), 61–81. https://doi.org/10.18806/tesl.v24i2.139

Bai, L., & Hudson, P. (2011). Understanding Chinese TEFL academics' capacity for research. *Journal of Further and Higher Education, 35*(3), 391–407. https://doi.org/10.1080/0309877X.2011.569014

Barkhuizen, G. (2009). Topics, aims, and constraints in English teacher research: A Chinese case study. *TESOL Quarterly, 43*(1), 113–125. https://doi.org/10.1002/j.1545-7249.2009.tb00231.x

Barkhuizen, G. (2021). Identity dilemmas of a teacher (educator) researcher: Teacher research versus academic institutional research. *Educational Action Research*, *29*(3), 358–377. https://doi.org/10.1080/09650792.2020.1842779

Barkhuizen, G., & Wette, R. (2008). Narrative frames for investigating the experiences of language teachers. *System (Linköping)*, *36*(3), 372–387. https://doi.org/10.1016/j.system.2008.02.002

Block, D., Gray, J., & Holborow, M. (2012). *Neoliberalism and applied linguistics*. Routledge.

Boice, R. (2000). *Advice for new faculty members : Nihil nimus*. Allyn and Bacon.

Borg, S. (2007). Research engagement in English language teaching. *Teaching and Teacher Education, 23*(5), 731–747. https://doi.org/10.1016/j.tate.2006.03.012

Borg, S. (2010). Language teacher research engagement. *Language Teaching, 43*(4), 391–429. https://doi.org/10.1017/S0261444810000170

Borg, S. (2013). *Teacher research in language teaching: A critical analysis*. Cambridge University Press.

Borg, S. (2015). *Teacher cognition and language education: Research and practice*. Bloomsbury Publishing.

Borg, S., & Liu, Y. (2013). Chinese college English teachers' research engagement. *TESOL Quarterly, 47*(2), 270–299. https://doi.org/10.1002/tesq.56

Brotherhood, T., Hammond, C., & Kim, Y. (2020). Towards an actor-centered typology of internationalization: A study of junior international faculty in Japanese universities. *Higher Education, 79*(3), 497–514. https://doi.org/10.1007/s10734-019-00420-5

Corbin, J. M., & Strauss, A. L. (2015). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (4th ed.). Sage.

Hiratsuka, T. (2018). Narrative frames as a course evaluation instrument. *The Language Teacher, 42*(1), 3–7. https://doi.org/10.37546/JALTTLT42.1-1

Kubota, R. (2016). The multi/plural turn, postcolonial theory, and neoliberal multiculturalism: Complicities and implications for applied linguistics. *Applied Linguistics*, *37*(4), 474–494. https://doi.org/10.1093/applin/amu045

Olssen, M., & Peters, M. (2005). Neoliberalism, higher education and the knowledge economy: From the free market to knowledge capitalism. *Journal of Education Policy*, *20*(3), 313–345. https://doi.org/10.1080/02680930500108718

Polkinghorne, D. (1995). Narrative configuration in qualitative analysis. *International Journal of Qualitative Studies in Education*, *8*(1), 5–23. https://doi.org/10.1080/0951839950080103

Powell, W. W., & Snellman, K. (2004). The knowledge economy. *Annual Review of Sociology, 30*(1), 199–220. https://doi.org/10.1146/annurev.soc.29.010202.100037

Reyes-Cruz, M. R., Murrieta-Loyo, G., & Perales-Escudero, M. (2018). Motivacion, creencias y percepciones hacia la investigation de profesores de lenguas extranjeras en tres universidades mexicanas [Foreign language faculty research-related beliefs, perceptions, and research motivation at three Mexican universities]. *Revista PROFILE: Issues in Teachers' Professional Development, 20*(2), 35–49. https://doi.org/10.15446/profile.v20n2.66008

Reyes-Cruz, M. R., & Perales-Escudero, M. D. (2016). Research self-efficacy and research motivation in a foreign language university faculty in Mexico: Implications for educational policy. *Higher Education Research and Development, 35*(4), 800–814. https://doi.org/10.1080/07294360.2015.1137884

Schreier, M. (2013). Qualitative content analysis. In U. Flick (Ed.), *The SAGE handbook of qualitative data analysis* (pp. 170–184). Sage. https://ebookcentral-proquest-com.simsrad.net.ocs.mq.edu.au/lib/mqu/detail.action?docID=1707694

Swenson, T., & Visgatis, B. (2011). Narrative frames to assess overseas experiences. In A. Stewart (Ed.), *JALT2010 Conference Proceedings* (pp. 441–452). JALT.

Wadden, P., & Hale, C. C. (2019). The landscape of Japanese higher education: An introduction. In P. Wadden & C. C. Hale (Eds.), *Teaching English at Japanese universities: A new handbook* (pp. 3–10). Routledge.

Xu, Y. (2014). Becoming researchers: A narrative study of Chinese university EFL teachers' research practice and their professional identity construction. *Language Teaching Research, 18*(2), 1–18. https://doi.org/10.1177/1362168813505943

Yuan, R. (Eric). (2017). "This game is not easy to play": A narrative inquiry into a novice EFL teacher educator's research and publishing experiences. *Professional Development in Education, 43*(3), 474–491. https://doi.org/10.1080/19415257.2016.1182936

Appendix

Questions Collecting Demographic Information

Survey

Q: Are you currently teaching English as a foreign, or second language at a tertiary institution in Japan? Y/N

Demographics

1. Gender: Male Female Other

2. Age: 20-24 25-34 35-44 45-54
55-64 65-74 75-84

3. Nationality
   - Japanese (including dual citizenship holders)
   - Other

4. Prefecture residing in:

5. Your main employer:
   - University
   - Technical institute
   - Vocational School
   - High School
   - Junior High School
   - Primary (Elementary) School
   - Preschool
   - Conversation school
   - Other

6. What is your highest level of education?
   - Undergraduate degree
   - Master's degree (coursework only)
   - Master's degree (coursework and dissertation)
   - Doctorate

7. Are you currently studying for a master's degree or doctorate? Yes/No

8. Are you planning to enrol in a master's degree or doctorate within the next 12 months? Yes/No

9. Do you have any publications in peer reviewed journals? Yes/No

Research Forum

The Relationship between Japanese EFL Learners' Perceived Fluency and Temporal Speech Measures in a Read-Aloud Task

Aki Tsunemoto
Concordia University

Pakize Uludag
Concordia University

Kim McDonough
Concordia University

Talia Isaacs
University College London

This study examined the relationship between holistic rater judgments of second language (L2) speech fluency (i.e., perceived fluency) and temporal measures of fluency (i.e., utterance fluency) in a read-aloud task. 63 L2 English Japanese secondary school students were audio-recorded while carrying out a 69-word read-aloud task. 11 L2 English-speaking instructors rated the speech for perceived fluency, and the speech samples were analyzed for utterance fluency. The linear regression model revealed that articulation rate and clause-internal pauses significantly predicted

https://doi.org/10.37546/JALTJJ45.1-5

JALT Journal, Vol. 45, No. 1, May 2023

perceived fluency. Findings are discussed in relation to the use of read-aloud tasks for the teaching and assessment of L2 speech fluency.

本研究では、音読タスクにおける第二言語音声の流暢性に関する総合的評価と、スピーチの言語的特徴の関係性を調査した。高校生の日本人英語学習者63名が、69語の音読タスクを実施する様子を録音した。11人の言語指導経験のある第二言語英語話者が、スピーチの流暢性について評価した。更に、スピーチを言語的特徴について分析した。重回帰モデル解析の結果、発声速度と節内のポーズが流暢性の重要な予測変数として算出された。これらの成果に基づき、音読タスクを利用した第二言語における流暢性の指導や評価について教育的な示唆を行う。

Keywords: English as a foreign language; fluency; read-aloud task; second language speaking; speech perception

Whereas fluency in a broad sense is often equated with general oral proficiency, fluency in a narrow sense refers to the temporal fluidity of speech (Lennon, 1990), specifically whether it is smooth and rapid (De Jong, 2018). One goal of fluency research has been to understand the relationship between *utterance fluency* (i.e., speech features), and *perceived fluency*, which captures raters' impressions of utterance fluency (Segalowitz, 2010). To gain insight into this relationship, utterance fluency has been measured in terms of speed fluency (e.g., speech rate), breakdown fluency (e.g., duration and frequencies of pauses) and repair fluency (e.g., frequency of self-corrections and repetitions) (Tavakoli & Skehan, 2005) while perceived fluency has been assessed through holistic rater judgments. Prior studies of fluency during spontaneous speech reported a positive relationship between perceived fluency and speed fluency measured as speech rate (e.g., Magne et al., 2019) and mean length of run (MLR: e.g., Kormos & Dénes, 2004; Trofimovich, et al., 2017). In contrast, perceived fluency has been negatively associated with breakdown fluency measured as the frequency and durations of silent pauses (e.g., Rossiter, 2009), pauses within clauses (e.g., De Jong & Bosker, 2013; Kahng, 2018; Suzuki & Kormos, 2020), and pauses between clauses (Saito et al., 2018). Finally, perceived fluency has shown both positive (Magne et al., 2019) and negative (Kormos & Dénes, 2004) relationships with repair fluency.

Although the relationship between perceived and utterance fluency has been widely examined in spontaneous speech, less is known about their relationship during read-aloud tasks, which are commonly used for both English proficiency testing and pedagogical activities. Several high-stakes English proficiency tests use read-aloud tasks, often combined with automated scoring, as part of their speaking assessment (e.g., Duolingo, EIKEN, GTEC, Pearson Test of English Academic [PTEA]), including new tests

developed in response to the COVID-19 pandemic (e.g., TOEFL Essentials, Isbell & Kremmel, 2020). In Japan, English learners may take such tests for admission to foreign universities or for immigration purposes. Among them, EIKEN, which includes a read-aloud task for most grade levels, is taken by three million people each year as a gatekeeping measure to demonstrate English proficiency for post-secondary education and employment in Japan (EIKEN, n.d.). Furthermore, in L2 classrooms, read-aloud tasks have been included in diagnostic pronunciation assessment to identify learner needs and create individualized instruction (Celce-Murcia et al., 2010). In Japan specifically, rather than using extemporaneous speech tasks, instructors often implement controlled tasks (e.g., reading aloud from textbooks) for teaching (Uchida & Sugimoto, 2018) and evaluating their students' speaking performance (for review, see Koizumi, 2022).

Unlike spontaneous speech, read-aloud tasks do not require speakers to conceptualize message content. Instead, they need to parse the textual information, encode phonological information, and execute the planned phonetic information into sounds using physiological mechanisms. Although read-aloud tasks require this complex processing, they do not require speakers to pre-plan content, retrieve words, or build grammatical structures as in spontaneous speech tasks. As a result, a speaker may produce more regulated speech patterns (Laan, 1997) and speak faster with fewer hesitations (Trofimovich, et al., 2017) during read-aloud tasks than spontaneous speech. The lower variability in speaker performance is conducive for machine scoring, making the read-aloud task attractive as a time-efficient, reliable, and inexpensive test item that can be scored automatically (Isaacs, 2018). Nevertheless, in languages like English with poor sound-symbol correspondence, read-aloud tasks may still pose challenges for speakers, such as mispronouncing words that have irregular written forms or hesitating before unfamiliar words (Hayes-Harb et al., 2010), and these challenges may influence rater perceptions of their fluency.

In light of the role of read-aloud tasks in L2 assessment and classroom practices in English L2 settings, it is important to investigate speech characteristics that are perceptually salient to L2 English speakers. The few studies that included read-aloud tasks with L2 Dutch and L2 French speakers found that perceived fluency was positively associated with speed and repair fluency but negatively related to breakdown fluency (Cucchiarini et al., 2002; Trofimovich, et al., 2017). However, both studies elicited evaluations of perceived fluency from first language (L1) speakers of the target language. Prior studies of perceived fluency during spontaneous speech found that both L1 and L2 English raters were influenced by speed and clause-internal pauses,

but only L1 raters were sensitive to clause-external pausing (e.g., Magne et al., 2019). Little is known, however, about whether these utterance fluency measures are equally important for L2 English speakers when assessing L2 fluency through a read-aloud task. Due to globalization, most English speakers are now L2 speakers (Pennycook, 2020) and many work as instructors and language test examiners (Carey et al., 2011), which highlights the need for further research to elicit their perceptions of fluency. Therefore, the current study examines the relationship between L2 English-speaking instructors' perceptions of fluency and temporal measures of Japanese English as a foreign language (EFL) students' read-aloud task performance. The research question was as follows:

RQ. What temporal measures of speech fluency (i.e., utterance fluency) are associated with L2 English-speaking teachers' holistic fluency ratings (i.e., perceived fluency) during a read-aloud task?

Method

L2 Speakers

As part of a larger study, L2 speech samples were elicited from 63 secondary school students in Japan (45 males, 18 females, M_{age} = 16.4, *SD* = 0.6). All students and parents were L1 Japanese speakers except for one Japanese-Korean bilingual student. The students began studying English around the age of 10.5 years (*SD* = 3.1) and except for the bilingual student, they had no experience living in English-speaking countries longer than a month. All but eight students self-reported their most recent EIKEN Grades (*range* = Grade 1–4), 80% of whom reported achieving Grade 2, Pre-2, or 3. Their English classes primarily targeted reading and writing skills, and speaking activities usually involved reading words and sentence aloud from a textbook, occasional paired or group discussions, and bi-weekly sessions with an assistant language teacher. Some students voluntarily participated in after-school English conversation groups.

Task and Speech Recording

During an individual session with the first researcher (15 minutes), the students completed a read-aloud task based on a passage from the Speech Accent Archive (Weinberger, 2015; see Appendix). The 69-word passage was selected because it contained all possible English sounds for eliciting the students' phonological encoding skills (Cucchiarini et al., 2002). Each

student was given the passage and were asked to read it silently within one minute. After having the opportunity to ask about the meaning or pronunciation of any unfamiliar words, each student read the passage aloud while being audio-recorded. The audio-recordings, which ranged in length from 22 to 47 seconds, were trimmed by removing initial pauses and hesitations and normalized for peak intensity. The recordings were organized into three lists with different orders to limit the possibility of ordering effects.

Raters and Rating Procedure

Reflecting our focus on L2 English-speaking raters, we purposefully recruited L2 English speakers who had teaching experience. To ensure consistency in their familiarity with the Japanese language (Carey et al., 2011), we recruited raters who had never lived in Japan and did not speak Japanese. Through convenience sampling, 11 L2 English raters (10 females, 1 male) with experience teaching English to L2 learners (M = 5.8 years, SD = 4.0) were recruited. They were adults (M = 31.4 years, SD = 6.5) enrolled in or recent graduates of Education programs at an English-medium Canadian university. As degree seeking students, they had met the university's minimum English language requirement for admission without additional language instruction, which was a TOEFL iBT score of 90 (or equivalent). On a background questionnaire, they reported varied L1 backgrounds, including Chinese, Dutch, Farsi, Polish, Portuguese, Russian, and Vietnamese. They all reported having normal hearing, and nine reported having previously taken a phonology course. They estimated the percentage of time that they used English in their daily life on a scale of 0 to 100% for both speaking (M = 69.1%, SD = 24.3) and listening (M = 74.6%, SD = 21.2). When asked to self-report familiarity with L2 accented English on a percentage scale (Tsunemoto et al., 2021; 0 = *not at all,* 100 = *very familiar*), the raters indicated that they were very familiar with L2-accented English (M = 77.8%, SD = 17.2), but not very familiar with Japanese accents specifically (M = 27.3%, SD = 26.1). None of the raters had previously lived in Japan and they reported spending little time in their daily lives' interacting with Japanese speakers (M = 9.1%, SD = 16.1) when the study was carried out.

The raters scheduled individual rating sessions (60 min) with the first or second researcher held in a quiet room on a university campus in Canada. All 11 raters evaluated the entire 63 speech samples on a computer connected to a headset using 9-point Likert-type fluency scales (1= *not fluent at all*, 9= *very fluent*) in accordance with L2 speech fluency research conventions (e.g., Suzuki & Kormos, 2020). In line with previous studies that have revealed

highly consistent fluency ratings among raters (e.g., Trofimovich, et al., 2017), raters were asked to judge how smooth the oral delivery was while focusing on temporal features (speech rate, fillers, pauses) in the speech (e.g., Kahng, 2018). After completing three practice ratings, they had opportunities to ask about the speech samples or rating scale. They were instructed to listen to an entire speech sample before providing a fluency rating. Raters were randomly assigned to one of three presentation orders to avoid possible ordering effects. The internal consistency of the raters' perceived fluency ratings was assessed by Cronbach's alpha, which was .91. Interrater reliability was assessed through two-way random, agreement, average-measure intraclass correlation coefficients. The obtained value was .88, which revealed acceptable rater agreement (Field, 2018; Kahng, 2018). As the consistency exceeded the threshold values of .70–.80 (Larson-Hall, 2010), fluency ratings were averaged to derive single mean scores for each speech sample.

Speech Analysis

The speech samples were analyzed for six temporal measures of speech that reflect speed fluency, breakdown fluency, and repair fluency. Although prior research has used several utterance fluency measures (e.g., Tavakoli, et al., 2020), we selected measures from previous studies with EFL Japanese speakers (e.g., Saito, et al., 2018) or read-aloud tasks (e.g., Cucchiarini et al., 2002). For speed, articulation rate was calculated as total syllables divided by total phonation time (subtracting the total silent pause duration from the total speech duration) (Prefontaine et al., 2016). Four pause measures were used to assess breakdown fluency (MLR, clause-external, clause-internal, and filled pauses). MLR (total syllables/utterances produced between silent pauses) has been examined as speed measure (Prefontaine, et al., 2016), but we considered the variable as breakdown measure as it incorporates pauses and may represent a speaker's hesitation (Towell et al., 1996). As for pauses, any silences longer than 200ms were operationalized as pauses. A shorter duration than De Jong and Bosker's (2013) recommended cut-off (250ms) was used because read-aloud tasks require shorter periods to produce speech as compared to spontaneous speech (e.g., Cucchiarini, et al., 2002). Silent pauses were manually coded using Praat (Boersma & Weenink, 2017) with the assistance of automated silence detection. Pauses were categorized as either clause-external or clause-internal to examine relative contribution of pause location to perceived fluency ratings (e.g., Bosker et al., 2013; Kahng, 2018). Filled pause frequency was obtained as total number of dysfluencies (e.g., uh and um) divided by total phonation time (Bosker et

al., 2013). Repair fluency was operationalized in terms of the repair ratio, which is the total number of dysfluencies (e.g., self-corrections and repetitions) divided by the total number of syllables in a passage from the Speech Accent Archive (Weinberger, 2015) to obtain a standardized measure that are comparable across speakers. A subset of the data (25%) was coded by the first researcher and an independent rater. Two-way mixed, agreement, average-measure intraclass correlation coefficients revealed high agreement values for clause-external pause frequency (.97), clause-internal pause frequency (.92), filled pauses (1.00) and total dysfluencies (.88). Having established coding reliability, the remaining speech samples were coded by the independent rater.

Results

The descriptive statistics for the perceived fluency ratings and utterance fluency measures are provided in Table 1. The raters provided a wide range of L2 fluency ratings (3.2–7.8 on a 9-point scale), with a mean score slightly above the scale midpoint (M = 5.3). Overall, L2 speakers produced all types of utterance fluency measures, but filled pauses and repairs occurred less frequently.

Table 1
Descriptive Statistics for Perceived Fluency and Utterance Fluency

Variables			*M*	*SD*	*Min*	*Max*
Perceived Fluency		Raters' ratings	5.30	1.08	3.18	7.82
Utterance Fluency	Speed	Articulation rate	3.21	0.38	2.17	4.52
	Breakdown	Mean length of run	5.75	2.07	3.00	13.80
		Clause-external pause frequency	0.29	0.08	0.08	0.47
		Clause-internal pause frequency	0.22	0.14	0.01	0.66
		Filled pause frequency	0.03	0.06	0.01	0.25
	Repair	Repair ratio	0.04	0.03	0.01	0.15

Half of the utterance fluency measures had skewness and kurtosis indices larger than ±2 and examination of the histograms suggested that the data were not normally distributed (Field, 2018). Therefore, a nonparametric Spearman's rank-order correlations were obtained to determine the relationship between utterance fluency and perceived fluency (see Table 2).

Table 2

Correlations Between Perceived Fluency Ratings and Utterance Fluency Measures

	1	2	3	4	5	6
Perceived Fluency	.71**	.67**	.26*	–.71**	–.21	–.23
1. Articulation rate	-	.46**	.28*	–.41**	–.16	–.17
2. Mean length of run		-	–.22	–.87**	–.26*	–.25*
3. Clause-external pause frequency			-	–.15	–.11	–.24
4. Clause-internal pause frequency				-	.31*	.37**
5. Filled pause frequency					-	.54**
6. Repair ratio						-

Note. $^{*}p < .05$, $^{**}p < .01$.

Based on the correlation coefficients, MLR was dropped from further analysis because it was strongly correlated with clause-internal pauses[1]. The three remaining variables that reached the benchmark for a small correlation coefficient of ±.25 (Plonsky & Oswald, 2014) were selected for inclusion in a hierarchical multiple regression model: articulation rate, clause-external pauses, and clause-internal pauses. Regarding assumptions and model fit, tests of multicollinearity showed that the model all tolerance values were above .20, and no VIF values were above 10 (1.00 to 1.24). The Durbin-Watson statistic indicated good model fit (1.84). The normality of residuals was determined by (a) visual inspection of histogram, scatterplots, and P-P plots, (b) fewer than 5% of cases with standardized residuals greater ±2, and (c) Cook's distance and DfBeta values were less than 1 (Field, 2018).

Because previous research has identified the importance of speed fluency, articulation rate was entered first followed by the two breakdown fluency

measures. As shown in Table 3, the first model with articulation rate was significant, but the second model with clause-external pauses and clause-internal pauses led to a significant *F* change and higher R^2 value.

Table 3

Summary of Hierarchical Regression Models for Raters' Ratings

Blocks	*R*	R^2	ΔR^2	ΔF	*p*
1. Articulation rate	.66	.43	.42	46.66	.001
2. Clause-external pauses & clause-internal pauses	.85	.71	.70	28.85	.001

Both articulation rate and clause-internal pauses were significant predictors of L2 raters' perceived fluency in the second model and they explained a combined 71% of the variance, $R^2 = .71$, $F(3, 59) = 48.99$, $p < .001$. (see Table 4).

Table 4

Summary of Predictor Variables for Regression Model with Blocks 1 and 2

Predictors	*B*	*SE B*	B	95%CI		*t*	*p*
Articulation rate	1.87	.27	.66	1.32	2.42	6.83	.001
Clause-external pause	1.15	.99	.09	–.82	3.12	1.17	.248
Clause-internal pause	–4.13	.55	–.55	–5.24	–3.02	–7.46	.001
Constant	1.71	.72		.26	3.15	2.36	.021

Discussion

This study determined which temporal measures of utterance fluency are associated with L2 English speakers' holistic ratings of students' perceived fluency during a read-aloud task. The positive relationship between articulation rate and perceived fluency is in line with previous read-aloud task studies that demonstrated a positive link between articulation rate (i.e., mean syllables per second excluding pauses) and L2 Dutch fluency ratings (Cucchiarini et al., 2002) or between MLR and L2 French fluency ratings

(Trofimovich et al., 2017). Put simply, these EFL speakers were perceived to be more fluent if they produced more syllables per second when reading aloud. Additionally, perceived L2 fluency was negatively associated with clause-internal pauses. Although prior read-aloud research identified a negative association between perceived fluency and the duration and frequency of silent pauses (Cucchiarini et al., 2002), the current findings indicate that only clause-internal pauses predicted perceived fluency. When reading aloud, pausing at clause boundaries may have occurred when these EFL speakers were organizing words into meaningful chunks, which did not influence these raters' perceptions. However, when they paused within clauses, such as when hesitating to pronounce unfamiliar words, they were perceived to be less fluent.

An example of clause-internal pauses is provided in the excerpt below ([*] represents a 200ms or longer clause-internal pause). This student received a low fluency rating (3.18 on a 9-point scale) and her speech contained numerous clause-internal pauses. Even though the student had chances to check the pronunciation of the unfamiliar words before reading aloud, clause-internal pauses seem to occur before unfamiliar words (e.g., slabs, plastic, scoop). There were pauses before more familiar words (e.g., big, bags, train), which suggests that the student did not put words into chunks, such as noun phrases (e.g., a big toy frog, three red bags) or prepositional phrases (e.g., at the train station).

> S56: Please [*] call Stella. Ask her to bring [*] these [*] things with her from the [*] store. Six [*] spoons of fresh snow [*] peas, five thi-[*]-ck [*] slabs [*] of blue cheese, and [*] maybe a snack for her brother Bob. We also need [*] a small [*] plastic snake and [*] a [*] big [*] toy frog for [*] the kids. She can s-[*]-coop [*] these things into three red [*] bags, and we will go meet her [*] Wednesday at [*] the [*] train station.

Finally, in contrast to speed and breakdown fluency measures, repair fluency occurred relatively infrequently and did not predict perceived fluency, which is in line with previous studies that demonstrated small negative correlations between repair fluency and perceived fluency in L2 Dutch (r = –.15) (Cucchiarini et al., 2002) and L2 French (r = –.24) (Trofimovich et al., 2017).

The current study raises some potential implications for L2 instruction and assessment. Instructors may help students increase their articulation rate and decrease their clause-internal pauses by having them read the

same text aloud repeatedly (Yoshimura & MacWhinney, 2007). For instance, instructors may include target formulaic sequences (Wood, 2009) in a text and then ask students to read it aloud repeatedly with increased time pressure over cycles, which may result in better retention of word chunks (Durrant & Schmitt, 2010). In addition, when using read-aloud or other scripted tasks, instructors can help students recognize where to pause and which words form a unit by using typographical enhancement, such as punctuation markers. However, the effect of such pedagogical interventions should be empirically examined in future research. When it comes to the use of read-aloud task in L2 fluency assessment, the current findings suggest that human raters (e.g., EIKEN) may be susceptible to the location of pauses (clause-internal vs. clause-external pauses), which should be reflected in the automated machine scoring in language tests (e.g., PTEA).

Although this study highlights how pause locations and articulation speed relate to perceived fluency during a read-aloud task, several factors may limit its generalizability. First, to minimize the influence of listeners' individual characteristics, we purposefully recruited L2 English-speaking raters who had L2 teaching experience but had little exposure to the Japanese language. Nonetheless, the raters had variation in their familiarity with Japanese-accented English (M = 27.3%, SD = 26.1). Although Kahng (2018) did not find any relationships between listeners' accent familiarity and L1 Korean speakers' fluency ratings, future research should explore if such relationships exist when different L1–L2 combinations are utilized (e.g., listeners with varying degrees of familiarity with Japanese accents evaluate L2 English fluency). In addition, fluency in this study was operationalized by having the raters judge how smoothly the speech was delivered while focusing on temporal speech features. Although inter-rater reliability among raters was high, it would be important to qualitatively investigate which temporal measures of speech the raters focused on when evaluating fluency in a read-aloud task to triangulate the current findings. Finally, although the use of read-aloud tasks was ecologically valid for the Japanese EFL setting where there is little L2 exposure outside the classroom (Uchida & Sugimoto, 2018), future investigations of speech fluency should explore the relationship between utterance and perceived fluency for other tasks and in other foreign and second language contexts.

Notes

1. As recommended by Suzuki et al. (2021), we used articulation rate rather than MLR because the latter reflects multiple dimensions of utterance fluency.

Acknowledgements

We gratefully acknowledge the assistance of Shungo Suzuki in data coding and analysis. This study was funded by the Canada Research Chairs program (950-231218) awarded to Kim McDonough.

Aki Tsunemoto received her PhD in Education from Concordia University. Her research interests include second language speech assessment and individual differences in speech perception.

Pakize Uludag is an assistant professor in technical communication in the Centre for Engineering in Society at Concordia University. Her interests include academic writing, language assessment and Corpus Linguistics.

Kim McDonough is a Professor of Applied Linguistics at Concordia University. Her research examines visual cues during task-based interaction, reverse linguistic stereotyping, and writing development.

Talia Isaacs is Associate Professor of Applied Linguistics and TESOL at the IOE, UCL's Faculty of Education and Society. Her research interests include speaking and assessment.

References

Bosker, H. R., Pinget, A. F., Quené, H., Sanders, T., & de Jong, N. H. (2013). What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing*, *30*(2), 159–175. https://doi.org/10.1177/0265532212455394

Boersma, D., & Weenink, P. (2017). *Praat: Doing phonetics by computer* (Version 6.0.40). http://www.praat.org

Carey, M. D., Mannell, R. H., & Dunn, P. K. (2011). Does a rater's familiarity with a candidate's pronunciation affect the rating in oral proficiency interviews? *Language Testing, 28*(2), 201–219. http://doi.org/10.1177/0265532210393704

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching pronunciation hardback with audio CDs: A Course book and reference guide*. Cambridge University Press.

Cucchiarini, C., Strik, H., & Boves, L. (2002). Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. *The Journal of the Acoustical Society of America*, *111*(6), 2862–2873. https://doi.org/10.1121/1.1471894

De Jong, N. H. (2018). Fluency in second language testing: insights from different disciplines. *Language Assessment Quarterly*, *15*(3), 237–254. https://doi.org/10.1080/15434303.2018.1477780

De Jong, N. H., & Bosker, H. R. (2013). Choosing a threshold for silent pauses to measure second language fluency. In R. Eklund (Ed.), *Proceedings of the 6th Workshop on Disfluency in Spontaneous Speech* (pp. 17–20). Royal Institute of Technology (KTH).

Durrant, P., & Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research, 26*(2), 163–188. https://doi.org/10.1177/0267658309349431

Eiken. (n.d.). *Juken no jōkyō* [Examinee statistics]. Retrieved February 22, 2022, from https://www.eiken.or.jp/eiken/merit/situation/

Field, A. (2018). *Discovering statistics using IBM SPSS statistics*. Sage.

Hayes-Harb, R., Nicol, J., & Barker, J. (2010). Learning the phonological forms of new words: Effects of orthographic and auditory input. *Language and Speech*, *53*(3), 367–381. https://doi.org/10.1177/0023830910371460

Isaacs, T. (2018). Fully automated speaking assessment: Changes to proficiency testing and the role of pronunciation. In O. Kang, R. I. Thomson, & J. Murphy (Eds.), *The Routledge handbook of contemporary English pronunciation* (pp. 570–584). Routledge.

Isbell, D. R., & Kremmel, B. (2020). Test review: Current options in at-home language proficiency tests for making high-stakes decisions. *Language Testing, 37*(4), 600–619. https://doi.org/10.1177/0265532220943483

Kahng, J. (2018). The effect of pause location on perceived fluency. *Applied Psycholinguistics*, *39*(3), 569–591. https://doi.org/10.1017/S0142716417000534

Koizumi, R. (2022). L2 speaking assessment in secondary school classrooms in Japan. *Language Assessment Quarterly.* Advance online publication. http://doi.org/10.1080/15434303.2021.2023542

Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, *32*(2), 145–164. https://doi.org/10.1016/j.system.2004.01.001

Laan, G. P. M. (1997). The contribution of intonation, segmental durations, and spectral features to the perception of a spontaneous and a read speaking style. *Speech Communication*, *22*(1), 43–65. https://doi.org/10.1016/S0167-6393(97)00012-5

Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. Routledge.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning, 40*(3), 387–417. http://doi.org/10.1111/j.1467-1770.1990.tb00669.x

Magne, V., Suzuki, S., Suzukida, Y., Ilkan, M., Tran, M., & Saito, K. (2019). Exploring the dynamic nature of second language listeners' perceived fluency: A mixed-methods approach. *TESOL Quarterly, 53*(4), 1139–1150. https://doi.org/10.1002/tesq.528

Pennycook, A. (2020). The future of Englishes: One, many or none? In A. Kirkpatrick (Ed.), *The Routledge handbook of world Englishes* (pp. 679–692). Routledge.

Plonsky, L., & Oswald, F. L. (2014), How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878–912. https://doi.org/10.1111/lang.12079

Prefontaine, Y., Kormos, J., & Johnson, D. E. (2016). How do utterance measures predict raters' perceptions of fluency in French as a second language? *Language Testing*, *33*(1), 53–73. https://doi.org/10.1177/0265532215579530

Rossiter, M. J. (2009). Perceptions of L2 fluency by native and non-native speakers of English. *The Canadian Modern Language Review*, *65*(3), 395–412. https://doi.org/10.3138/cmlr.65.3.395

Saito, K., Ilkan, M., Magne, V., Tran, M., & Suzuki, S. (2018). Acoustic characteristics and learner profiles of low-, mid- and high-level second language fluency. *Applied Psycholinguistics*, *39*(3), 593–617. https://doi.org/10.1017/S0142716417000571

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. Routledge.

Suzuki, S., & Kormos, J. (2020). Linguistic dimensions of comprehensibility and perceived fluency: An investigation of complexity, accuracy, and fluency in second language argumentative speech. *Studies in Second Language Acquisition, 42*(1), 143–167. http://doi.org/10.1017/S0272263119000421

Suzuki, S., Kormos, J., & Uchihara, T. (2021). The relationship between utterance and perceived fluency: A meta-analysis of correlational studies. *The Modern Language Journal, 105*(2), 435-463. https://doi.org/10.1111/modl.12706

Tavakoli, P., Nakatsuhara, F., & Hunter, A. M. (2020). Aspects of fluency across assessed levels of speaking proficiency. *The Modern Language Journal, 104*(1), 169–191. https://doi.org/10.1111/modl.12620

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239–276). John Benjamins.

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics*, *17*(1), 84–119. https://doi.org/10.1093/applin/17.1.84

Trofimovich, P., Kennedy, S., & Blanchet, J. (2017). Development of second language French oral skills in an instructed setting: A focus on speech ratings. *Canadian Journal of Applied Linguistics*, *20*(2), 32–50. https://doi.org/10.7202/1042675ar

Tsunemoto, A., Lindberg, R., Trofimovich, P., & McDonough, K. (2021). Visual cues and rater perceptions of second language comprehensibility, accentedness, and fluency. *Studies in Second Language Acquisition.* Advance online publication. http://doi.org/10.1017/S0272263121000425

Uchida, Y., & Sugimoto, J. (2018). A survey of pronunciation instruction by Japanese teachers of English: Phonetic knowledge and teaching practice. *Journal of the Tokyo University of Marine Science and Technology*, *14*, 65–75. https://ci.nii.ac.jp/naid/120006402286/

Weinberger, S. (2015). *Speech Accent Archive*. Retrieved from http://accent.gmu.edu

Wood, D. (2009). Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. *Canadian Journal of Applied Linguistics*, *12*(1), 39–57. https://journals.lib.unb.ca/index.php/CJAL/article/view/19898

Yoshimura, Y., & MacWhinney, B. (2007). The effect of oral repetition on L2 speech fluency: An experimental tool and language tutor. *SLaTE-2007*, 25–28. https://psyling.talkbank.org/years/2007/fluency.pdf

Appendix

Read-Aloud Passage from Speech Accent Archive (Weinberger, 2015)

Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station.

Expositions

Reflective Practice for TESOL Teachers: "What, Why, When and How"

Thomas S. C. Farrell
Brock University

In this invited paper for *Expositions* in the *JALT Journal* I will address all aspects of reflective practice for TESOL teachers so that readers can consider implementing it in their daily practice in Japan. The paper outlines and discusses what reflective practice is, why it is important, when, and how language teachers can do it. More specifically I outline and discuss two different frameworks I developed for teachers wishing to reflect on their practice that I developed over the past 30 years. The first is an early framework I developed has five interrelated components and is useful for groups of teachers coming together to reflect on their practice. The second more recent framework also has five interrelated stages and suitable for individual teachers as well as groups when wishing to reflect on their practice. I believe that both frameworks may be useful for teachers to consider when wishing to engage in reflective practice in Japan.

Keywords: language teachers; reflective practice; TESOL

What Is Reflective Practice?

I remember the excitement and fear I felt the first day I walked into a classroom in Dublin, Ireland as a trainee "teacher" (actually, I was teaching for a year for 2 hours a day as part of my teacher qualification diploma). I remember the room and can still to this day nearly 40 years ago, see all those faces looking at me as I said "good morning" to them all. Then I also

https://doi.org/10.37546/JALTJJ45.1-6

remember that it suddenly hit me that I did not know what to do at that moment after entering the room; I remember wondering for example, 'do I stand up or sit down?' Do I ask them to open their books, and/or write on the board (yes, we had chalkboards in those days), and many more issues related to how I would conduct the class. I realize that this may seem trivial to most seasoned teachers, but to a neophyte like myself, those opening moments on my first day were some of the most frightening experiences of my teaching career. To be sure these dilemmas were critical incidents, some of many I was to experience that were not only going to shape me as a teacher but also as a reflective practitioner because as I learned after, experiences by themselves are of no use unless we engage in "reflection" so that we can learn from them. But what do I mean by the term "reflection" and is it the same as "reflective practice"?

I would say that "reflection" and "reflective practice" are a bit different, because reflection in its everyday connotation can be considered fleeting or reflection in passing based on our hunches, or intuition like the reflection I mentioned in the paragraph above about my first day as a teacher in Ireland all those years ago. It is a good start and perhaps one in which many teachers experience, however, we cannot be sure of what has really occurred beyond our hunches or intuition because the events have come and gone. So although engaging in some kind of reflection may be a good beginning when considering what we do as teachers, it is not enough to help us really see what is happening in our classes and lessons. That is why I use the term "reflective practice" because it means much more that thinking about what has happened in our lesson as we are going home on the subway or bus after a class. Engaging in reflective practice is a more systematic analysis of gathering evidence about what has actually happened in your lessons and also examining who you are as a person, what you do in the classroom, why you do it, and what the result is. It includes not only examining our teaching plans before class, our teaching actions during the class, and what we think we achieved in each class after the lesson, but also who we are as a human being, because I believe the person you are cannot be separated from the teacher you are and the act of teaching (Farrell, 2022). In other words, you bring *all of you* (your past and present) into each class you teach—for more on this see below under *philosophy* in the second framework for reflective practice I present. I provide more details on this evidence–based approach to reflective practice in the sections below.

Why Is Reflective Practice Important?

Teachers may ask why they should engage in "reflective practice" I mentioned above when they say that they always do so after teaching and mention to other teachers in the staff room that they had "a good/bad class!" or that their "students were not very responsive today!" In other words, most teachers think they already reflect already. While I agree most teachers do "reflect" in such a manner as we are not robots and we are happy after an activity or a class if we perceive these to have gone well, we can also be overly depressed or angry if we perceive them to have gone badly and then we engage in "beating ourselves up" too much. The operative word here is "perception" or what we think went well or not so well in our lessons. Some teachers base such perceptions on the way the students respond (e.g., yawning) or do not respond during class (e.g., sitting in silence). They may consider this as "a critical event" for them; however, that yawn may have nothing to do with the class or teaching and everything to do with that student's lack of sleep or an illness. So, teachers need to know why classes go "well" and some other classes do not go so "well" and how they define what this "well" means. How do you *know* it went well or not so well? So how do you collect this evidence?

Teachers can collect evidence about what they do through recordings of what actually happens in classroom lessons rather than what we *think* happens. As Walsh (2015) notes, we can only get a real understanding of the complexities of interaction when we have a precise representation of what is really occurring by recording the communications and a record of this recording in the form of a written classroom transcript. This is mostly because we all have selective memories and these are not *real evidence* of what has occurred. We can collect this type of evidence by placing an audio recorder or video recorder in our classroom. Once the classroom communication data has been collected, the teacher then needs to transcribe the recording; this can be the most painful part of the whole process because it can take a long time to transcribe a one–hour class. It may not be necessary to transcribe the entire recording; teachers can decide what aspect of the classroom communications they are interested in knowing more about. In his excellent book, Fanselow (1987) suggested that transcriptions be made at certain intervals or at special events that the teacher wants to investigate. For example, teachers may only be interested in reflecting on the impact of their verbal instructions in their classes, so all they need to do is listen to and transcribe those parts of the tape that show the teacher giving instructions and then the turns immediately after this (for about five minutes) to see what impact these have on instruction.

Teachers can also collect evidence by writing about their practice because writing has its own built–in reflective mechanism; the process entails that writers must stop to think and organize their thoughts before writing (either with a pen or computer) and then decide on what to write. After this they can 'see' (literally) their thoughts and reflect on these for self–understanding. This I call reflective writing and I use it all the time to help me with my own reflections (such as writing this article). For teachers, such reflective writing can include written accounts of teachers' thoughts, classroom observations, assumptions, beliefs, attitudes, and experiences about their practice both inside and outside the classroom (Farrell, 2013a). By writing regularly teachers are able to identify and address issues critical to their practice within their teaching contexts, and as a result provide more learning opportunities for their students. Teachers can use this evidence based on concrete evidence systematically collected over a period of time to make more informed decisions about teaching rather than relying on hunches or the like. As such, teachers will need to get solid data about what is really happening in their classroom rather than what they think is happening. This brings us to the next important question related to engaging in reflective practice, how do I do it?

When And How Do I Reflect?

To answer the "when" question about reflective practice, there is no correct answer as teachers can reflect at any time during the day. That said, a lot depends on how you reflect. As mentioned above, just thinking about your teaching will probably naturally occur at most times as you teach, as well as before you enter the classroom and when you leave the classroom. This may not be deliberate reflection, and it may be in reaction to something that the teacher perceived to have occurred. As mentioned above, we need all the evidence we can get to make informed decisions about what happened, why it happened and what we want to do next.

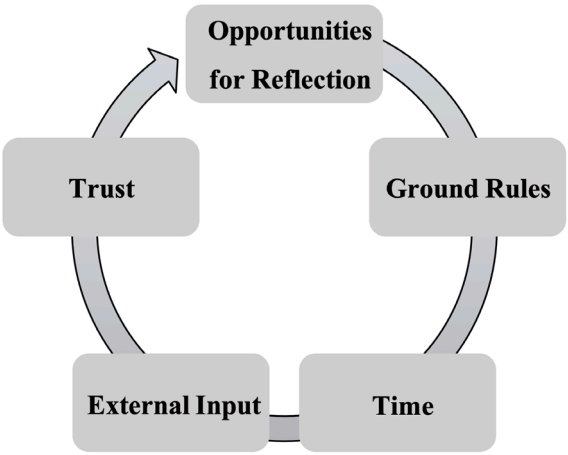
There are many different models and approaches about how teachers can reflect, too numerous to cover in this article (but see Farrell, 2019 for a comprehensive review of many of them). So, in this *Expositions* article I will outline two different approaches that I have developed over the years. An early model of reflective practice I developed emphasized a practical approach with the idea that practicing TESOL teachers would be better able to "locate themselves within their profession and start to take more responsibility for shaping their practice" (Farrell, 2004, p. 6) rather than relying on publisher produced materials and books that were rampant in the TESOL profession

at that time. I saw a need for teachers to be able to break away from relying on these badly produced textbooks along with teacher guides to *tell* them what they should be doing rather than taking responsibility for their own direction while teaching *their* students.

My initial framework was crafted to encourage teachers to look at their own practice with other teachers and decide their own future direction in terms of providing opportunities for their students to learn. This framework (Farrell, 2004) of reflective teaching is composed of five components: (a) a range of opportunities and activities; (b) ground rules; (c) provision for four different times or categories of reflection; (d) external input, and (e) trust. Figure 1 outlines this model.

Figure 1

Farrell Reflective Practice Framework (2004)



This framework (Farrell, 2004) illustrated above, is explained as follows:

1. *Opportunities*. A range of activities should be provided for teachers to reflect on their work. In this model the activities that were emphasized were group discussions, journal writing and classroom observations. These activities can be carried out alone, in pairs, or as a group. A group of teachers may decide to do one of the activities or a combination of any or all of them.

2. *Ground rules*. In order to avoid groups or individual teachers just drifting off into something other than reflection, this framework suggests a need for a negotiated set of built–in–rules or guidelines that each group or pair should follow in order to keep the drifting to a minimum. The model can be adjusted to individual group needs. Indeed, suggestions three through five are actually ground rules that can be built in to the activities. For example, who will chair the meetings and other such related question? For observations, certain understandings need to be negotiated ahead of time. For example, what are the responsibilities of the observer? Is intervention possible or desirable in the class? Will the class be videotaped, audiotaped, or neither? If you use a video, how will this be analyzed and why? What is to be observed and how? For journal writing, groups/pairs should negotiate the number of frequency of entries and the type of entries. The following list of general questions may help get a writer started: Describe what you do with no judgment? Why do you do it? Should you continue to do it or change it? What do others do? To suggest a set of built–in rules for critical friends while observing is not easy because there must be an element of trust and openness present in order to avoid putting emphasis on the critical while overlooking the friend. The friend can provide another set of eyes that both support and challenge us to get at deeper reflections of our teaching. To encourage this openness, the initial conversations between critical friends (or all conversations) should be taped and analyzed. This analysis can include the use of questions in their relationship, in terms of type, power structures established, focus of observation, and usefulness. In this way critical friends can negotiate what they want to achieve. Of course, all of the above activities and built–in guidelines cannot be accomplished quickly; like all valuable things, they take time. This introduces the next component of the model: time.
3. *Time*. For practicing teachers to be able to reflect on their work, time is a very important consideration. Groups can consider four different views/types of time: *Individual*, *Activity*, *Development, Period of Reflection*

- Individual: A certain level of commitment by individual participants in terms of time availability should be negotiated by the group at the start of the process.
- Activity: Associated with the time each participant has to give the project is the time that should be spent on each activity.
- Development: Another aspect of time that is important for teacher self–

development groups is the time it takes to develop. Analytical reflection takes time and only progresses at a rate which individual teachers are ready to reflect critically.

- Period of reflection. The time frame for the project as a whole is important to consider. How long should a group, a pair, or an individual reflect? Having a fixed period in which to reflect allows the participants to know what period during the semester they can devote wholly to reflection.

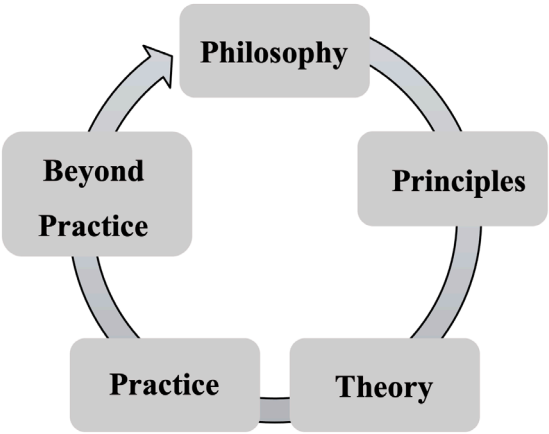
4. *External input*. The previous three suggestions utilize the idea of probing and articulating personal theories, which is at the center of teacher professional self–development. This involves process of constructing and reconstructing real teaching experiences, and reflecting on personal beliefs about teaching. However, at this level, reflection only emphasizes personal experiences but what do these mean in the greater professional community? Thus, external input of some kind is necessary to see what other teachers and groups have done. This external input can come from professional journals, other teachers' observations, and book publications of case studies.
5. *Trust*. The above four components of the model all pose some threat and associated anxiety for practicing teachers. Inevitably, there will be a certain level of anxiety present. Therefore, trust will be a big issue when teachers reflect together so a non–threatening environment should be fostered in the group by the individuals themselves.

The most important aspect of this early framework (Farrell, 2004) is to encourage reflection and to give teachers the opportunity to reflect, and I believe this framework is still relevant today: I have used this framework successfully and very recently with experienced TESOL teachers in a teacher reflection group in Canada (e.g., see Farrell, 2014), and it is still worthwhile for teachers wishing to reflect on their practice and especially with a group of teachers. In fact, the main topics the teachers talked about in order of frequency was their students (46% of the time) and how they had successes with them as well as challenges, the school context (44% of the time) in which they were teaching and mostly negative experiences with the administration and to a much lesser extent, their own teaching methods (10% of the time). I urge you to read this book and compare their experiences to your own in Japan.

In more recent times I began to work on a different framework that focused more on individual teachers reflecting holistically on their practice rather

than a group of teachers reflecting together as the early model above focused on. I call this the Framework for Reflecting on Practice (Farrell, 2015). As outlined in Figure 2 below, the framework has five different stages/levels of reflection: *Philosophy; Principles; Theory; Practice; and Beyond Practice*.

Figure 2
Farrell Framework for Reflecting on Practice (2015)



1. *Philosophy*. This first stage of reflection within the framework examines the "teacher–as–person" and suggests that professional practice, both inside and outside the classroom, is invariably guided by a teacher's basic philosophy and that this philosophy has been developed since birth. Thus, in order to be able to reflect on our basic philosophy, we need to obtain self–knowledge and we can access this by exploring, examining and reflecting on our background – from where we have evolved – such as our heritage, ethnicity, religion, socioeconomic background, family and personal values that have combined to influence who we are as language teachers. As such, teachers talk or write about their own lives and how they think their past experiences may have shaped the construction and development of their basic philosophy of practice. Reflecting on one's philosophy of practice cannot only help teachers flesh out what has shaped them as human beings and teachers but can also help them move onto the next level of reflection, reflecting on their principles.

2. *Principles*. The second stage/level of the framework, principles, includes reflections on teachers' assumptions, beliefs, and conceptions of teaching and learning. All three are really part of a single system, and thus difficult to separate because they overlap a lot, and, although I treat them separately in the framework, I see them as three points along the same continuum of meaning related to our principles. Teachers' practices and their instructional decisions are often formulated and implemented (for the most part subconsciously) on the basis of their underlying assumptions, beliefs and conceptions because these are the driving force (along with philosophy reflected on at level/stage one) behind many of their classroom actions.
3. *Theory*. Theory explores and examines the different choices a teacher makes about particular skills taught (or they think should be taught) or, in other words, how to put their theories into practice. Influenced by their reflections on their philosophy and principles, teachers can now actively begin to construct their theory of practice. Theory in this stage/level means that teachers consider the type of lessons they want to deliver on a yearly, monthly or daily basis. All language teachers have theories, both "official" theories we learn in teacher education courses and "unofficial" theories we gain with teaching experience. However, not all teachers may be fully aware of these theories, and especially their "unofficial" theories that are sometimes called "theories–in–use." Reflections at this stage/level in the framework include considering all aspects of a teacher's planning and the different activities and methods teachers choose (or may want to choose) as they attempt to put theory into practice.
4. *Practice*. Reflecting on practice begins with an examination of our observable actions while we are teaching as well as our students' reactions (or non–reactions) during our lessons. Of course, such reflections are directly related to and influenced by our reflections of our theory at the previous level and our principles and philosophy. At this stage/level in the framework, teachers can reflect while they are teaching a lesson (reflection–in–action), after they teach a lesson (reflection–on–action) or before they teach a lesson (reflection–for–action). When teachers engage in reflection–in–action they attempt to consciously stand back while they are teaching as they monitor and adjust to various circumstances that are happening within the lesson. When teachers engage in reflection–on–action they are examining what happened in a lesson after the event has taken place and this is a more delayed type

of reflection than the former. When teachers engage in reflection–for–action they are attempting to reflect before anything has taken place and anticipate what may happen and try to account for this before they conduct the lesson.

5. *Beyond Practice*. The final stage/level of the framework entails teachers reflecting beyond practice. This is sometimes called critical reflection and entails exploring and examining the moral, political and social issues that impact a teacher's practice both inside and outside the classroom. Critical reflection moves the teacher beyond practice and links practice more closely to the broader socio–political as well as affective/moral issues that impact practice. Such a critical focus on reflections also includes teachers examining the moral aspect of practice and the moral values and judgments that impact practice.

The framework can be navigated in three different ways: theory–into–(beyond) practice, (beyond practice–into–theory or a single stage application. Thus, it is a descriptive rather than a prescriptive framework. Teachers can take a deductive approach to reflecting on practice by moving from theory–into–practice or from stage/level 1, philosophy through the different stages to stage/level 5, beyond practice. Some may say that pre–service teachers who do not have much classroom experiences, would be best suited to take such an approach because they can first work on their overall philosophical approach to teaching English to speakers of other languages and work their way through the different stages of principles (stage/level 2), theory (stage/level 3) when they reach the practicum stage, they will be well placed then to reflect on their practice (stage/level 4) and eventually move beyond practice (stage/level 5). This theory–driven approach to practice where philosophy and theory have an initial influence on practice is probably a natural sequence of development for novice teachers because they do not have much teaching experience. When their early practices are observed, it is most likely that theory can be detected in their practice; however, over time, and with reflection, it is possible that their everyday practice will begin to inform and even change their philosophy and theory and they may come up with new principles of practice.

Experienced teachers too can also choose to begin their reflections at stage/level 1, philosophy especially if they consider their philosophy as a significant basis of their practice with principles second, theory third and so on through the framework. For experienced teachers some of whose practice can be theory–driven if they have been reading and experimenting with

applications of particular theories throughout their teaching careers, most likely describe their work in terms of their overall philosophical approach to teaching English to speakers of other languages and this description probably embeds a lot of their values, beliefs, principles and well as theories behind their practice. When such teachers are observed teaching their lessons, we are likely to see that their approaches, methods and activities often reflect the influence of these theories.

Attached to the "when" and "how" of reflective practice is the time teachers have to reflect. Many teachers are very busy and as such may consider the above approaches too time consuming for them to engage in. I agree to a certain extent that it can be time consuming, but it would be time well spent. I would also suggest that teachers begin at whatever stage they feel comfortable with above (e.g., your philosophy or your principles) when you have the time and work your way around the framework as you see fit. In this way teachers can use the framework as a lens through which they can view their professional (and personal) worlds—what has shaped their professional lives—as they become more aware of their philosophy, principles, theories, practices and how these impact issues inside and beyond practice. I believe that such a holistic approach to reflection produces more integrated second language teachers who have self–awareness and understanding to be able to interpret, shape and reshape their practice throughout their careers. The information that is produced from reflecting during each stage can be compiled into a teaching portfolio and used for collaborative teacher evaluation purposes. In such a manner the teacher is not separated from the act of teaching when reflecting or being evaluated.

Implementing Reflective Practice in Japan

So far in this paper I have outlined and discussed two major frameworks that language teachers can implement individually or in groups to facilitate their reflections. The first framework I outlined was a broad implementation of reflective practice that most likely serves groups of teachers reflecting together rather than individual teachers reflecting alone. I would recommend a group of three or four teachers come together weekly (or whenever possible) for one semester and consider using that early model when considering the (a) range of opportunities and activities they intend to follow, (b) the specific ground rules the group wants to follow when engaging in reflection for one semester, (c) provision for four different times or categories of reflection (*individual*, *activity*, *development,* and *period of reflection*), (d) what kind of external input they will use (see next sentence), and (e) how they

will develop trust in each other throughout the process. In this regard, I urge interested groups of teachers to read a paper I wrote for a short version of how this all works and what the teachers focused on in Farrell (2014a), and/or a longer version that details everything in book form in Farrell (2014b).

When implementing the second framework for reflecting on practice you can read how it was used recently in the case studies outlined in the work of Farrell & Kennedy (2019), Farrell & Avejic (2020), Farrell & Macaplinac (2021), Farrell (2022), and most recently Farrell & Moses (2023). Indeed, in a recent published review of 92 studies on reflective practice in second language education, Sarab and Mardian (2022) highlighted the usefulness and importance of the second framework for reflecting on practice in all global contexts that include Japan when they noted that "one central benefit of Farrell's framework is its specific and holistic nature" (p. 13). They continue: "Besides, another striking feature of the model is that it functions in a reflective–reflexive manner, meaning that the model not only views 'reflection as an analytical process' but emphasises 'the mirroring of practice, and thereby undertaking a self–analysis'" (p. 13). The authors especially recommend the use of the framework in all contexts (such as Japan) because it includes critical reflection that I call beyond practice, or the fifth stage of the framework outlined above. Sarab and Mardian continue:

> It is through critical reflection or beyond practice – the last stage in Farrell's framework – that the benefits of reflection can be applied to social contexts. With such a critical focus on reflection, research can provide insights into how L2 teachers around the globe explore the moral, political, and sociocultural issues that impact their performance inside and outside the classroom. (p. 14)

In this paper I outlined and discussed my approaches to reflective practice that I believe will be useful for teachers wishing to engage in reflection on their work in Japan. I should also point out that I fully recognize that the concept of reflection is certainly not new to Japan with its rich history of Buddhist practices that has existed for centuries (Watanabe, 2016). In her important work on the concept of reflective practice in a Japanese context, it is interesting to note that Watanabe (2016) has pointed out that there is no agreed Japanese translation for the term "reflective practice" which suggests it is still new(ish) in education circles. Watanabe (2017) used the term *kotodama* or "word spirit [for] "putting one's inner thoughts into words" (p. 98) as a reflective communication convention among the Japanese people.

Watanabe included this interesting concept in her study of seven in–service high school teachers of English reflections that show that reflection is highly contextualized. In her study, Watanabe conceptualizes teacher reflection and development as 'expansion' rather than 'change' and she places teachers, who she notes are equipped with different strengths and weaknesses, at the centre or the core of the activity of their own reflection and development. Watanabe notes that rather than shedding their old practices, teachers in Japan she says are encouraged to expand their repertoires of use. Watanabe (2016) continues:

> The 'expansion' model, which places teachers in the centre, also allows teachers more autonomy in taking responsibility for both student learning and their own growth. In the study, reflective practice helped my participants to recognise that they were driving forces in leading the students to learning. Their notion of themselves as teachers also expanded to include a new awareness that they had agency. They acknowledged that the locus of control for their own growth was themselves and expanded their sense of being agents of their own development. (p. 289)

Another interesting approach to the implementation of reflective practice in Japan was a recent study by Chris Harwood and Dennis Koyama (2022) where they implemented reflection within an onboarding process for hiring new faculty at universities as a way of facilitating success in and acclimatization to their new work environments. Specifically, they outlined how they successfully implemented a reflective practice process that included a routine of reflecting *in*, *on* and *for* action. Harwood and Koyama's (2022) four stage framework (pre–class, in–class, post–class and meta reflection) were used to evaluate the efficacy of existing curricular materials to inform adjunct–faculty in an undergraduate English composition program's onboarding and professional development. They cite several benefits of implementing such a system such as more rapid troubleshooting before the lesson occurs in the pre–class stage, a high level of teaching engagement in the in–class stage, and more in–depth discussions among teachers in the post–class stage. In addition, in the meta–stage students' perceptions about materials were included in reflections and of course such inclusions lead to more student reflections on their own learning which should always be included in any reflective practice process. Harwood and Koyama (2022) also include an important aspect of such meta–reflections by their writing

up of their study (as did Watanabe, 2016, 2017) leading to its publication where they can share their experiences with others.

I urge readers to investigate both these studies when wishing to engage in reflective practice in Japan as well as the following publications on this interesting yet complex topic of reflective practice in language.

Barnard, R., & Ryan, J. (Eds). (2017) *Reflective practice: Voices from the field*. Routledge.

Barnard and Ryan's (2017) collection contains reflective practice studies of TESOL teachers (preservice and inservice) on topics such as (collaborative) lesson planning, classroom observation, lesson transcripts, post–lesson discussions, journal writing, reflection on action, reflection in action, critical friends, and focus groups. The aim of the book is to explain a range of options for implementing the reflective practice cycle in educational settings in various international contexts. Written by international academics, these studies show how reflection can be interpreted in different cultural contexts.

Mann, S., & Walsh, S. (2017). *Reflective practice in English language teaching*. Routledge.

Mann and Walsh's (2017) book outlines an empirical, data–led approach to reflective practice and uses excellent examples of real data along with reflexive vignettes from a range of contexts in order to help teachers to reflect on their practices. Mann and Walsh also note the importance of dialogue as crucial for reflection as is allows for clarification, questioning and enhanced understanding.

Tajeddin, Z., & Watanabe, A. (Eds.). (2022). *Teacher reflection: Policies, practices, and impacts*. Multilingual Matters.

This edited book has been compiled in honor of Thomas S. C. Farrell, one of the most distinguished scholars in theorizing and researching language teacher reflection. It examines teacher reflection in three main areas: policies, practices, and the impact of teacher reflection on teachers' practices and professional development. The data–driven chapters shed light on concerns and challenges experienced by teachers in diverse international contexts and institutions and discuss the practical implications of their findings across a variety of policy settings. The book addresses aspects of reflective practice including macro and micro policies and constraints, as well as opportunities in the engagement of reflective practice. In addition, it

explores teachers' identity, cognition, emotion and motivation, areas which are relevant but often not discussed in the literature on reflective practice (from the publisher's webpage: https://www.multilingual–matters.com/page/detail/Teacher–Reflection/?k=9781788921022)

Conclusion

Reflective practice as it is outlined in this article is much more than taking a few minutes to think about our teaching. Most teachers do this regularly after a class, or on the way home from school. Reflective practice as it is outlined here is evidence–based because involves teachers' systematically gathering data about their teaching and using this information to make informed decisions about their practice. Reflective practice is more than a method, it is really a way of life. Teachers can engage in reflective practice at any stage of their careers and at any time of the teaching day as they continue to construct their own personal theories of teaching and improve their instructional practice. Teachers who engage in life long reflective practice can develop a deeper understanding of their teaching, assess their professional growth, develop informed decision–making skills, and become proactive and confident in their teaching and possibly their personal life as well. I wish all the readers of the *JALT Journal* a happy reflective journey.

References

Fanselow, J. (1987). *Breaking rules*. Longman.

Farrell, T. S. C. (2004). *Reflective practice in action: 80 reflection breaks for busy teachers.* Corwin Press.

Farrell, T. S. C. (2013). *Reflective writing for language teachers*. Equinox.

Farrell, T. S. C. (2014a). "I've plateaued…gone a little stale." Mid–career reflections in a teacher reflection group. *Reflective Practice. 15*(4), 504–517. https://doi.org/10.1080/14623943.2014.900029

Farrell, T. S. C. (2014b). *Reflective practice in ESL teacher development groups: From practices to principles*. Palgrave Macmillan. https://doi.org/10.1057/9781137317193

Farrell, T. S. C. (2015). *Promoting teacher reflection in second language education: A framework for TESOL Professionals*. Routledge. https://doi.org/10.4324/9781315775401

Farrell, T. S. C. (2018). *Research on reflective practice in TESOL*. Routledge. https://doi.org/10.4324/9781315206332

Farrell, T. S. C. (2021). *TESOL teacher education: A reflective approach*. Edinburgh University Press. https://doi.org/10.1515/9781474474443

Farrell, T. S. C. (2022). *Reflective language teaching*. Cambridge University Press.

Farrell, T. S. C., & Avejic, V. (2020). "Students are my life": Reflections of one novice EFL teacher in Central America. *TESL Canada Journal, 37*(3) 47–63. https://doi.org/10.18806/tesl.v37i3.1345

Farrell, T. S. C., & Kennedy, B. (2019). Reflective practice framework for TESOL teachers: One teacher's reflective journey. *Reflective Practice, 20*, 1–12. https://doi.org/10.1080/14623943.2018.1539657

Farrell, T. S. C., & Macapinlac, M. (2021). Professional development through reflective practice: A framework for TESOL teachers. *Canadian Journal of Applied Linguistics, 24*(1) 1–25. https://doi.org/10.37213/cjal.2021.28999

Farrell, T. S. C., & Moses, N. (2023). "Class is like a family": Reflections of an experienced Canadian TESOL teacher. *Korea TESOL Journal*. Advance online publication.

Harwood, C., & Koyama, D. (2022). Employing reflective practice to inform adjunct–faculty onboarding. *Innovation in Language Learning and Teaching, 16*(2), 191–203. https://doi.org/10.1080/17501229.2021.1895800

Sarab, M., & Mardian, F (2022). Reflective practice in second language teacher education: A scoping review. *Journal of Education for Teaching*, Online first Oct 28, 2022.

Walsh, S. (2015). *Classroom interaction for language teachers*. TESOL Press.

Watanabe, A. (2016). *Reflective practice as professional development experiences of teachers of English in Japan.* Multilingual Matters. https://doi.org/10.21832/9781783096985

Watanabe, A. (2017). Journal writing as a space for reflection: The concept of kotodama. In R. Barnard & J. Ryan (Eds.), *Reflective practice: Voices from the field* (pp. 98–107). Routledge. https://doi.org/10.4324/9781315397665-11

A Journal on Student, Teacher, and Researcher Journal Writing

Christine Pearson Casanave
Temple University, Japan Campus
Middlebury Institute of International Studies at Monterey

Journal writing is an open, unevaluated form of free writing that can be used by L1 and L2 students, teachers, and researchers to help them develop language fluency and proficiency, to promote reflective and critical thinking, to contribute to research activities, and to communicate with self, peers, and teachers. The main point of journal writing is to promote evaluation-free reflection and communication, and when written in the L2, to encourage fluency rather than be used as a graded and corrected assignment. This article consists of a series of journal entries about my experiences with different kinds of journal writing over many years. The entries describe "dissertation journals," language learning journals, private and shared journals, journals for promoting critical and analytical reflection, and journals written for research purposes.

Keywords: evaluation-free writing, fluency, language development, reflection, response to reading, research tool

A Short History of My Journal Writing Experience

Journal Entry 1

Today I got an interesting email from Dennis Koyama and his associates at the *JALT Journal*, inviting me to contribute a short piece to a new section

https://doi.org/10.37546/JALTJJ45.1-7

of *JJ* (Expositions) that will appear in an issue in May 2023. He was familiar with my work on journal writing (much of it done while I was in Japan). As I thought about Dennis's invitation, I wondered first if I should accept, and then when I did, how to organize an article for *JJ* and how to write in a style compatible with my messages. It occurred to me the other day that I could write it as a number of journal entries on various aspects of journal writing that have struck me as important over the years. So that is what I decided to do. Following a journal writing style that I am comfortable with, I am trying not to be bound by rigid academic conventions, and instead composing this article as a series of journal entries.

This project for the *JALT Journal* takes me back over much of my academic career and up to the present moment, and pushes me to reflect on the place of various kinds of journal writing in my personal and professional life and in the lives of past and present students. This first entry is too long to be considered a "short" history, but it covers the background that I wanted to include about how my interest in and experiences with journal writing developed over time.

I began decades ago, before I went to Japan, writing what I called "dissertation journals" while I was in my PhD program. My dissertation project concerned how writing helped socialize graduate students into their disciplines, and as I wrote in my journals about my developing ideas and my responses to readings, I was doing for myself what I would write about in my dissertation—getting socialized into a discipline. I filled three hardbound casebooks with handwritten entries over several years. These journal entries included lots of reflections on and stories of my case study participants and their professors whom I was interviewing and observing. I wrote about more participants than I needed too, but I think that in this kind of journal writing for case study research purposes, there is no way to know ahead of time which participants will become central cases and which will drop out, so good to write about all of them. The best thing about using my journals to write about individual participants was that I came to know and respect them deeply and to care about progress they were making in their own doctoral studies. As case study participants in my own project, they came alive, and I could write about them with compassion and understanding. It was during this time that I began to realize the value for researchers of keeping journals, a realization that has influenced my later work as a dissertation advisor with students in Japan.

In my dissertation journals, I included not just reflections on my participants, but also some responses to readings and reflections on my own dis-

sertation progress and interactions with two advisors. Some of the responses to readings, in revised and synthesized form, would find their way into the dissertation literature review chapter. The reflections on my progress and on my interactions with advisors helped me monitor and adjust my moods and attitudes, allowing me to express privately what I might not have been able to share with them or with classmates (no matter how much we liked to gossip about our professors). These lessons I learned about the personal value of writing private journals during difficult academic work stayed with me and I hope have been passed on to my own graduate students in Japan.

But not all journal writing needs to be private. Whether writing in their L1 or their L2, students who write journals to their teachers and professors are helping themselves move their work and their academic language forward. (Vanett & Jurich, 1990, wrote about this a long time ago.) The journal writing helps students turn ideas into language, and helps teachers understand what their students are doing and thinking. During my doctoral program and dissertation preparation, I wrote what I called "academic letters" to one of my advisors. I didn't call these journals at the time, but they served the same purpose, including serving as a forum for my main advisor and me to communicate. I wrote about questions I had, ideas I was developing for the dissertation project, and comments on some of the readings I was doing. I somehow wrote more freely when I was not being evaluated by her for a grade or writing required academic papers for a class or independent study. I am not sure how many dissertation advisors have time to read and respond to journal-like "academic letters," but in my own current work with doctoral students in Japan, I find that dissertation students who write me long journal-like emails or send Word files of reflections on readings and progress get pretty prompt attention from me, including encouragement to keep working. The most important part of this kind of journal writing for students of any kind is that it is evaluation-free. How liberating!

My early experiences with journal writing continued after I completed my PhD, but switched to journal writing for language development (see the next section). I began reading about the benefits of journal writing and got very interested in journal writing as a way help to my students in Japan develop fluency in English and to lose their fear of making mistakes. Drawing on my classroom practices and on my readings, I published a few articles (Casanave, 1992, 1994, 1995), a small edited collection with Keio SFC teachers, (Casanave, 1993a), and a book on journal writing (Casanave, 2011), and made a few conference presentations. I was inspired early on by Joy Kreeft Peyton's work on dialogue journals (Peyton, 1990; Peyton & Reed, 1990;

Peyton & Staton, 1991, 1993), Rebecca Mlynarczyk's (1998) book on journal writing with second language learners, and others who devoted their teaching, research, and writing to issues in journal writing.

During my years in Japan, I asked my undergraduate EFL students to write journals regularly in English, hoping they would develop fluency and lose their fears of making mistakes. Some students, predictably, hated this journal writing experience and others blossomed. But all of them developed fluency and speed and expressivity, much to their surprise by the end of a term. I also, for a time, kept a handwritten journal in my baby Japanese (which I had never studied previously) that documented my efforts at learning Japanese kanji, kana, and syntax by writing about my daily life. I let my EFL students know about this effort, and shared some of my awkward Japanese journal entries with them. I was hoping to write these entries in the style of a dialogue journal with a Japanese friend outside the university context, but when he read these journals, he mainly made some language corrections, even though I had asked him to write back in the style of a conversation. (See my 2012 "Diary of a Dabbler" article.) The eye-opener of this experience was that I was trying to do what I had been requiring my own EFL students to do, and came to appreciate the burdensome but ultimately gratifying task, as well as the importance of a teacher's substantive responses, not meticulous language corrections, to students' journals.

During my early years in Japan, I also began teaching master's and doctoral students at an American university campus there, and periodically asked the graduate students (all very advanced L2 English or native English speakers) to write a research journal or a dissertation journal, to which I would respond. This early experience with graduate students convinced me of the value of journal writing or its equivalent in helping students formulate and refine ideas, grapple with problems with methods or participants, and interpret findings (see more below, in the section Journal Writing for Research).

Fast forward to the present day: I continue to write a perfunctory handwritten personal journal in English, but with little of the beneficial reflection I have long touted and almost none of the L2 language practice, but for the recent emails with a friend in Spanish that I mention below. My current handwritten journal is mainly a record of things happening around me or in the world, including connections I have with a few friends around the world or occasional health matters, details that I record so as to consult and remember them if needed. My email correspondence with a few friends, some of which is journal-like, continues to serve purposes that have been

touted in the journal writing literature as beneficial for promoting reflection and developing ideas. As of this writing, one of these correspondents writes me in Spanish, which as an L1 English speaker he is trying to improve, and I respond in Spanish (my strong L2). In general, these email "journals," both those in English and Spanish, provide an ongoing record of the lives and thoughts of my correspondents and me—a lot of narratives, a few rants, and some good language practice. Such practices benefit students, teachers, and researchers alike. My remaining entries are organized by theme, not by chronology.

Journal Writing for Language Development

Journal Entry 2

A basic question about journal writing for language development concerns whether the mere act of writing in an L2 extensively and over time (as I tried to do with my Japanese journal, as I asked my EFL students in Japan to do in English, and as I am trying to do now in Spanish) will contribute to that development, or whether teacher or peer feedback (comments or corrections) is necessary for this development. In my undergraduate EFL classes in Japan, students started out writing journals weekly. I felt strongly that my feedback comments (not corrections) would help students develop their thinking and conceptualizing and would provide models for vocabulary and grammar that fit what they were trying to say. The commentary from me demonstrated to the students that they had a real reader who was paying attention to what they were saying, not just how they were saying it. Fluency soared, and students often commented that they were able to complete a journal in less time than when they started. But after a few semesters of trying to read and comment on 90 journals a week, I switched to bi-weekly submissions, and still barely managed to keep up. But it was worth it.

My writing speed in my Japanese journals improved too. I was hoping to get substantive feedback on my Japanese journals, of the sort I was providing my EFL students, but my Japanese friend-tutor responded only with corrections to my grammar, kanji, and kana. I was grateful, but the tutorial experience differed from what I was hoping to experiment with on myself and from what much journal writing literature recommended (i.e., free, uncorrected writing). There really is a case to be made for free writing that teachers or friends possibly read but do not comment on, as well as free writing that is for the writer's eyes only (as Peter Elbow [1973, 1999] told us long ago about free writing in L1). In both cases, if students (and teachers along with them) are writing regularly in their L2, the mere practice suppos-

edly improves their fluency of expression, if not the grammatical accuracy of their expression. And if journal writers are writing about topics they have read about in the target language and/or topics they are interested in, so much the better. In all cases in the school context, they are using their journal writing to develop their academic literacy (Fogal & Koyama, 2022).

I always liked the idea of reading response journals to help students expand their academic literacy. Readings are a primary source of learners' vocabulary and syntax, even if items are only copied verbatim into their journals, but especially if incorporated into their own descriptions and discussions. By observing closely how the L2 works in an interesting reading (including in blogs and emails with competent L2 users) and in feedback and commentary they might receive, students shift their focus from learning rules and memorizing words to using patterns and words (even in nonstandard ways) to interact with authors and with correspondents and to make meaning. How does an author or correspondent express this or that idea? What words (nouns and verbs in particular) are used to express an idea, describe a process, or make an argument? Can students use some of those same words and parts of sentences in their own journals and other writings, without plagiarizing? Can they foster their language development by patchwriting (Pecorari, 2003)?

Journal Entry 3

Some months after completing the previous entry, I was deeply into rereading a long novel in Spanish by Carlos Ruiz Zafón, *La Sombra del Viento*. Even on the second reading, which admittedly was going faster than the first attempt, I continued to find the reading difficult, mostly due to a great deal of vocabulary that I could not even find in my tattered pocket dictionary. I occasionally wrote email in Spanish about this book to my friend who was trying to learn more Spanish. Even though I did indeed make a few notes on interesting patterns and phrases in this book, as I had hoped my EFL students in Japan would do with their reading, I neglected to write in a journal about my responses to this complex story and to the language. Thinking about my students in Japan forced to write journals in response to difficult readings, I wrote to my friend about my inability to read smoothly if I had look up every other word in my Spanish-English dictionary and about my wilting motivation to persist: "Quisiera avisarte que continuo leyendo la novela muy larga de Carlos Ruiz Zafón (*La Sombra del Viento*) con interés, pero me siento un poco desanimada a causa de todas las palabras nuevas. Es imposible leer si busco cada palabra desconocida en el diccionario" (email, October 10,

2022). I think my Spanish and my motivation would have improved had I been able to make myself write a reading response journal of some kind in Spanish, even in the form of more emails in Spanish to my friend.

Journal Entry 4

I asked my friend who is trying to improve his Spanish, now that he is retired and has more time, to explain how he was doing this, and he wrote me that he is using what he calls "diarios" (what I am calling journals) to help him with his language development. I was curious as to how his experiences compared with those of some of my EFL students in Japan. He wrote me, in English, that in the past, his Spanish language development was hindered because he hesitated to speak or write if his grammar and vocabulary were not perfect (the dilemma that many EFL students in Japan face). He has found a way more recently to use Google Translate to help him express himself in his "diarios." In explaining how he tried to overcome his fear of making mistakes, he told me this:

> "So I decided to start writing in a diary every day. I've been doing that for about 5 months now. I write just 8-10 lines of Spanish about anything that comes into my mind. And I make extensive use of Google Translate in the following way: After I've written a sentence or two in Spanish, I have GT translate it into English. If the English doesn't reflect what I was trying to say, I work with the Spanish until it does. And I'll use GT to help me say it. Then I'll switch the process and have GT translate the translated English into Spanish. In the early days of the diary, I was using GT to do a lot of the translating from English to Spanish. As I went along, however, I got better at having the Spanish->English translation say what I was trying to say – without having to use GT to do it. And my vocabulary and phrases for everyday things improved. […] What has surprised me about writing a diary in Spanish is that it has helped me get over my fear of being misunderstood when speaking Spanish by not having perfect vocabulary and grammar."

This is a lesson he has learned late in life, and one that is central to the benefits of journal writing for language learning: first, that our L2 does not have to be perfect for us to be able to communicate and to improve, and second, that writing in the L2 does indeed contribute to language develop-

ment. My hope in my own teaching career has been that if EFL students in Japan and elsewhere can use journal writing to overcome their own fears of making mistakes, their fluency and their motivation will flourish, and language development will happen even without their needing to perfect every phrase.

Journal Writing for Educational Development

Journal Entry 5

By "educational development," I refer loosely to growth in the kind of reflection and critical and analytical attention to content, ideas, and arguments that we expect from a student, teacher, or professor in an educational setting (Casanave, 1995; Lee, 2008). I'm not sure, but it is possible that this kind of development does not happen in journal writing without intervention of a more competent interlocutor, in the style of Vygotsky's (1978) scaffolding within the Zone of Proximal Development. Interactions with more competent others can occur in a variety of ways: oral conferences for discussing journal entries, substantive feedback from teachers with revisions and responses from students, and dialogue journals. However, as some of my students in Japan told me, even though journal writing "changed my college life" for the better because they were reflecting on things they would not have thought about without the journals (Casanave, 1993b), they needed to be writing on topics they were somewhat familiar with. None of us, writing in L1 or L2, can write fluently and meaningfully on topics we know nothing about—ones that are outside our Zones of Proximal Development.

Oral Conferences

In the context of Japan, oral conferences and consultations between English teachers and students can provide the kind of engagement and attention that young university students as well as graduate students can benefit greatly from. These conferences do not have to be one-on-one, but can be between two or three students and their teacher or even just among students. The point is to have a small gathering focused on students' journal entries, where students can read their journal entries aloud and engage the interest and commentary of one or more listeners. Of course, this activity can be done in a whole-class format as well, but some students might be fearful to display their possibly imperfect English in front of all their classmates, or would hesitate to express any private thoughts in such a setting. In pairs or small groups, if there are no serious privacy issues, the hope is that listeners-

readers would be attentive for bits they might not have understood, or that they found particularly interesting, or that they have ideas and information about that will further the journal writer's knowledge and understanding. Brave teachers who are writing journals in their L2 (their students' L1) can also share their own L2 journal entries with student listeners or readers. Students in this way become empowered as "teachers" of their teacher, a confidence booster if there ever was one.

Teacher Feedback and Student Response

Teacher or peer feedback on journals (NOT corrections) is what enables an "educational discussion" to take place, particularly if teachers or peers ask journal writers to continue the discussion with follow-up responses to the feedback. For this feedback-response discussion to succeed, both teachers and students need to set aside the conventional attitude that student written work needs to be graded and corrected. (In my undergraduate classes in Japan, I would mark the journals as done or not done, without a grade, but added numerous comments and questions on each one, time permitting.) It might be easier to shift to this interactional attitude if both students and teachers can conceptualize the journal writing activity as a dialogue, both between-among students, and between teachers and students. It can help to label the activity with the well-known moniker from many years ago, the "dialogue journal."

Dialogue Journals

As I mentioned in the introduction, dialogue journals (Peyton, 1990) consist of ongoing written communication between students and teachers or peers, whether done electronically or by hand and in oral conferences. If students are not confronted with too many rules and regulations for how to write their journal, and if they are not required to revise them unless they wish, they can greatly improve their L2 written fluency, and depending on the topics they have chosen (or been asked) to write about, can also deepen their thinking about themes in their lives or about topics they are learning about in their classes.

The main problem I faced with this interactive journal activity in Japan was the size and number of my classes. As I mentioned in Journal Entry 2, I recall one term in which I was receiving 90 journals a week to read and comment on, clearly an impossible load to keep up with, even nowadays if done all electronically. In such cases, it seems that teachers have only two

choices: Ask students to write fewer journals, or cut back on the number and depth of responses to them.

Peer reading and responding helps, but in my experience, students really appreciate the one-on-one personal responses from a teacher. This means that the responses cannot be generic and perfunctory (like the classic comment "Interesting!"). They need to be a sincere act of communication, a real dialogue. To this end, teachers who make real comments and ask real questions of students, in the hope of a response of some kind in the next journal entry, are setting themselves up as genuine communication partners in the target language. Likewise, students who ask real questions in their journals of both their teachers and their peer readers are communicating purposefully in their L2 perhaps in ways they would not be able to do as easily in their L1. (How many students have we known who ask genuine questions of their teacher!)

However, in my experience in Japan, I found it difficult to break the expected pattern of "teacher question—student response," particularly with undergraduate students. The teacher is supposed to know everything, correct everything, control everything, in the conventional view. And some students believe that if a piece of written work is not graded and corrected it does not count as an educational activity to be done as part of a language class. There is also a deep-seated belief by many that one's L2 cannot improve without errors being corrected. One of the initial challenges for teachers is thus to help students understand the purposes and procedures of communicative journals. It can help if teachers can present evidence to students of positive changes and development of the L2 over time, perhaps with data from previous classes or from the literature. And teachers who write journals in their L2 along with students are ideal models for the benefits of journal writing.

One caution—the fatigue factor. By the end of a term, both students and teachers are likely to be tired, pressed for time, and ready for the term to end. Journal entries might become shorter and be done less carefully, possibly adding to rather than reducing language errors. Adjustments can thus be made at any time during a term, to maintain the freshness and purposefulness of the (dialogue) journal activity.

Journal Writing for Research

Journal Entry 6

Journal writing is a fabulous activity for both students and teachers who are doing research projects, no matter how simple or complex the project.

Research journals can record all kinds of valuable information, in both the L1 and the L2. As I wrote in the introduction, my dissertation journals formed a central part of my PhD research activities, helping me to process readings, to pose and clarify ideas and questions that were curiosities and puzzles to me, to record what happened at various stages, and to work through occasional difficult encounters with participants or advisors. Moreover, in a private research journal, we can rant with impunity.

In a research journal that takes the form of email exchanges between students and teachers, as I have done for many years now with students in Japan who are writing doctoral dissertations, both students and their teacher-advisor have an ongoing record of how a project develops over time, how problems and confusions are confronted and overcome, and how drafts of written work with substantive feedback develop into finished products. We also have a record of contextual and environmental factors that might be influencing motivation and engagement.

Hence, from undergraduate to doctoral level, as well as part of a teacher's professional activities, journal writing can record numerous kinds of information, activities, and feelings that benefit research projects, either for actual use later in the project or its write-up, or for processing privately the confusions and complexities that go hand in hand with both short-term and longer-term research projects.

Journal Entry 7

I don't know how many teachers write journals, but teachers who do engage in journal writing do so for numerous reasons. If their journals are personal, the writing might simply be a private way to decompress after a stressful day, to plan for the next day, or to reflect on how a day, a week, or even a term has gone. What worked and what didn't work in today's class? What readings have I come across that might address some of the issues that I face in my teaching? What might work better tomorrow, or next term, or next year? What teachers do I know with whom I could share some of my journal entries and maybe get their perspectives? A teacher's journal can be used for both planning and reflection on current teaching and for developing writings, conference presentations, and research projects that might be connected with that teaching.

Researchers too benefit greatly from keeping journals, and indeed, researcher journals constitute a kind of data in some kinds of projects. Ideas and quotes can be drawn directly from them that include field observations, summaries of conversations and interviews with participants, and descrip-

tions of research sites and activities. Such journals also record responses to readings, initial ideas for projects, hypotheses and speculations, initial analyses of data, problems and successes with participants, and the development of arguments and interpretations.

Some teachers and researchers—particularly overworked ones in Japan—might protest that they have no time or energy for journal writing, on top of their normal work lives and snatches of private life if they can even manage these. It's true that journal writing requires some time (though not necessarily a lot), and a place and time to write that afford the writer chances to think, reflect, and compose, even if just for 15 minutes a day, without interruptions. Journal writers thus need to know themselves—where and when are they able to write? An office, a coffee shop, a room at home, or a noisy bumpy bus or train? With or without music? How about on a walk, without anything except a device for recording oral musings? What about a pictorial journal (a good idea for students too?) using digital images, or hand-done drawings, dated, and labeled with or without extensive commentary, and discussed or not with others? The point is that teachers can apply the same purposes to their own researcher journals as they do to their students' research journals. In all cases, questions, ideas, and experiences get transformed into language.

Concluding Comments

Journal Entry 8

As this journal on journal writing comes to its end, I wonder what Words of Wisdom, Hope, and Motivation I might convey to students and teachers in the Japan context and to myself as an L2 language user. In Japan, it continues to be challenging for EFL students to use English outside the confines of the classroom, and for me, living in California, with little regular access to the L2 (some Spanish, a little Japanese, and even less French since my ancient cousin in France died), I have to create opportunities to use my L2s. As I discovered long ago through my struggles to become a competent user of Spanish and a survival user of Japanese, languages are rarely learned deeply in a classroom context or in sporadic tutorials. It can even be challenging for English teachers in Japan who are L2 learners of Japanese to acquire Japanese in a steady and naturalistic way, unless one happens to be blessed with a patient and understanding Japanese partner. I recall that my attempts at using my survival Japanese in naturalistic contexts when I was living in Japan were often met by responses in English, or as was the case with my attempts at journal writing in Japanese, by language corrections.

So whatever these final words are, they must also apply to myself. Real communication partners help greatly in the journal writing effort, providing both an audience of interested (we hope) listeners-readers, and intentional and unintentional feedback of various kinds, not just corrections. What I have learned most recently in my efforts to improve my Spanish is that real communication in an L2 can come from unexpected places, including from communication partners who are themselves learners of the L2. As I think back on these efforts to communicate in Spanish in this past year, I have come to understand the form and function of journals in expanded ways. My main "journal writing" in Spanish is in the form of email, with some entries just a few lines long. I don't currently write journals in Spanish by hand, although I certainly could. (Why could I not write my daily excessively boring and often very short journal entries in Spanish, or partly in Spanish instead of in English? Why could I not throw in a few words and phrases in Japanese kana or in my disappearing French? Why could I not ponder some of my writing topics, challenges, and dilemmas using a mix of all these languages? Why could I not communicate occasionally with my L1 and L2 doctoral students in Japan, even briefly, in a bit of Japanese?)

Hmmm. So at the end of this journal on journal writing, and at this late stage of my academic career, it seems there is quite a bit left I could do to practice what I preach. *Demo mokuteki ga nan desuka?* Role model? Lifelong learning? Enjoyment and intellectual stimulation? *Zenbu? Ganbarimash*ō.

Author Bio

In Japan post-PhD, I taught at Keio University's SFC campus, and part time at the Japan campus of Columbia University's Teachers College. My next post was in TESOL and applied linguistics at Temple University's Japan Campus (TUJ), work that continues online as doctoral dissertation advising from my home base in Monterey, California. My publications concern academic writing, and I am currently co-book review editor of the *Journal of Second Language Writing*.

References

Casanave, C. P. (1992). Educational goals in the foreign language class: The role of content-motivated journal writing. *SFC Journal of Language and Communication, 1*, 83–103.

Casanave, C. P. (Ed.) (1993a). *Journal writing: Pedagogical perspectives*. Keio University SFC Monograph No. 3. (ERIC ED423682)

Casanave, C. P. (1993b). Student voices: The insiders speak out on journal writing. In C. P. Casanave (Ed.), *Journal writing: Pedagogical perspectives* (pp. 95–115). Keio University SFC Monograph No. 3. (ERIC ED423682)

Casanave, C. P. (1994). Language development in students' journals. *Journal of Second Language Writing, 3*(3), 179–201.

Casanave, C. P. (1995). Journal writing in college English classes in Japan: Shifting the focus from language to education. *JALT Journal*, *17*, 95–111.

Casanave, C. P. (2011). *Journal writing in second language education*. University of Michigan Press.

Casanave, C. P. (2012). Diary of a dabbler: Ecological influences on an EFL teacher's efforts to study Japanese informally. *TESOL Quarterly*, *46*(4), 642–670. https://doi.org/10.1002/tesq.47

Elbow, P. (1973). *Writing without teachers*. Oxford University Press.

Elbow, P. (1999). In defense of private writing: Consequences for theory and research. *Written Communication, 16*(2), 139–170.

Fogal, G. G., & Koyama, D. (2022). A study of co–adaptation through journaling. *Journal of Second Language Writing, 55,* Article 100873. https://doi.org/10.1016/j.jslw.2022.100873

Lee, I. (2008). Fostering preservice reflection through response journals. *Teacher Education Quarterly*, *35*(1), 117–139.

Mlynarczyk, R. W. (1998). *Conversations of the mind: The uses of journal writing for second–language learners*. Lawrence Erlbaum Associates.

Pecorari, D. (2003). Good and original: Plagiarism and patchwriting in academic second language writing. *Journal of Second Language Writing, 12*(4), 317–345.

Peyton, J. K. (Ed.). (1990). *Students and teachers writing together: Perspectives on journal writing*. Teachers of English to Speakers of Other Languages.

Peyton, J. K., & Reed, L. (1990). *Dialogue journal writing with nonnative English speakers: A handbook for teachers*. Teachers of English to Speakers of Other Languages.

Peyton, J. K., & Staton, J. (Eds.). (1991). *Writing our lives: Reflections on dialogue journal writing with adults learning English*. Prentice–Hall Regents.

Peyton, J. K., & Staton, J. (Eds.). (1993). *Dialogue journals in the multi–lingual classroom: Building language fluency and writing skills through written interaction*. Ablex.

Vanett, L. & Jurich, D. (1990). The missing link: Connecting journal writing to academic writing. In J. K. Peyton (Ed.), *Students and teachers writing together: Perspectives on journal writing* (pp. 21–33). Teachers of English to Speakers of Other Languages.

Zafón, C. R. (2001). *La sombra del viento* [Shadow of the wind]. Planeta.

Reviews

Task-Based Language Teaching Theory and Practice. Rod Ellis, Peter Skehan, Shaofeng Li, Natsuko Shintani, and Craig Lambert. Cambridge University Press, 2020. xv + 417 pp. https://doi.org/10.1017/9781108643689

Reviewed by
Kathryn Akasaka
Morioka Shirayuri Gakuen

What is a task? Though tasks, naturally, lay at the core of task-based language teaching (TBLT), even experts in the field cannot agree on what exactly comprises a task. On the surface, this might seem problematic. From another viewpoint, however, the flexibility to interpret what a task consists of might broaden the appeal of TBLT and lead to greater diversity in its applications. As TBLT has become a mainstream approach to teaching a second or foreign language, there is a welcome addition to the literature with *Task-Based Language Teaching Theory and Practice*, from the prestigious Cambridge Applied Linguistics series. This volume provides a comprehensive and deep examination of TBLT by presenting unique perspectives from its collection of five authors.

This dense volume has three ambitious aims. The first is "to provide a broad-based and accessible state-of-the art account of TBLT by considering the pedagogical aspects of this approach and reviewing relevant theories and research" (p. xiii). However, I doubt if readers unfamiliar with TBLT would choose this as a first text to read on the subject, and I would certainly not recommend it as an introduction to TBLT as it is clearly not written with the novice reader in mind. Whilst it does provide an overview of the history behind TBLT, there is also an assumption that the reader will have a strong foundation in cognitive theory, which not every teacher has. For those lacking in this area, there is however an extensive bibliography to draw on to

help fill in any gaps. The second aim is "to examine the effectiveness of TBLT in relation to other mainstream approaches to language teaching" (p. xiii). TBLT has now been around long enough for there to be substantial research showcasing its effectiveness, and so the authors provide a number of studies comparing this approach with said mainstream approaches and methods. The third aim is "to examine the criticisms of TBLT that have been advanced by advocates of traditional language teaching and then to identify a number of 'real' issues that need to be addressed" (p. xiii). It is this final section where the book is at its best, as it tackles critics of TLBT head-on, while also looking carefully at critiques of the TBLT approach by researchers who, though not critics of the approach itself, have found valid points of concern in its theory and application. The authors divide the book into five sections: the "Introduction," "Theoretical Perspectives," "Pedagogical Perspectives," "Investigating Task-Based Programmes," and finally, "Moving Forward."

For readers new to the TBLT approach, the opening part provides a broad overview of the pedagogic background, starting by introducing the history and developments of the approach thus far, and concluding with a long list of questions about TBLT and some provisional answers. Part 2 concerns theory, and it contains five chapters, giving the reader a variety of lenses through which to view TBLT: "Cognitive-Interaction Perspectives," "Psycholinguistic Perspectives," "Sociocultural Perspectives," "Psychological Perspectives," and "Educational Perspectives." With several decades of second language acquisition having been dominated by cognitive theory, Chapter 4 on sociocultural perspectives is a welcome addition to the research as its absence has been noted in other seminal volumes on the topic. Drawing mainly on the work of Swain et al. (2011), this chapter stresses how task implementation can be used to mediate effective learning. The three major areas of graduated feedback, collaborative dialogue, and dynamic assessment are covered in detail. The chapter concludes with a note that pedagogy should involve both cognitive and sociocultural frameworks in order to provide a long-term vision of TBLT and decisions on task implementation.

Part 3 deals with pedagogical perspectives which are taken up in three chapters: "Task-Based Syllabus Design," "Methodology of Task-Based Language Teaching," and "Task-Based Testing and Assessment." Chapter 7 on "Task-Based Syllabus Design" provides a comparison and contrast of four perspectives on scope, task selection, and sequencing from researchers Prabhu, Long, Robinson, and Ellis. Further, how the approaches have been put into practice in universities in Japan using actual task-based syllabuses is introduced. Finally, the advantages and disadvantages of each are discussed,

including operational or illuminative syllabuses. It concludes by advocating Ellis's modular implementation as one that allows for more flexibility and better outcomes, than any single implementation would achieve on its own (p. 207). Chapter 9 on task-based testing and assessment will interest many educators, as assessment outcomes drive a major portion of the teaching industry. This chapter begins with an introduction to the basics of testing and the role of "ability for use" (p. 241) as a model for assessment. Four case studies are presented linking task-based research to assessment. For those invested in TBLT, it is a positive validation that what happens in the classroom can be successfully transferred to the testing arena.

Part 4 "Investigating Task-Based Programmes" has only two chapters: "Comparative Method Studies" and "Evaluating Task-Based Language Teaching." The first question addressed in the chapter is how effective task-based language teaching is compared to more traditional approaches. The authors note that studies comparing methods or approaches are challenging to design and often contain design flaws (p. 281). However, the research points to the benefits of a task-based program while calling for more studies comparing traditional approaches with TBLT. Chapter 11 illustrates how TBLT works in actual classrooms with both micro-evaluations (on particular lessons or tasks) and macro-evaluations (on whole courses) presented. These provide evidence for the effectiveness of TBLT in various teaching contexts, concluding that micro-evaluations are an especially effective way for teachers to better understand TBLT and implement it in their classrooms (p. 330).

The fifth and final section "Moving Forward" also has two chapters: "Responding to the Critics of Task-Based Language Teaching," and "Questions, Challenges, and the Future." In keeping with the theme of this book, this section explores the interface between research-based and pedagogical-orientated perspectives on TBLT. Presenting critiques from both insiders and outsiders, this "choir" of TBLT authors presents valid arguments for how to progress further with TBLT, openly detailing areas that require more research and discussing how to prepare for a future in which they envision TBLT being more frequently mandated by educational authorities.

I originally read this book as part of an online professional development course. One of the common challenges we all faced was the lack of a glossary. It is understandable, in that having one would make the text even longer than the hefty four hundred pages it stands at. However, many terms are not well defined, and an even more significant number are abbreviated in ways that can quickly get confusing. This is especially problematic if you are not reading the text from front to back, but flipping from place to place as

relevant situations occur in your context. Although the "Introduction" did an excellent job of outlining the history of TBLT, this volume is not targeted as the ideal first resource for those wanting to learn more about implementing TBLT as a teaching approach. The authors clearly state that it is not a "how-to" text, and certainly, those looking for practical ways to implement TBLT into their courses and lessons may not find many immediate answers in this book as they might in the introductory guide to teaching with TBLT from Willis and Willis (2007) or in an activity resource workbook such as the one published by Anderson and McCutcheon (2019). True to its aims however, this book does provide a wealth of information and insight into the creation of a task-based syllabus based on rationale and evidence. Japan-based teachers can also get insight on how the order and sequence of tasks in TBLT could align with recent educational directives from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT). As such, it might also prove invaluable for materials designers and curriculum creators. Well-organized and with a strong overall flow and cohesion, this dense and theoretical read is nevertheless comprehensible. The argumentation for and against aspects of TBLT by a variety of researchers in the book's final section was, for me, its most compelling aspect. *JALT Journal* readers who teach at the high school or tertiary level and are interested in a comprehensive text covering the theory of TBLT in-depth, with support from up-to-date research, will find lots to ponder in this book.

References

Anderson, N., & McCutcheon, N. (2019). *Activities for task-based learning: Integrating a fluency first approach into the ELT classroom.* Delta.

Swain, M., Kinnear, P., & Steinman, L. (2011). *Sociocultural theory in second language education: An introduction through narratives*. Multilingual Matters.

Willis, D., & Willis, J. (2007). *Doing task-based teaching.* Oxford University Press.

Classroom-Based Conversation Analytic Research: Theoretical and Applied Perspectives on Pedagogy. Silvia Kunitz, Numa Markee, and Olcay Sert (Eds.). Springer, 2021. x + 426 pp.

Reviewed by
Tim Greer
Kobe University

Although based in sociology, Conversation Analysis (CA) has become a cross-disciplinary approach that seeks to account for the ways people understand each other through interaction. CA is both an analytical method and a social theory (Heritage, 2008), and interventionist CA researchers are increasingly proposing ways to apply its findings to practical problems (Antaki, 2011). The contributors to this volume, *Classroom-Based Conversation Analytic Research*, edited by Silvia Kunitz, Numa Markee, and Olcay Sert, take up this aim in relation to the second language (L2) classroom.

The book is divided into four parts focusing on CA in relation to (a) L2 classrooms, (b) content-based language classrooms, (c) teacher education, and (d) assessment, with each comprising an overview from the editors and three or four chapters from leading researchers in the field of Conversation Analysis for Second Language Acquisition (CA-SLA). In line with the CA approach, the analysis is grounded in a detailed description of interactional practices that are made visible through transcripts of naturally occurring conversations. Each chapter also includes a section on the pedagogical implications of the research and each part begins with an overview from the editors, making the contributions of value to language teachers as well as specialist researchers. A fifth and final section provides two additional discussant chapters to conclude the volume.

Part I features four chapters on CA research in L2 classrooms, and a recurring theme throughout them is that learning involves not just language, but also other phenomena like the sequential context, the interactants' bodies, and the complex interactional ecology. Ali Reza Majlesi draws our attention to the way objects of learning emerge by and through multimodal interaction and how teachers and students jointly treat an unknown word as learnable. Similarly, Søren W. Eskildsen explores how one English learner's use of the word "can" is embedded within the recurring daily routines of the classroom and how embodied interactional resources enable him to expand the way he uses that word over time. The learner achieves this not through language

alone, but through a repertoire of visibly and audibly available means, with Eskildsen proposing that "the semiotic resource known as 'language' is a residual of social sense-making practices" (p. 71). Nigel Musk documents spelling-related repair practices among pairs of Swedish students writing an English composition on a computer. He demonstrates that various distributions of knowledge (between the partners or from the spellchecker) help make up an epistemic ecology that allows for opportunities for language learning within collaborative writing. Kunitz considers instruction-giving sequences in an Italian class, showing that the teacher provides detailed directions to the first group, but those instructions gradually become briefer as they are directed to other overhearing groups. One valuable take-away from this section is that language learning can happen while we are doing other things (i.e., while jointly undertaking actions and activities with others), and may not always conform to the teacher's lesson plan or agenda.

Part II uses the CA approach to account for interactional practices in content-based language classrooms. For example, two of the chapters look at interaction in CLIL classes: Natalia Evnitskaya explores facework and collaborative learning among Spanish primary school students learning Math in English, and Leila Kääntä provides a multimodal analysis of the way Finnish teachers accomplish English definitions during History and Physics classes. These chapters include a variety of multimodal resources in their analyses, including gestures, objects, and the students' first language. A third chapter by Yo-An Lee draws on data from content-based EFL classes at a university in South Korea. Going beyond the familiar IRF (Initiation-Response-Feedback) sequence, Lee tracks longer extracts of talk to show how teachers adjust their explanations based on unfurling evidence of the students' real-time understandings.

In Part III, the attention turns to some of the contributions CA can offer language teacher education. In his chapter, for example, Sert compares the way that various teachers respond to L1 use in their classrooms, then goes on to demonstrate a mobile application known as IMDAT that allows teachers to notate video recordings of their classes. Hansun Zhang Waring uses transcripts of a class to undertake an analysis of the various voices that teachers draw on and then offers a plan for using such extracts for teaching training. Finally, Younhee Kim and Rita Elaine Silver look at how post-observation conversations between expert and novice educators provide opportunities for reflection and feedback. Overall, this section offers some very practical pedagogical applications of the CA approach, not just for students but for teachers themselves.

The theme of Part IV is language assessment in L2 classrooms. It begins with a chapter by Nilüfer Can Daşkin, in which she traces the ongoing, contingent, and flexible nature of formulative assessment, illustrating assessment in practice through reference to a past learning event. The next two chapters explore assessment in relation to interactional competence (IC). Thorsten Huth outlines how interactional learning targets can be conceptualized within a language curriculum and F. Scott Walters puts forward some considerations regarding validation in CA-informed oral testing.

Finally, in Part V two preeminent scholars from the field reflect on each of the previous chapters and suggest future directions for CA-SLA. Junko Mori highlights the links between researchers and classroom practitioners, advocating for further CA investigations into pedagogical practices, for curriculum reform that recognizes the difference between language and interaction, and for a deeper understanding of the multilayered interactional ecology of the classroom. Likewise, Simona Pekarek Doehler calls for greater awareness of the notion of L2 interactional competence and outlines some crucial points of departure between it and the dominant view of language held in L2 education. She sees competence as jointly accomplished by both parties and therefore locally contingent and adapted in each instance of interaction. This reconceptualization (from language to interaction) calls for nothing short of a change to the central object of the field of Applied Linguistics.

CA was not originally intended as an approach to second language learning or teaching, but over the past 25 years, it has certainly become a significant research tool for accounting for social interaction in such settings. While this book focuses on language use in the classroom, a companion volume in the same series (Hellermann et al., 2019) employs CA to look at second language use "in the wild" and may also be of interest to readers of this later Kunitz et al. volume. This book offers an accessible overview of CA research in a range of different classrooms and will be of interest to both researchers and teachers.

References

Antaki, C. (Ed.). (2011). *Applied conversation analysis: Intervention and change in institutional talk*. Springer.

Hellermann, J., Eskildsen, S., Pekarek Doehler, S., & Piirainen Marsh, A. (Eds.). (2018). *Conversation analytic research on learning-in-action: The complex ecology of L2 interaction in the wild*. Springer.

Heritage, J. (2008). Conversation analysis as social theory. In B. Turner (Ed.), *The new Blackwell companion to social theory* (pp. 300-320). Wiley-Blackwell.

How Languages Are Learned (5th ed.). Patsy M. Lightbown and Nina Spada. Oxford University Press, 2021. xiii + 280 pp.

Reviewed by
Martin Hawkes
The University of Shiga Prefecture

This is the 5th edition of this popular title introducing the theoretical and empirical research background to second language acquisition and its relevance to language teaching. The first edition was published back in 1992, and I have been familiar with the title since its third iteration when I was informed of it by a colleague who was beginning a master's degree in TESOL. I then used it myself when I was doing my own Master's and found it to be a very readable and easy introduction to several key topics in the field. I feel its main strength is the accessibility in terms of technical language, concepts, and constructs. Because of this, when I was asked to start teaching *zemi* cohorts a few years ago, I chose this title to use with my students to help them build towards writing a graduation thesis. With the growth in English as a medium of instruction (EMI) courses in Japanese universities (Bradford, 2018), teachers of such classes may find this type of book to be highly appropriate.

How Languages Are Learned is divided into an Introduction, with an opening task, and seven main chapters. As with some other introductions to SLA (e.g., Hummel, 2021), Chapter 1 begins by looking at first language acquisition and describing some of the key theoretical approaches from the fields of psychology and linguistics including behaviourist, innatist, and interactionist perspectives. In Chapter 2, the authors shift toward second language acquisition and discuss key concepts such as developmental sequences and cross-linguistic influence. While the main focus here is on grammar, there are also sections covering the acquisition of vocabulary, pragmatics, and pronunciation. Chapter 3 identifies the key individual differences that can influence SLA success, including aptitude, motivation, and age. Chapter 4 outlines some of the main theoretical approaches to SLA from behaviourism through to sociocultural and complex dynamic systems theory. In Chapter 5, attention turns to the findings of empirical, mainly classroom, investigations into L2 learning and teaching. The book is very much designed to build towards Chapter 6, in which some of the most important and influential approaches to language teaching are described and appraised. Finally, Chapter 7 returns the reader to their responses to the opening task from the introduction and summarizes the key concepts and ideas of the book.

The book also contains activities and guidance for further study. Chapters 1 to 7 conclude with three to four useful suggestions for supplementary reading. Throughout the book, references to seminal studies allow the reader to explore and discover more about their specific areas of interest. There are also a few activities including the analysis of learner language and reflection on one's own experiences and preconceptions of L2 learning. In addition, each chapter closes with *Questions for reflection*, which might be useful for motivated solo readers but are more likely to be effective when taken up in a class as they are likely to stimulate discussions on diverse views. While there is not a large number of tasks contained within the book, it is possible to access more through the online resources (https://www.oup.com/elt/teacher/hlal). Among the most useful expansion contents are the 126 supplementary discussion questions, which are organized by chapter and sometimes sub-sections of a chapter. The ten extra activities provided online are also helpful, especially when the book is used as the main text for an introduction to SLA course.

As explained by Spada in one of the videos on the companion website, it is not the case that new material has simply been added to the 5th edition, but previous content from the older 4th edition – judged to have become less relevant – has been omitted. First of all, it must be said that while the revisions from the 4th edition are welcome, they are often rather modest and essentially woven into the existing structure of the book. There are no new chapters (not that I am contending any are necessarily needed) and section changes within the chapters are minimal. Mostly, the revisions involve additions, of varying length but typically brief, describing topics (e.g., translanguaging in bilinguals, complex dynamic systems theory, EMI) or pedagogical approaches and techniques (e.g., CLIL, TBLT, corrective feedback) that have attracted attention of late; the reorganization of some sections and paragraphs (e.g., sections on language aptitude and motivation have been reworked); and the updating of further reading suggestions (e.g., three recommendations in Chapter 2 are all new). There are very few sections that have actually been cut; however, one example where this has been done is with the discussion of intelligence as an individual difference determining L2 learning success, which had its own section in previous editions, but has now been eliminated.

One of the greatest benefits of the 4th edition for me was the Japanese-language version that was released alongside it. At the time of writing, it appears that there is no plan to release a Japanese version of this latest edition. For my purposes, I found that by having both the English and Japanese versions of the book, I was able to handle mixed-ability classes with more

ease. Those confident or more highly motivated students would tend to use the English version exclusively. However, less confident students could turn to the Japanese version to confirm their understanding of a given section or chapter in English, or perhaps read the Japanese version first so that they could more confidently tackle the English-language version and contribute more actively in tasks and discussions. It is slightly disappointing for me in my particular context that a Japanese version is seemingly not on its way.

For teachers wishing to build their knowledge of SLA research, or for graduate students of TESOL, this book continues to provide both an accessible introduction to the field and a springboard to reading further to explore specific areas of interest. I have primarily used the book as the main text for my seminar course with students at an English proficiency level usually between IELTS 5.5 and 6.0. Although the concepts can still be somewhat challenging to process in their second language, I have found the book to be quite suitable for my *zemi* students. Compared to other introductory SLA books, this title from the Oxford Handbooks for Language Teachers series is written in relatively simple language, and the extensive glossary of key terms is helpful for students and has also provided me with an additional resource for assessment. In addition, my students seem to have found the tasks (particularly the supplementary ones) and the discussion questions to be helpful and engaging. These have also lifted some of the burden off me when planning how to deliver the courses and approach each chapter.

Overall, *How Languages Are Learned* remains a comprehensive and reader-friendly introduction to second language acquisition and L2 teaching. Although the changes from the 4th edition are not particularly substantial, each chapter has been moderately revised and updated from the previous edition with a discussion of the most significant recent developments in research and examples of empirical studies. So, although it is certainly useful for anyone embarking on a graduate programme in TESOL, I would also recommend it for any university instructors who are teaching an introductory SLA or TESOL course, in a seminar, or perhaps an EMI course.

References

Bradford, A. (2018). It's not all about English! The problem of language foregrounding in English-medium programmes in Japan. *Journal of Multilingual and Multicultural Development, 40*(8), 707-720. https://doi.org/10.1080/01434632.2018.1551402

Hummel, K. M. (2021). *Introducing second language acquisition: Perspectives and practices* (2nd ed.). Wiley-Blackwell.

Pedagogical Stylistics in the 21st Century. Sonia Zyngier and Greg Watson (Eds.). Palgrave Macmillan, 2022. xxxii + 405 pp. https://doi.org/10.1007/978-3-030-83609-2

Reviewed by
Paul Hullah
Meiji Gakuin University

Emerging in the late 20th century, with the establishment of the Poetics and Linguistics Association (PALA) in Sheffield, England, in 1983, Pedagogical Stylistics (PS) has sought to activate and consolidate learner acquisition and deeper understanding of (especially literary) language worldwide via the tailored teaching of and research into discourse analysis (PALA, 2019). Dedicated to PALA founding member Ronald Carter (1947-2018), who transformed and elevated the study of literary linguistics by championing and systematizing the linguistically-informed interpretation of literary texts, the opening sentences of the "Preface" to this new collection make substantial claims for PS and for seeking silver linings in our COVID-clouded recent past: "Never have Ron Carter's words ["*seeing new horizons is always the hardest part of the journey*": the section's epigraph] been more pertinent than now. Just as this volume had been commissioned, the world came to a standstill, so new ways of learning, working and communicating had to be devised" (p. vii).

These assertions are worth considering, thence upholding, since knowing and re-evaluating the realities of the past is an effective yet seldom discussed go-to procedure along any self-developing educator's journey to consistent classroom success. The latter contention—that historical and technological progress necessitates novel modes of human interaction—is indisputable, though whether new means better is debatable. The former claim, with its insinuation that Carterian PS (gleaning insightful communicative praxis from the literary by "using linguistic description to substantiate textual interpretation" (p. viii)) remains relevant and inspirational today, and is what the variously informed but unanimously pragmatic chapters of this book seek to confirm and celebrate. The editors have been here before: this intelligently plotted book critically revisits and updates their equally ambitious *Literature and Stylistics for Language Learners* (Watson & Zyngier, 2007), and is divided into four sections.

Part I "Assessing and Broadening the Scope of Pedagogical Stylistics" instructively illustrates the current landscape. Geoff Hall (Chapter 1) sum-

marizes recent and ongoing research and lists essential literature, paving the way for Violeta Sotirova (Chapter 2) to delineate the traditional and residual disconnects and frictions between linguistic and literary criticism in a chapter both informative and playful, syntactically parsing a Pound poem by means of the skillful, linguistic-critical prestidigitation that was Carter's own forte. In Chapter 3, David I. Hanauer explores the efficacy of course-based research (CURE) on learner agency, demonstrating how it stimulates meaningful student engagement with socio-political issues, and in Chapter 4, Marcus Bridle and Dan McIntyre also advocate learner empowerment, reporting Japanese EAP students' heightened alertness to linguistic appropriacy as a direct result of hands-on application of principles drawn from corpus stylistics.

Part II is titled "Cognitive Perspectives." Analyzing Shakespeare and Waugh, via Barthes and Jerome McGann, Peter Stockwell proposes an experiential "readerly" (p. 107) notion of textuality in Chapter 5 "The Principle of Moments" (*moments* plural; not singular, as it is mis-rendered in the "Preface"). His essay offers theoretical and rhetorical redefinitions of what a "text" is and is not, and should be compulsory reading for educators hoping to use literature (poetry in particular) effectively in EFL. Stockwell proposes measured introspection as a "schematic of the textual moment … [which] can serve as an account of textual momentum in readerly experience" (p. 120). Marcello Giovanelli and Chloe Harrison (Chapter 6) next reclaim the intuitive aspect of Cognitive Grammar in an illuminating trio of case studies that finds them workshopping ideas, methodologies, and "different models of grammars that have been used in stylistics; and some overarching principles of cognitive grammar" previously outlined in their *Cognitive Grammar in Stylistics: A Practical Guide* (Giovanelli & Harrison, 2018, p. 1). Like the previous chapter, but here practically rather than theoretically, Giovanelli and Harrison's project suggests that *redefining* — carefully bridging the disconnect between outmoded negative preconceptions and the newly repurposed usefulness of PS to teacher, student, or reader, thereby decompartmentalizing language and literature — is key to this discipline's continued development (and can be said to be the main theme of this book). Keeping an eye on conceptualization, in Chapter 7 "A Text-World Pedagogy for Young Stylisticians," Ian Cushing examines Text World Theory — the notion that language is processed and understood based on personalized imaging schemata: "mental representations, or text-worlds" (Gavins, 2007, p. 2) — explaining measurable benefits of this undersubscribed cognitive-stylistic pedagogic modality to teachers in training. In Chapter 8, Esmeralda

V. Bon and Michael Burke assess the effects that the analogue-to-digital shift has had on situational and cognitive aspects of reading. Their study discovers that contemporary "literary reading behavior is more traditional than assumed" (p. 184). As with all the best parts of this book, a logical argument lucidly expounded here makes for a pleasant, persuasive, and productive reader experience.

Frank Hakemulder opens "Part III: Reader Engagement and Feelings" with a solid, empirically-driven exposition in his chapter subtitled "Reader Response Research in the Classroom," encouraging intersectional envisioning of classroom dynamics to promote EFL learner self-awareness. Anna Chesnokova and Sonia Zyngier (Chapter 9) then offer pragmatic comparative cultural reflections on usage of poems in translation, in a smooth fusion of literary and linguistic techniques. In Chapter 10 "Teachers' Intertextual Identities and English Education," Jessica Mason ruminates on identity and applies an intersectional frame to reader sense of self. This part of the volume showcases dedicated teachers who are committed to student empowerment, and the book is better for it.

Part IV addresses "Innovations in the Educational Setting" and contains two out of the three chapters of the collection's main 14 that specifically discuss EFL in Japan. The exception, Jane Spiro's evaluative survey of patterns and developments in tertiary teaching of academic writing, concludes in recommending adapted discourse awareness (via peer reviewing) as fertile ground for flexible educators wanting to "enable writers to acquire ownership of their target writing community" (p. 290). Paul Sevigny (Chapter 13) methodically but innovatively evaluates Japanese EFL literature circles in terms of collaboration, relevance, evidence, and alignment, deducing from his findings that instructor-instigated extension of role-based circles into spaces beyond the classroom generates "gains in self-efficacy for all" (p. 339). Finally, in Chapter 14, Azumi Yoshida, Masayuki Teranishi, Takayuki Nishihara, and Masako Nasu controversially, but convincingly via meticulous qualitative cross-linguistic stylistic analysis of EFL learner written output, argue that evident correlation between L1 (Japanese) and L2 (English) proficiency confirms the importance of a structured primary duo-lingual education (with emphasis on reading) for optimum subsequent English proficiency. Admitting the limitations of their relatively modest study, these researchers nevertheless offer provocative and engaging insight.

There is much to admire in this anthology. Most major aspects of PS are serviceably tackled: contributors address corpus stylistics, creative writing, literary-linguistic criticism, student research, critical discourse, cognitive

and multimodal stylistics, classroom discourse, literary language, and more, including "the virtual world, and … how pedagogical stylistics can promote political and social awareness" (p. ix). For the *JALT Journal* readership, EFL-related considerations surface usefully throughout. Though careless lapses such as "language patterns, especially those that turn into verbal art" (p. viii) and "the scope of this book is quite comprehensive in terms of contexts" (p. ix) infuriate, arguments are clearly articulated and studies precisely described throughout. Minor gripes? Co-authors of joint-authored books are occasionally omitted in biodata: unforgivable in academic work. "Keywords" (which some chapters include, but others do not) seem a hasty afterthought: often woefully redundant (e.g., "language acquisition"), yet, in a book about Pedagogical Stylistics, "pedagogical stylistics" is listed in "Keywords" for only 4 of 15 chapters.

Given the task of writing the "Afterword," Michael Toolan asks: "Where might—or *should*—pedagogical stylistics go from here, in the next few decades? How might it adapt and thrive, and what might be its priorities? The editors of this volume have invited me to use this space to give my views" (p. 372). Surely the editors' own predictions at such a juncture would have been more pertinent and more aptly satisfied Carter's original vision? Nonetheless, this valuable book goes a good way towards answering, from various angles, the question posed by Michael Burke (2010) in his seminal paper "Why Care about Pedagogical Stylistics?" and, for that reason alone, it merits attention.

References

Burke, M. (2010). Why care about pedagogical stylistics? *Language and Literature: International Journal of Stylistics, 19*(1), 7-11. https://doi.org/10.1177/0963947009356805

Gavins, J. (2007). *Text world theory: An introduction.* Edinburgh University Press. https://doi.org/10.3366/edinburgh/9780748622993.001.0001

Giovanelli, M., & Harrison, C. (2018). *Cognitive grammar in stylistics: A practical guide.* Bloomsbury. https://doi.org/10.5040/9781474298957

PALA. (2019). *Constitution of The Poetics and Linguistics Association (PALA).* https://www.pala.ac.uk/uploads/2/5/1/0/25105678/pala_constitution.pdf

Watson, G., & Zyngier, S. (2007). *Literature and stylistics for language learners: Theory and practice.* Palgrave Macmillan. https://doi.org/10.1057/9780230624856

Language Teacher Noticing in Tasks. Daniel O. Jackson. Multilingual Matters, 2021. xii + 192 pp. https://doi.org/10.21832/9781800411241

Reviewed by
Martyn McGettigan
Hiroshima Shudo University

The concept of noticing will be familiar to anyone with even the most cursory knowledge of SLA literature. The term was first used by Schmidt (1990) in his "noticing hypothesis", and forms a central component of Skehan's (2002) model of language aptitude. In this accessible, evidence-based study, Jackson aims to give an account of another form of noticing: language teacher noticing. Jackson defines language teacher noticing as "a form of reflection entailing processes of attending to events, interpreting them, and deciding how to act on them, which occurs during engagement with learners" (Chapter 1, para. 2). As he points out, noticing, as opposed to reflection or cognition, has been a neglected area in the SLA literature. He argues for the need to address this gap in the literature on the grounds that a deeper understanding of language teacher noticing can help develop rapport, aid acquisition, foster participation and reflection, and guide observation.

Jackson sets out his five research questions, looking at whether task complexity or perspectival memory affects recall of teacher noticing, what teachers notice about students' use of embodied and verbal resources, and finally, as this is a mixed methods study, how quantitative and qualitative data can be integrated to form an understanding of language teacher noticing. In the remainder of the first part of the book, Jackson gives an overview of the literature on teacher noticing in general and language teacher noticing in particular (including literature from education fields other than language teaching). Jackson skilfully brings together data and insights from various disciplines, forming an account of the key concepts underlying his study. This coverage is thorough and clearly structured to be accessible to the general reader while remaining insightful and instructive to anyone already well-versed in the areas he draws on.

In Part 2, which forms the largest part of the book, Jackson gives a detailed account of his study. He begins by outlining the context and describing the study design, goes on to give the quantitative results in full and an overview of the qualitative results, and ends the section with an analysis of the study's

findings. The study was conducted at a private university which, judging by Jackson's description, seems to be an institution that places a great deal of emphasis on peer tutoring and other forms of collaborative learning. The participants were pre-service English teachers on a teacher education programme. The study involved pairs of these pre-service teachers taking the roles of both teacher and student, and the participant in the teacher role leading the participant in the student role through a map-based task (provided in Appendix A). This was followed by the teacher participant watching a recording of the task and commenting on what they remembered noticing at different stages. Both simple and complex tasks were used, and the recordings were made in both field (i.e., from the view behind the teacher) and observer (from behind the student) perspectives. In this way, both quantitative and qualitative data were generated. Quantitative data was produced by setting certain criteria for what constituted incidents of noticing and then totalling the number of these in each task iteration. Qualitative data, on the other hand, was derived from the content of the comments made by the participants. Jackson argues that this "embedded" (Chapter 5, para. 16) design, with both qualitative and quantitative data collected from the same source, allowed for a richer understanding of the study's findings.

Jackson appears to have been somewhat disappointed with his study's quantitative findings. Many of the results that initially looked meaningful were found not to be statistically significant. In fact, only the effect of task complexity on recall was found to be significant, with complexity having no significant effect on noticing and perspective having no effect on either recall or noticing. Despite this, as Jackson points out, the results raise interesting questions. The results suggested, for example, that there might be a greater effect for perspective in more complex tasks, and this is something Jackson suggests would benefit from further research. Despite these weaker-than-hoped-for results, Jackson's account of his study's design offers many insights that would be useful to novice researchers or those taking up a mixed methods study in this or related areas. It is clear that a great deal of thought went into task design and how to code for noticing, to ensure that the stimulated recall gave as true a picture as possible of what was noticed. Even where mistakes were made, Jackson admits to them, and subsequently, the reader should be able to learn from them also.

It is in the qualitative results that Jackson's study really comes into its own. Jackson first looks at what the pre-service teachers noticed in terms of "embodied resources" (Chapter 7, para. 3) (e.g., gestures, facial expressions). He gives a detailed analysis of several excerpts from his transcripts (which

included detailed annotation covering these embodied resources). His analysis provides fascinating insight into the role that embodied resources play in student-teacher interaction. Jackson also raises the issue of teachers' sensitivity to embodied resources as an area where further study could be beneficial. This is followed by his account of what the teachers noticed in terms of "verbal resources" (Chapter 8, para. 2). Here, too, a number of well-selected excerpts are provided by Jackson which illustrate commonly occurring themes. In particular, his findings led to a fascinating reflection on the difficulties faced by teachers facilitating tasks in progress, and the concerns over their choices in terms of when and how to intervene. This section of the book is likely to be very useful to teachers at any stage of their career, from pre-service to highly experienced. The situations and issues noticed (and not noticed) will resonate as familiar, and should serve to encourage the teacher to reflect on their own noticing in class.

In the final part of the book Jackson looks to the future, suggesting a number of areas for further research, and also giving his views on how language teacher noticing research can be applied in language teacher education and training. For those looking for new areas of research, there is ample inspiration here for potential research projects in an exciting, emerging field. There is also a great deal that should be of interest to those involved in language teacher education, whether as future teachers or their instructors. For this reviewer, apart from the ideas for possible future projects, the main takeaway was the opportunities for reflection occasioned by the in-depth analysis into episodes of interaction between student and teacher. Although always familiar, these interactions tend to be the sort of thing that teachers may take for granted, or perhaps have their own, almost unconscious ways of dealing with. This reviewer noticed his teaching and practice in these areas challenged, and believes he grew as a teacher as a result.

References

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11(*2), 129-158. https://doi.org/10.1093/applin/11.2.129

Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 69-94). John Benjamins. https://doi.org/10.1075/lllt.2.06ske

Bridging the Humor Barrier: Humor Competency Training in English Language Teaching. John Rucynski Jr. and Caleb Prichard (Eds.). Lexington Books, 2020. xxi + 275 pp.

Reviewed by
Cathrine-Mette Mork
Miyazaki International College

Bridging the Humor Barrier: Humor Competency Training in English Language Teaching, edited by John Rucynski Jr. and Caleb Prichard, came into my hands in 2021 at an opportune time. I had just finished skimming through Bell and Pomerantz's (2016) *Humor in the Classroom: A Guide for Language Teachers and Educational Researchers*, which although interesting, did not prove directly helpful to a project I had been contemplating starting: an educational video explaining sarcasm in the English language to non-native English speakers. I had received several questions about this type of humor from students and had encountered enough personal experiences in the classroom relating to my intentional use of sarcasm going unrecognized, misunderstood, or unappreciated, that I thought it potentially useful to create a video demystifying this type of humor. Propitiously, one of the JALT YouTube channels published a video presentation created by Rucynski and Prichard. The editors then hosted a follow-up Q&A session to their video for an online JALT conference. After keenly digesting their presentation and talk, I immediately purchased their book, *Bridging the Humor Barrier.* Chapter 7, co-authored by the editors themselves and titled "Humor Competency Training for Sarcasm and Jocularity," included exactly the information I had been searching for.

After completing my video project (Mork, 2021), I sat down to read the remainder of the book, which includes three chapters in Part 1 on humor competence development outside the class, three chapters in Part 2 on integrated humor instruction, and four chapters for Part 3, on explicit humor competency training, two of which were written by the editors.

In the introductory section, Rucynski and Prichard state that although there is an expanding body of research on humor in English language instruction, much of the concern has been with the use of humor as a tool, rather than as an objective (p. viii). They explain that humor competence requires several steps: recognizing, comprehending, appreciating, responding, and producing. They then outline six guidelines useful for teachers in

implementing humor competency training (p. xii – xiv) and offer seven key areas for research into humor competency (p. xv).

Anne Pomerantz opens Part 1 with a chapter on preparing language learners to not only interpret humorous discourse in a second language, but also to "position learners as creators and enactors of interactional humor" (p. 2). Pulling from research literature dealing with descriptions of learners using humor of their own accord, in addition to international comedian Gad Elmaleh's successes and failures in performing for different linguistic and cultural groups, she posits that four areas of humor production are important: (a) creating, establishing, and checking for shared knowledge; (b) playing to and with expectations; (c) making intentional repertoire choices; and (d) focusing on pronunciation and timing in the delivery. She also delves into the intersectionality of humor competence and intercultural communicative competence.

Jules Winchester (Chapter 2) discusses the role of sociopragmatic knowledge when expressing humor in intercultural contexts. He reminds us that in addition to having language knowledge, we need to be social and cultural to successfully comprehend and produce humorous discourse.

In Chapter 3, Maria Ramirez de Arellano relates second language humor competency development to cross-cultural adaptation, especially emotional adaptation. Her findings show a close interrelation between the communicative, social, and psychological effects brought about through humor in intercultural interactions.

Mohammad Ali Heidari-Shahreza opens Part 2 on humor instruction with a chapter on teaching with and about humor. His research-supported classroom experience prompted him to coin the term *humor-integrated language learning* (HILL), a methodology that he positions as an attempt to improve learners' humor competence and L2 proficiency.

Scott Gardner describes the interactional humor (humor intentionally spoken by a character in a dialogue) found in 32 English language textbooks targeting learners in Japan, Greece, and global audiences at the junior high school level. He shares the underlying reasons for the inclusion of interactional humor, argues that conversational humor is cultural and a valuable pragmatic resource despite the often-contrived nature of the dialogues.

The reading of jokes in English is the topic of Nadezda Pimenova's chapter, in which she explains how English language learners understand and appreciate humor. After outlining two theories of verbal humor, she explains how the cultural background and L2 proficiency of the participants in two studies she conducted are related to humor comprehension and appreciation.

Part 3 of the book was the most interesting to me, as it deals with explicit humor training for competency. Prichard and Rucynski (Chapter 7) describe how they go about training Japanese students to recognize, understand, and appreciate sarcasm (irony) and jocularity in English. They recommend avoiding, or at least putting off, training learners to produce sarcasm, since it might easily have harmful effects if not conducted well.

Richard Hodson follows with a summary of three studies he conducted which contain four pedagogical elements he deems important in the teaching and learning of humor competence: (a) familiarity with theories of humor, (b) cultural content knowledge, (c) the provision of input models, and (d) the provision of output opportunities. For instructors with little time, confidence, and materials for humor competency training, he recommends single-skills training such as structured activities with model examples (input) that allow learners to create and evaluate their own humorous output.

In Chapter 9, Maria Petkova suggests the use of diaries to research and develop second-language humor competence, since journaling, among several other benefits, "can encourage more and deeper reflection on past humor or humor attempts with interlocutors and in the media" (p. 219). Her study made use of a curriculum that included humor diaries as an instructional tool in addition to being a data collection method. After writing in response to prompts about humor for eight weeks in addition to receiving humor competency instruction, quantitative results showed definitive gains in humor comprehension and appreciation.

The final chapter by Rucynski and Prichard focuses on training Japanese learners to recognize satirical news. This comprised the second part of their JALT conference presentation mentioned earlier, the first being about sarcasm recognition. The authors argue that the ability to identify satire is important for improved media literacy and critical thinking, as well as to have a better understanding of the cultural bases for humor in the target language. This is especially true in Japan, where humor in the form of *manzai* (stand-up comedy duos) and *rakugo* (solo storytelling) is popular, but satire is less so. The authors shared the results of a study involving a tool they developed to test the extent to which Japanese learners of English can identify satire, as well as the results of an experiment testing the efficacy of implementing humor training into a reading course.

This publication will be of obvious interest to humor scholars and interculturalists, but also to any second or foreign language instructors who have ever asked themselves, "Why aren't they laughing?" after cracking a joke, making a snide quip, or joshing around—all in the simple hope of deliver-

ing instruction or guidance in a more enjoyable way. Through this volume, second language educators might become inspired and better equipped to integrate humor competence training into their own teaching.

References

Bell, N., & Pomerantz, A. (2016). *Humor in the classroom: A guide for language teachers and educational researchers.* Routledge. https://doi.org/10.4324/9780203082690

Mork, C. [MORKSENSEI]. (2021, June 27). *What's so funny? How to understand and respond to spoken sarcasm in English*. [Video]. YouTube. https://youtu.be/uRyAhXdqpMM

JALT Journal Aims and Scope

JALT Journal is a bi-annual, Scopus-approved research journal of the Japan Association for Language Teaching (全国語学教育学会). JALT's larger mission is to support the research programs and professional development of JALT members, promote excellence in language learning, teaching, and research, and provide opportunities for those involved in language education. In line with this mission, *JALT Journal* publishes high-quality English- and Japanese-language, quantitative and qualitative, theoretically-informed and empirically-grounded studies of relevance to second/foreign language education in Japan. Although emphasis is placed on the Japanese context, *JALT Journal* values contributions which also transcend geographical boundaries to illuminate the complex interaction between language, language use, people, education, and society across cultural and socio-political contexts.

When possible, submissions to *JALT Journal* should aim to be both descriptive (*What is my data?*) and explanatory (*Why is my data like this and not otherwise?*) in purpose, and further stimulate scholarly debate, to hopefully improve existing applied linguistic scholarship around the world. Areas of interest include but are not limited to the following:

- Bilingualism and multilingualism
- Classroom-based language education
- Cognitive linguistics
- Contrastive linguistics
- Conversation/discourse/critical discourse analysis
- Critical language pedagogy
- Curriculum design and teaching methods
- Intercultural communicative competence
- Language acquisition/learning
- Language policy and planning
- Language testing/evaluation
- Phonetics
- Pragmatics
- Psycholinguistics
- Semantics
- Sociolinguistics
- Syntax
- Teacher training
- Translation and interpretation
- Vocabulary

Guidelines

Style

Authors are encouraged to submit manuscripts in five categories: (1) full-length articles, (2) short research reports (*Research Forums*), (3) theory-grounded essays which may include analysis of primary or secondary data (*Perspectives*), (4) comments on previously published *JALT Journal* articles (*Point-to-Point*), and (5) book/media reviews (*Reviews*) either requested by the Book Reviews editor or suggested by the author. Articles should be written for a general audience of language educators; therefore, statistical techniques and specialized terminologies must be clearly explained and their use clearly justified.

Authors are responsible for obtaining permissions for any copyrighted material included in their manuscript submission. When submitting a manuscript based on a thesis or dissertation published in an institutional repository, authors are responsible for checking with the copyright holder for permission to publish portions of the original text.

Style

JALT Journal follows the *Publication Manual of the American Psychological Association, 7th edition* (available from <https://apastyle.apa.org/products/publication-manual-7th-edition>). A downloadable copy of the *JALT Journal* style sheet is also available on our website at <https://jalt-publications.org/jj/>. Instructions for online submissions can also be found at this website.

Format

Full-length articles must not be more than 8,000 words, including references, notes, tables, and figures. *Research Forum* submissions should not be more than 4,000 words. *Perspectives* submissions should not be more than 5,000 words. *Point-to-Point* comments on previously published articles should not be more than 1,000 words in length, and *Reviews* should generally be around 1,000 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 and for obtaining any permissions for copyrighted material contained in the manuscript.

Submission Procedure

Please submit the following two documents in MS Word format to the appropriate editor indicated below:

1. Cover sheet with the title and author name(s), affiliation(s), and contact information of the corresponding author.
2. A blinded manuscript, including title, abstract, and keywords, with no citations referring to the author. Do not use running heads and do not include your biographical information. Follow the *JALT Journal* style sheet.

If the manuscript is accepted for publication, a Japanese translation of the abstract will be required. Authors will also be asked to provide biographical information. Insert all tables and figures in the manuscript. Do not send them as separate files.

Submissions will be acknowledged within 1 month of their receipt. All manuscripts are first reviewed by the 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. The first round of review 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. Plagiarism, including self-plagiarism, will result in articles not being published or being retracted and may also result in the author(s) being banned from submitting to any JALT publication. For further information, see the JALT Publications Statement of Ethics and Malpractice at: https://jalt-publications.org/jalt-publications-statement-ethics-and-malpractice

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

Please upload submissions in these categories to the following website:

https://jalt-publications.org/content/index.php/jj

Manuscripts should follow the **American Psychological Association (APA) 7th Edition style**. Please indicate if your submission is a 1) Full-length article; 2) *Perspectives* article; 3) *Research Forum* article or 4) *Point-to-Point* submission.

For any general inquiries about English-language submissions, please contact:

Dennis Koyama, *JALT Journal* Editor
jaltpubs.jj.ed@jalt.org

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 (following page) for details. Please send Japanese-language manuscripts to:

Kiwamu Kasahara, *JALT Journal* Japanese-Language Editor
jaltpubs.jj.ed.j@jalt.org

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* and can be found online in each issue at <https://jalt-publications.org/tlt/>. Review authors receive one copy of the *Journal*. Please send submissions, queries, or requests for books, materials, and review guidelines to:

Greg Rouault, *JALT Journal* Reviews Editor
jaltpubs.jj.reviews@jalt.org

Special Issues

Special issues often make an important contribution to the development of academic discourse in a specific field, because they allow researchers and practitioners to (a) identify an issue or topic of particular relevance to the context in which the journal is read, (b) summarize the key concepts and debates shaping that issue, (c) bring further sophistication to existing academic discourse and identify new research possibilities, and (d) identify key readings for the journal readership. Special issues can also attract new authors and readers to an academic journal, and can be an effective means of finding new editors for that journal.

We strongly encourage *JALT Journal* readers to submit special-issues proposals. When submitting such proposals, please make sure that they adhere to the aims and scope of *JALT Journal*. Proposals should include: (1) a title which clearly captures the special issue topic, (2) a brief description of the special issue, (3) an account of the motivation behind the special issue and its importance to the field at large, (4) a list of no more than three guest editors with short biographical information, including editorial work experience, and (5) a list of article contributors, with a short description of each article contribution.

Inquiries about Subscriptions, Ordering *JALT Journal*, or Advertising

JALT Central Office
Marunouchi Trust Tower Main Building 20F
1-8-3 Marunouchi, Chiyoda-ku, Tokyo 100-0005 JAPAN
Tel.: (+81) 3-5288-5443
Email: jco@jalt.org URL: https://jalt.org

日本語論文投稿要領

*JALT Journal*では日本語で執筆された（a）論文、（b）研究報告、（c）展望論文、（d）*JALT Journal*に掲載された著作物へのコメント・考察、（e）書評を募集しています。（a）論文と（b）研究報告の違いは、以下の通り字数制限による違いです。（c）展望論文は、言語教育研究に関する課題に焦点をあてた短い論文で、先行研究の検証、理論や1次2次データに基づく議論などを含むものです。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*（7th edition）の定める方式に合わせて下さい。*JALT Journal*書式シート（日本語原稿用）を以下からダウンロードできます<https://jalt-publications.org/jj/>。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、（a）論文は25,000字、（b）研究報告は13,000字、（c）望論文は16,000字、（d）*JALT Journal*に掲載された著作物へのコメント・考察は2,000字、（e）書評は1,500~3,000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:
*JALT Journal*書式シート（日本語原稿用）を参考に作成の上、電子メールの添付書類でお送りください。　なお、上記（a）論文~（e）書評のどのカテゴリーへの投稿かを明記ください。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願います。

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:

070-8621 北海道旭川市北門町9丁目　北海道教育大学旭川校
JALT Journal 日本語編集者　印南　洋 笠原　究
電話: 0166-59-1273
jaltpubs.jj.ed.j@jalt.org

JALT Journal 第 45 巻　第 1号

2023年 4月20日　印刷
2023年 5月1日　発行
編集人　小山デニス
発行人　クレア・カーネーコー
発行所　全国語学教育学会事務局
〒100-0005 東京都千代田区丸の内1-8-3 丸の内トラストタワー本館20階
TEL（03) 5288 5443
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

LOOKING FORWARD

Sustainable futures in language education

May 12-14, 2023, Kyoto Sangyo University

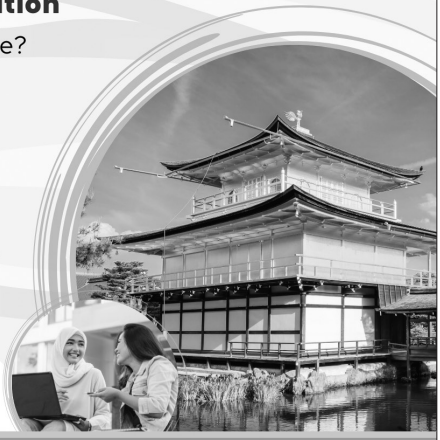
PanSIG

Building the future of language education

How will language learning communities evolve?
Let's get together and think it through,
in Kyoto, the cultural capital of Japan.

古都 京都で、語学教育の未来へ

言語学習コミュニティはどう進化するのでしょうか。
今こそ、皆と交流しながら考えましょう。



VISIT THE WEBSITE AND SNS FOR MORE INFO

www.pansig.org JALTpansig @pansig2023 @JALT_PanSIG

CALLING THE FUTURE JALTCALL 2023

JALTCALL 2023 Conference: Special invited Friday night workshops

Dr. Olena Prysiazhna
visiting professor from Ukraine

Brought to you by Sojo University
Workshop contents:
Ukrainian history and culture, followed by a Ukrainian CALL lesson.

Alexandra Burke
3x Best of JALT winner

Brought to you by BRAIN SIG and CUE SIG
Workshop contents:
Using the Immersive reader and other text-to-speech tools to bring text to life for readers of all levels, increasing accessibility.

Conference dates: June 2 - 4. F-2-F in Kumamoto with online attendance option.
https://jaltcall.org/jaltcall2023/