

# 外国語教育研究ジャーナル

JOURNAL OF FOREIGN LANGUAGE EDUCATION AND RESEARCH

第4巻

2023年12月



Vol.4

December 2023



立教大学 外国語教育研究センター

RIKKYO UNIVERSITY

Center for Foreign Language Education and Research

## 目次 / Table of Contents

### 【研究論文・Research Articles】

Prompt-Related Challenges:

Using a Mobile Application to Measure Spoken Fluency  
..... Heather Woodward & Joshua Rappeneker 1

Content, Language, and Rock n' Roll:

Finding the Harmony Between Content and Language in a CLIL Course  
..... Kevin Thomas 28

To What Extent Does Form-Based Priming Account for the Mnemonic Effect  
of Phonological Patterns? A Preliminary Investigation

..... Michael Green 47

実例でみる「SV了〈我〉時間」構文

..... 李 菲 66

### 【研究報告・Research Briefs】

Enhancing Receptivity in EFL Classrooms:

Exploring Teacher Development Opportunities Through Follow-up Reflective Diaries  
..... Anna Belobrovy 79

Incorporating the Elements of Fiction Into a CLIL-Based Extensive Reading Course

..... Michael Peragine & Christopher Mattson 88

English for Creative Purposes:

Implementing a CLIL Framework for Tertiary Creative Writing Courses in Japan  
..... Paul Curran 100

投稿規程・Manuscript Submission Guidelines ..... 111

執筆者・Authors ..... 119



# Prompt-Related Challenges: Using a Mobile Application to Measure Spoken Fluency

Heather Woodward

Joshua Rappeneker

---

## Abstract

The language-based justification for Rikkyo University's English discussion class (EDC) is to build second language (L2) spoken fluency. However, instructors lack a quantitative and objective method for monitoring their students' spoken fluency. Thus, assessing the EDC's ability to increase L2 spoken fluency is challenging. One possible solution is to develop a mobile application that measures students' in-class L2 spoken fluency in terms of speech rate, pause frequency, and peer-based perceived fluency during the 3/2/1 activity, a modified version of the 4/3/2. When testing our mobile application in class, one issue concerns the textbook prompts for the 3/2/1 activity. In this paper, we investigate one participant's spoken fluency for six weeks to exemplify the prompt-based challenges, which include a) a lack of topic repetition, b) no strategy training, and c) an under-utilization of formulaic utterances presented in the textbook. Possible solutions include recycling previous prompts, adding strategy training to the mobile application features, removing the third prompt, and embedding target formulaic utterances. Implementing these changes to the textbook might mitigate some issues experienced by mobile application users.

**Keywords:** *Discussion, Speaking, 4/3/2*

## Introduction

There has been a growing interest in spoken performance assessment in the field of second language (L2) development (Ogawa, 2022). During the speaking part of English proficiency exams such as the International English Language Testing System, candidates deliver opinion-based monologues. Spoken fluency, as defined by the speed and ease of speech, accounts for a significant portion of the variance in human ratings of these opinion-based monologues, whereas accuracy and complexity account for only a small amount (Ogawa, 2022). Students at Rikkyo University desire to live abroad, work for international companies, and make foreign friends, and by extension, they wish to build L2 spoken fluency to meet the demands of real-world communication (Hurling, 2012). Throughout the Spring 2023 semester, we have been conducting tests on a beta version of our mobile application

designed to measure spoken fluency in classroom settings during their 3/2/1 activity. The activity comprises speakers delivering the same monologue three times to three partners for three minutes, two minutes, and one minute (Appendix A). The activity's conditions of time constraint and content repetition help speakers increase their utterance fluency, which, theoretically, streamlines their L2 speaking process.

English discussion class (EDC) instructors implement the activity for at least 13 of the 14 weeks. In the EDC textbook, there are three prompts for the activity. The first two prompts are always opinion-based or experience-based questions (Appendix B). The third prompt includes sharing three interesting ideas or facts from the reading passage (Sturges et al., 2023), which refers to the passage assigned each week as homework; it is designed to introduce the topic for the next lesson (Hurling, 2012). As classes consist of 11 students, the instructors participate in the activity as well; therefore, they might not notice the strategies implemented by the students. For example, participants seem to approach the third prompt (i.e., share three ideas or facts from the EDC textbook's reading passage) with differing strategies, one of which is reading aloud from the passage in their textbooks. However, reading aloud inflates their fluency scores on the mobile application. Having accurate representations of their current performance is important for a) monitoring students' progress, b) providing useful instructions for improvement, and c) setting realistic spoken fluency goals. Thus, our objective is to investigate participants' transcripts and data to uncover issues and present potential solutions. In the following literature review, we review the pivotal studies on and theories for L2 spoken fluency development.

## Literature Review

### Earlier 4/3/2 Studies

Early studies have manipulated the 4/3/2 activity's conditions of content repetition and time constraint to investigate the efficacy of 4/3/2. This activity, similar to 3/2/1, consists of speakers delivering the same monologue thrice to three different partners, but the duration is four, three, and two minutes (Maurice, 1983). In de Jong and Perfetti's (2011) study on content repetition, participants who repeat the same content for all three rounds retain spoken fluency gains in a delayed posttest four weeks after the activity, whereas the participants who deliver monologues on different content do not. They suggest that participants who repeat language proceduralize the linguistic knowledge, thereby shifting their fundamental cognitive processes, which leads to observable increases in fluency. The term "proceduralization" refers to one of the processes in skill acquisition theory (SAT), a learning theory used to describe skill development.

SAT proponents, such as DeKeyser (2017), believe that the process of acquiring any skill, including the development of L2 spoken fluency, comprises three stages. The first stage is referred to as the declarative stage, wherein students develop knowledge about the processes necessary for task completion. In the procedural stage, they begin to practice these processes, and in the automatic stage, they perform the task consistently without any effort by repeated practice. When researchers refer to repeated practice, the optimal level of difficulty varies with students' prior learning, but generating skills or retrieving knowledge

from memory is desirable as long as students are equipped “by virtue of prior learning to succeed at that generation” (Bjork & Bjork, 2020). Content repetition plays an important role in moving formulaic utterances from the procedural stage to the automatic stage as it assists with long-term memory formation. Certain aspects of speaking performance such as formulaic utterances, when automatized, allow speakers to focus their attention on necessary features (Gatbonton & Segalowitz, 1988).

In other studies, Boers (2014) and Thai and Boers (2016) investigate the condition of time constraint. Participants who deliver monologues under the decreasing time condition (i.e., 4/3/2) outperform participants’ spoken fluency who have performed under a constant time condition (i.e., 3/3/3). Researchers use Levelt’s (1993) model of speech production to explain the effects of manipulating task variables on cognitive processes involved in producing speech. The main processes are the conceptualizer, formulator, and articulator. In the first process of speech production, the conceptualizer organizes thoughts and ideas into a coherent message, cultivating a pre-verbal message. Next, the formulator then transforms the intended message into linguistic representations including selecting words and organizing the structure of utterances. The articulator coordinates the motor movements for speech production, involving muscle activation, timing, and communication of phonemes, syllables, and words. In addition to utilizing the same cognitive processes of speech production, Gatbonton and Segalowitz (1988) recommend mimicking psychological pressures of real-world communication.

For example, the act of conceptualizing can be cognitively demanding for both first language speakers and L2 speakers as they carefully consider their intended speech before expressing it (Bui et al., 2019). Formulating and articulating can also be challenging for L2 speakers due to the heightened cognitive demands of retrieving relevant L2 lexical and syntactic information necessary for pre-verbal messages (Bui et al., 2019). Therefore, L2 speakers must allocate their attentional resources across the three stages of speech production (Ahmadian & Tavakoli, 2011; Skehan, 2014). L2 speakers might pay greater attention to formulating and articulating rather than conceptualizing by content repetition; this can lead to automatized changes in fundamental cognitive processes. In addition to the benefits of content repetition, Boers (2014) and Thai and Boers (2016) reason that the added time pressure condition of 4/3/2 might create the impetus to increase L2 spoken fluency automating the cognitive processes of conceptualizing, formulating, and articulating.

Without any pedagogical intervention, de Jong and Perfetti (2011) acknowledge that there exists uncertainty regarding exactly which lexico-grammatical items articulated during students’ monologues have been proceduralized. To address the ambiguity, researchers recommend supplying students with exemplars and encouraging integrating input into monologues (Boers, 2014; Thai & Boers, 2016). From these early studies, researchers interpret results based on SAT and Levelt’s (1993) model of speech production. They show that the conditions of content repetition and time constraint help students build L2 spoken fluency while acknowledging a need to incorporate pedagogical interventions to monitor the process of proceduralizing linguistic knowledge.

### Later 4/3/2 Studies

Reacting to these early studies' demand for focusing on target linguistic forms, researchers have explored the ability of pedagogical intervention to further L2 development. In Tran and Saito's (2021) study, providing corrective feedback on students' lexicogrammatical accuracy after each round not only improves accuracy but also does not diminish L2 spoken fluency gains. Researchers believe that incorporating formulaic utterances (i.e., automated language chunks) can increase students' L2 spoken fluency (Tavakoli & Wright, 2020). Ogawa (2019, 2021) investigates the extent to which encouraging the use of formulaic utterances by way of pedagogical interventions improves L2 spoken fluency. She uses three types of pedagogical interventions: input flooding, input enhancement, and peer feedback. For input flooding, she provides the participants with many instances of target language structures before the activity in exemplar monologues. For input enhancement, she highlights the target language structures in these exemplar monologues via italicizing, bolding, or emphasizing voice. Lastly, for peer feedback, during the activity, listeners use a checklist to indicate target language structures speakers use while delivering monologues. In the study, the group with all three pedagogical intervention types shows the most spoken fluency gains compared to the comparison group and the input-only group.

Unlike early studies of time pressure and repeated practice, pedagogical intervention is not an inherent condition of the 4/3/2 activity. Students can complete three iterations of their monologues without receiving feedback, whereas without the decreasing time condition and content repetition condition, the activity is no longer considered to be 4/3/2. However, feedback as a pedagogical intervention can be one of the most effective tools for enhancing students' academic achievements (Hattie & Timperley, 2007). Effective feedback can help students a) raise awareness of desired performance, b) know their current performance, and c) utilize strategies to bridge the disparity between their current performance and their goal, thereby enabling them to adjust their future actions and behaviors (Brooks et al., 2019). These three aspects of effective formative feedback are exemplified in the studies by Tran and Saito (2021) and Ogawa (2019, 2021) insofar as researchers raise awareness of desired performance, provide feedback on participants' current performance, and help participants utilize strategies to bridge the disparity.

One gap in the literature is to explore how instructors can use technology to provide in-class feedback on students' 3/2/1 performance. Using technology might assist the process of providing feedback by automating specific tasks such as analyzing speech rate, detecting pauses, or mean length of run while also providing a platform for classmates' perceived fluency ratings. Currently, instructors can only feel that their students are becoming more fluent. By leveraging mobile applications, instructors can streamline the feedback process while providing detailed, objective assessments on current and desired performances. During the pilot study, some challenges stem partially from the textbook prompts (Appendix B). Therefore, the following research question might shed light on these issues: What textbook-related challenges associated with the 3/2/1 prompts arise from providing participants' utterance fluency scores to track their monologic L2 speaking performance?

## Methods

### Participants

Participants ( $n = 35$ ; 20 females and 15 males) of the survey were first-year university students at a private university in Tokyo, Japan, and enrolled in the authors' EDC during Spring 2023. They were at least 18 years old and had completed at least six years of formal English as foreign language education prior to entering university (mean = 8.4). The participants ranged from B1 to C1 on the common European framework of reference (CEFR). The participant that we selected, pseudonym Jun, was from this pool of participants. The criteria for selection were that the participant had to have signed the consent form, attended all lessons, use the mobile application during the 3/2/1 activity, and exhibit patterns of participants' recordings and transcription. Before the semester, he passed Level 2 of the Eiken and had completed B1 on the CEFR. Jun's classmates ( $n = 11$ ; 4 females and 7 males) also participated in the study; they were B1 on the CEFR (i.e., a low Level 2 by Rikkyo University's system of labeling English proficiency). In total, 43 students from the authors' classes participated in the study. Three participants from Jun's class participated in the survey.

### Instruments

#### *Mobile Application*

Our spoken fluency application is designed to measure the utterance fluency of English language learning students. The features include measuring untrimmed and trimmed speech rate, pause frequency, articulation rate, and listener-based perceived fluency to provide feedback on students' spoken fluency performance across the semester. The application is in beta testing for iOS and thus has not yet been validated for in-class use. This semester, we have gathered data to validate the mobile application for measuring in-class L2 spoken fluency.

#### *Audacity*

Audacity is a free software by Audacity Team (<https://audacityteam.org/>) and is used for multi-track audio editing and recording digital audio. It is compatible with macOS, Linux, Windows, and other Unix-like operating systems. As of December 6, 2022, Audacity holds the title of the most downloaded software at FossHub, with over 114.2 million downloads since March 2015 (Appendix C).

#### *WhisperX*

Whisper is an open-source automated speech recognition system that has been trained using 680,000 hours of supervised data from various languages and tasks from the Internet (Radford et al., 2022). This has resulted in enhanced resilience to accents, background noise, and technical jargon. WhisperX, a variant of Whisper, provides word-level time stamps to incorporate precise temporal information (Bain et al., 2023). This system combines forced phoneme alignment and voice activity detection techniques to achieve its performance in word segmentation and long audio transcription, thereby giving practitioners and researchers



a tool for quick and accurate analysis of spoken language data (Bain et al., 2023).

### ***Praat***

Praat is a free phonetics software tool used for speech analysis (Boersma & Van Heuven, 2001). It is compatible with multiple operating systems and can examine and reconstruct acoustic speech signals. It provides a wide array of functions, such as speech analysis, manipulation, and synthesis, along with the capability to generate visuals.

### ***Survey***

The survey was in English and Japanese (Appendix D). The participants (n = 35) responded to our survey regarding reading aloud, which had the following two questions:

- a) When you were using the spoken fluency mobile application today, did you read aloud from the textbook's reading passages?
- b) If you read aloud from the textbook while using the mobile application, what is your reason for doing so? Reply in English or Japanese.

After the first survey, we created a second survey using the replies received to create an updated checklist (Appendix D).

- a) Below is a list of possible reasons for reading aloud while using the mobile application during the 3/2/1 activity. If you read aloud, please check all the reasons that apply.

### **Procedures**

After receiving approval from the Center for Foreign Language Education and Research at Rikkyo University, participants met once per week for 14 weeks. During Week 1, they learned how to use the mobile application. They participated in the 3/2/1 activity for the first time in Week 2. The instructions from the textbook were, "Talk to a partner. Say as much as you can. Don't worry about grammar or vocabulary" (Sturges et al., 2023). The goal was for participants to incorporate the formulaic utterances introduced in the EDC textbook and to speak as quickly as they could (Appendix E). The survey was distributed via Google Forms to participants during Weeks 6 and 7. The secondary author also conducted a follow-up interview in Week 8 with Jun and his classmates regarding their transcripts and survey results.

For every lesson, we used the prompts in the textbook (Appendix B). However, for Week 7, we recycled the same prompts from Week 3 to compare spoken fluency performances. We compared Week 3 performances because during Week 2, participants were still unfamiliar with 3/2/1. Participants had no pre-planning time to replicate real-world speaking conditions. After collecting the recordings, we transcribed them using WhisperX, then used a program written by the secondary author to convert transcripts to TextGrid, a visual aid that matches audio to transcriptions (Appendix F). We also used Audacity to view audio waveforms when doing the initial analysis. TextGrids were manually reviewed in Praat to a) align word

boundaries with the audio, b) add any missing fillers (e.g., uhs and ums), and c) delete any background voices. Then, we converted the file to JSON data to analyze the speech event and analyze the data (e.g., mean length of run). These newly analyzed speech events were collated into a single CSV file, further subject to statistical analysis.

## Analysis

For the quantitative analysis, six utterance fluency features were chosen from Suzuki et al.'s (2021) study. In this study, certain features seemed more highly correlated to perceived fluency than others. Hence, we selected trimmed speech rate, untrimmed speech rate, pause frequency, length of run, and pause location. We added pause length to better explain the results. For all graphs, we used untrimmed speech rate in syllables per minute (SPM) because speech rate had been the strongest indicator of perceived fluency. We analyzed the spoken fluency results using descriptive statistics. Table 1 explains the spoken fluency features.

**Table 1**

*Utterance Fluency Measures*

Type	Utterance Measure	Description
Composite (speed and breakdown)	Average Number of Trimmed SPM	Delete needless repetition (i.e., as in clearly a result of disfluency, not for purposes of emphasis) and fillers (e.g., uh, um), then divide the total number of spoken words by the duration, in minutes, of the speech.
Composite (speed, breakdown, and repair)	Average Number of Untrimmed SPM	Divide the total number of spoken words by the duration, in minutes, of the speech. Total number of words includes fillers, false starts, and repetitions.
Composite (speed, breakdown, and repair)	Mean Length of Run	Calculate the average number of untrimmed syllables in each uninterrupted speech segment between pauses, then add one to the total number of pauses, indicating the number of fluent sequences between silent breaks to determine the run count. The total syllable count would be divided by the run count. The syllable count can be verified on <a href="http://syllablecounter.net">syllablecounter.net</a> .
Pure (breakdown)	Average Frequency of Pauses per Minute	Fillers (e.g., um, uh) and silence were counted as pauses. Pauses were determined using Praat. Calculate the average number of pauses divided by the time.
Pure (breakdown)	Average Number of Mid-Clause Pauses per Minute	Count the number of filled or silent pauses located in mid-clause and divide the number by the total speaking time.
Pure (breakdown)	Mean Length of Pause	The length of silent pause was 300 ms as used in Ogawa's (2021) study. Pauses were determined using Praat. Calculate the average length of pauses by adding all pause times and divide by the number of pauses.

For the qualitative analysis section, we categorized potential problems into themes (e.g., participants who read aloud during the third prompt and lack of formulaic utterance use).

Survey questions and follow-up interviews with Jun and his classmates were categorized using thematic analysis to identify common themes, and the results were checked by both authors to reach an agreement regarding the categories that the individual responses should be placed in. Last, we used descriptive statistics to rank the most popular responses for the “select all that apply” question regarding their rationale for reading aloud.

### Results and Discussion

Figure 1 shows the average untrimmed speech rate of Jun and his classmates (n = 11) for all rounds of the 3/2/1 activity over a period of six weeks from Week 2 to Week 7.

**Figure 1**  
*Average Untrimmed SPM of Jun and His Classmates*

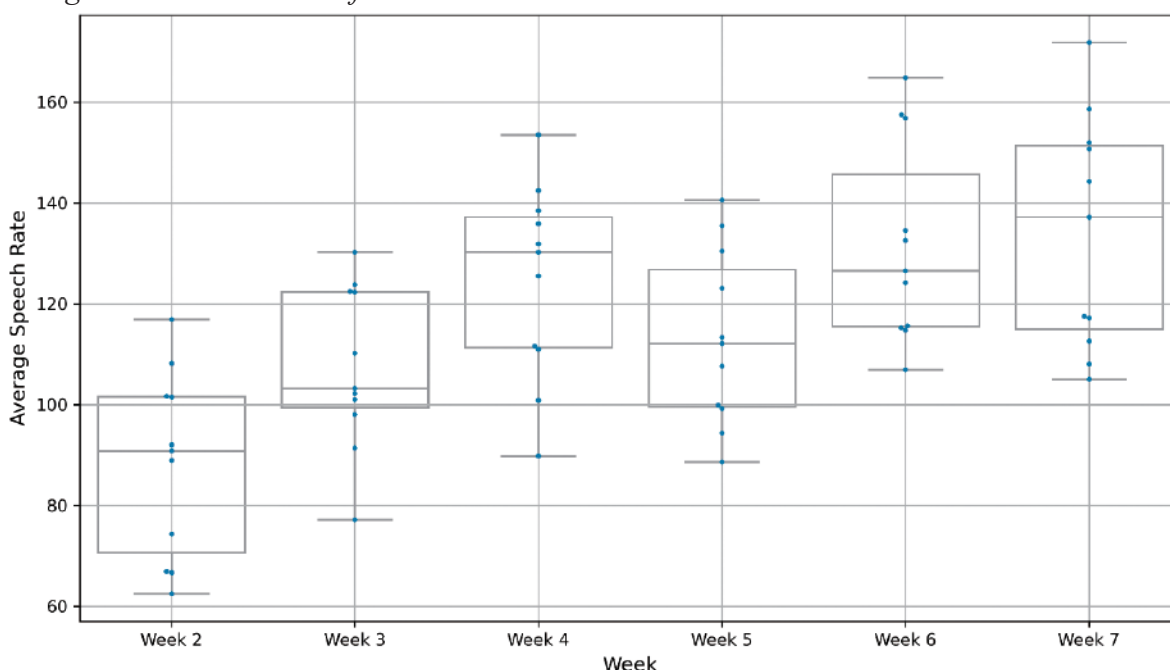
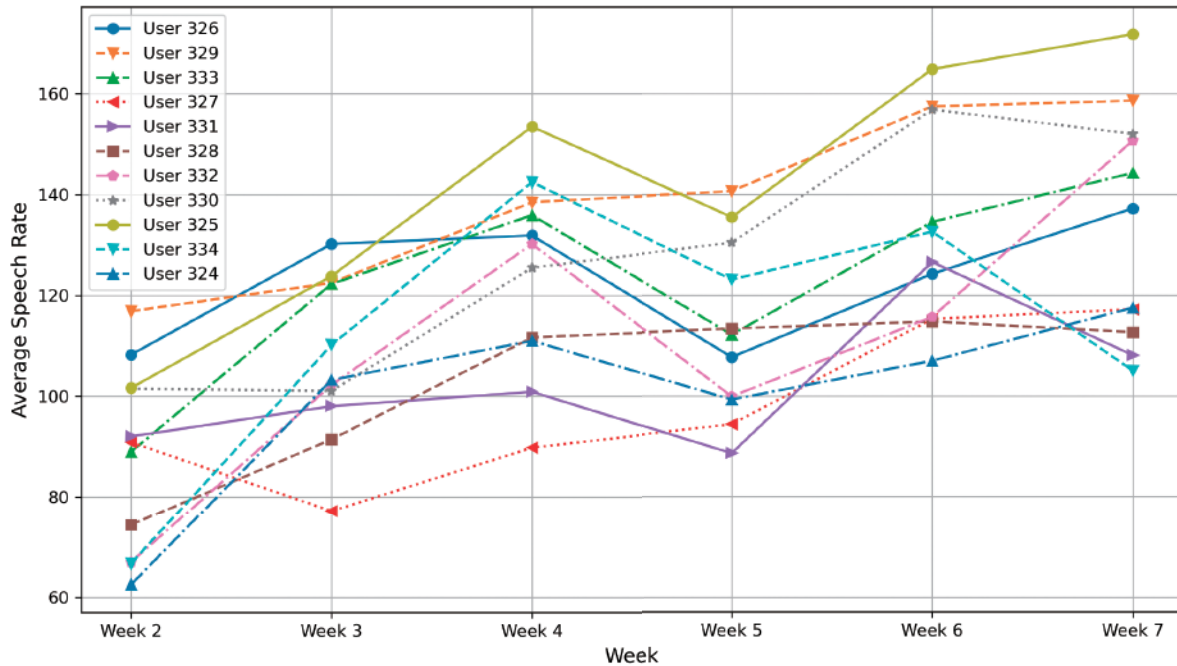


Figure 1 revealed an uptrend in speech rate, with a slight decline observed in Week 5. Based solely on this information, we cannot determine whether the 3/2/1 activity caused the speech rate increase as this sample was limited to 11 participants from a single site (Appendix G). The results aligned with previous 4/3/2 studies as they also showed that participants increased speech rate. From Week 2 to Week 3, the increase might be because Week 2 was the participants’ first time completing the 3/2/1 activity, so their initial performance gain might be the result of increased activity familiarity. The decrease from Week 4 to Week 5 might be caused by unfamiliarity with the prompt topics (i.e., part-time jobs and self-sufficiency) and possibly due to the anxiety participants might have felt as Week 5 was their first discussion test. Lastly, another possible explanation apart from the efficacy of 3/2/1 for the increase from Weeks 3 to 7 is the amount of L2 exposure participants had throughout the semester in other university English classes.

Although all participants increased their speech rate from Week 2 to Week 7, their

increases varied in degree. Figure 2 shows the individual progress of Jun and his classmates in terms of untrimmed speech rate in SPM over six weeks for only the last round of 3/2/1.

**Figure 2**  
*Individual Untrimmed SPM of Jun and His Classmates*

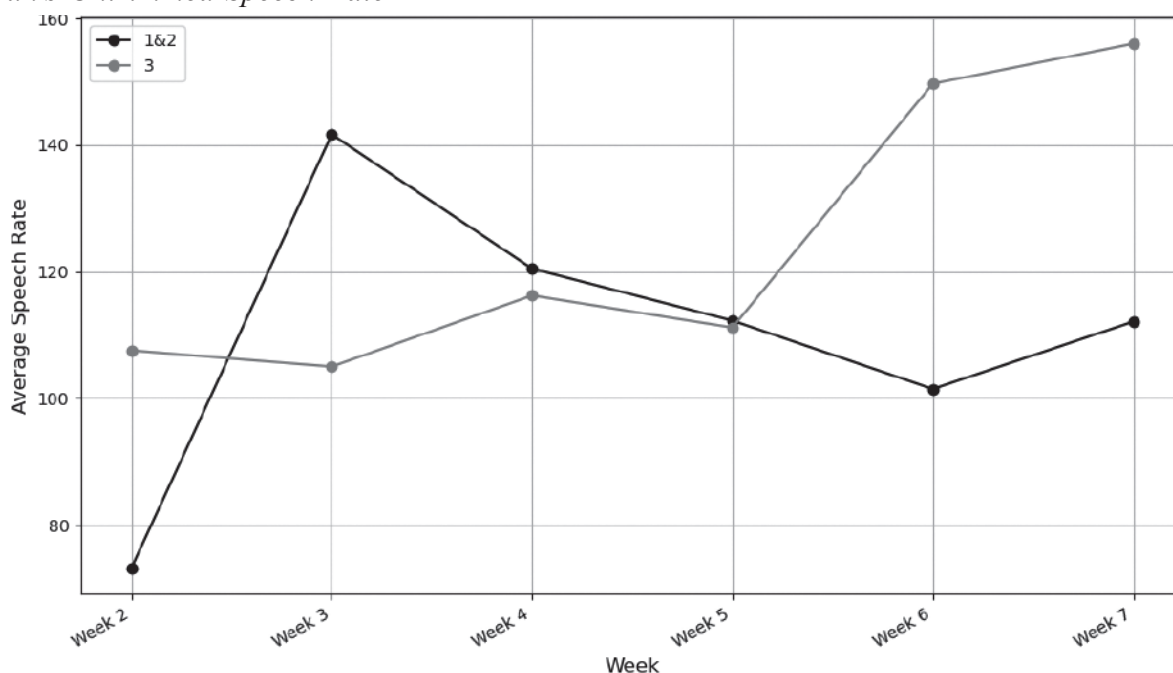


All participants had faster SPM in Week 2 than Week 7 (Appendix G). Moreover, from Week 3 to Week 7, nearly all participants except User 334 (light blue triangles) performed at a higher SPM. User 334 received a 110 SPM in Week 3 but 105 SPM in Week 7. Some participants recorded greater increases than others in SPM from Week 3 to Week 7. For example, User 325 (lime green circles) received 124 SPM on the mobile application in Week 3 and increased to 172 SPM in Week 7, which was an increase of 48 syllables. According to the secondary author, this participant worked hard and did relatively little reading aloud for the third prompt as compared to the others. Jun, User 324 (dark blue triangles), received 103 SPM in Week 3, and 118 SPM for Week 7. Users 331 (purple triangles) and 326 (dark blue circles) respectively received an SPM of 98 and 130 in Week 3 and 108 and 137 in Week 7; therefore, only a 10-syllable and 7-syllable increase was recorded.

To gain a better understanding of these results, we divided Jun’s monologue into the first and second prompts (personal opinions and experience-based prompts) versus the third prompt (find three interesting ideas from the reading) to investigate Jun’s untrimmed speech rates (Figure 3).

**Figure 3**

*Jun's Untrimmed Speech Rate*



Jun's untrimmed speech rate for the first and second prompts increased sharply from Week 2 to Week 3, perhaps in part due to the increased familiarity with the activity. After Week 3, his SPM had steadily decreased, whereas for the third prompt, he had increased his untrimmed speech rate from Weeks 5 to 7.

We further investigated this difference between the prompts across six utterance features. In Table 2, we present a comparison of Jun's performance in the third round of the 3/2/1 activity from Week 3 and Week 7, differentiating between the first and second prompts from the third prompt.

**Table 2**

*Jun's Utterance Fluency for the First, Second, and Third Prompts*

	Prompts							
	First and Second				Third			
	Weeks		Totals		Weeks		Totals	
	3	7	Raw Change	% of Change	3	7	Raw Change	% of Change
Untrimmed Speech Rate SPM	156.5	136.7	19.8	(12.7)	139.5	209.9	70.4	50.5
Trimmed Speech Rate SPM	118	108	10	(8.475)	140	180	40	28.571
Avg. Pause Frequency	25.0	21.5	3.5	(16.3)	17.2	19.8	2.6	15

	Prompts							
	First and Second				Third			
	Weeks		Totals		Weeks		Totals	
	3	7	Raw Change	% of Change	3	7	Raw Change	% of Change
Avg. Pause Length	0.8	1.4	0.6	75	0.9	0.8	0.1	(11)
Avg. Mid-Clause Pauses	6	13.333	7.333	122.217	1	0	1	(100)
Mean Length of Run	5.9	5.5	0.2	(6.8)	7.1	11.7	4.6	64.8
Time in Seconds	19.2	39.1	19.9	103.646	34.8	16.3	18.5	(53.2)

During Weeks 3 and 7, Jun experienced an undesirable outcome in five utterance fluency measures for the first and second prompts, despite addressing the same 3/2/1 questions as he did in Week 3. These measures were untrimmed and trimmed speech rates, pause length, mid-clause pauses, and length of run. However, one positive development was that he slightly reduced pause frequency. The secondary author noted that in class, he seemed to become increasingly confident in his ability to answer the first two prompts. For the third prompt, Jun always read aloud, resulting in a dramatic increase in his utterance fluency measures. All but pause frequency were positively impacted. Moreover, without separating the prompts, the application recorded 103 SPM in Week 3 and 118 SPM in Week 7. To explain this result, although he answered the first and second prompts at a slower pace in Week 7 compared to Week 3, he compensated by reading aloud during the third prompt more quickly for Week 7 than Week 3.

To provide further clarification regarding the results, we included transcripts of Jun’s 1-minute speeches. The results were categorized into different prompts (Table 3). The bold italics in the table signified instances when Jun borrowed lexico-grammatical items from the prompts. The prompts were as follows: “Did you study *hard to get into university*?” “Why did you *decide to go to this university*?” and “Share three interesting ideas or facts from the reading.”

**Table 3**  
*Jun’s Untrimmed 1-Minute Transcripts*

	Week 3	Week 7
First Prompt	I didn’t study <i>hard to get into university</i> because I don’t like studying.	Yes, I did. I studied <i>hard to get into university because uh I failed to get into university</i> last year. So I uh must study hard last year.

	Week 3	Week 7
Second Prompt	I <i>decided to go to this university because</i> uh <i>I passed</i> only this university in March.	uh I <i>decided to go to this university because I passed</i> uh Rikkyo, Seijo, and Toyo, so I uh selected Rikkyo University.
Third Prompt	Uh <i>I am interested in</i> uh reading this students who take entrance exams learn time management skills which are useful not just at university but for the rest of their lives. Universities are reviewing their wall interesting systems that may lead to new kinds of tests and vocational schools are quicker and cheaper and.	uh <i>I'm interested in</i> first as the world becomes more connected to the Internet, there will be many more examples in the future. Second, movie from Studio Ghibli and popular hits like Demon Slayer.

As can be seen from the transcriptions, Jun changed the content of his answers. For example, he said that he did not study hard to get into university because he did not like studying; however, in Week 7, he responded that he studied hard because he failed to get into university the previous year.<sup>1</sup> For the second prompt, he said that he only passed Rikkyo University, but in Week 7, he stated that he passed three universities and selected Rikkyo. The secondary author asked Jun for his rationale for changing his response to these questions to which he replied that formulating the response of not studying hard to get into university was less challenging than failing to get into university. In other words, he created answers that were inauthentic because the real answer was perceived by him to be too difficult to explain in English. There might also be external reasons such as Jun feeling more familiar with his classmates to share this information. Changing the content of his answers might be problematic as according to the results of de Jong and Perfetti's (2011) study, content repetition with repeated practice helps build spoken fluency. From our results, students might need additional planning time to provide monologues authentic to their beliefs with linguistic support from their instructor or peers to formulate their monologues. They also might benefit from knowing the rationale behind 3/2/1 (e.g., advantages of content repetition).

Regarding reading aloud for the third prompt, we had expected participants would, to an extent, extract language from the first and second prompts for support, but we had not expected that participants would rely on reading aloud from the textbook for the third prompt. As a result, the feedback Jun received via the mobile application on his current performance did not reflect his ability because reading aloud artificially boosted his fluency as he could read aloud more quickly than he could generate speech. In the survey, we asked participants, "How often do you read aloud from textbook reading passages during the 3/2/1 activity?" and 48.57% stated sometimes (n = 17), 25.71% stated rarely (n = 9), 14.29% said often (n = 5), 8.57% said never (n = 3), and 2.86% said always (n = 1). In Table 4, the survey question was as follows: "If you read aloud from the textbook while using the mobile application during the 3/2/1 activity, what is your reason for doing so?" (Check all that apply).

<sup>1</sup> We double checked to ensure that this participant was in fact the same participant as Week 3 because his answers to the same questions were contradictory.

**Table 4***Participants' (n = 35) Rationale for Reading Aloud*

<b>Rationale for Reading Aloud</b>	<b>No. of Participants</b>	<b>% of Agreement</b>
My mobile application score improves by reading aloud.	18	58.06
Reading aloud helps me to improve my spoken fluency.	16	51.61
I can answer the third prompt of three interesting ideas and facts.	14	45.16
It is difficult to recall the ideas from the reading passages.	14	45.16
I do not have enough of my own ideas to speak for the entire time.	9	29.03
I can help others understand my ideas better.	9	29.03
Speaking English is difficult.	8	25.81
There is no rule against reading aloud from the textbook.	7	22.58
I cannot summarize the ideas from the reading passage.	3	9.68
I have written my answers in the textbook.	1	3.23
Reading aloud is more enjoyable than speaking.	1	3.23

From the results, 58.06% of participants stated that reading aloud improves their score on the mobile application, followed by 51.61% who felt that reading aloud during 3/2/1 helps improve their spoken fluency. Further, 45.16% stated that they read aloud to answer the third prompt tied with the belief that recalling ideas from the reading was challenging. From the results, students might need an explanation of why they are not allowed to read aloud as they believe that doing so helps them build their spoken fluency. In terms of Bjork and Bjork's (2020) desirable difficulties, reading aloud would be less challenging than generating ideas from memory, but low proficiency students might not be equipped to succeed in generating contents of the textbook unless they have pre-planning time and additional strategies to expand their answers. In the Rikkyo University (2021), textbook writers stated that for low proficiency classes, instructors might consider allocating one minute pre-planning time to either write down or think about their ideas, whereas higher proficiency classes can begin speaking without planning time. Moreover, they wrote that instructors should consider the temporary strategy of reversing the timing (i.e., 1/2/3 rather than 3/2/1), which would equate to temporarily eliminating the condition of time constraint to support content expansion, to help low proficiency students speak for three minutes.

In terms of Levelt's (1993) model, by reading aloud, students were not engaged in the same cognitive speech production processes as they would be decoding, rather than encoding, conceptualizations and formulations; therefore, when answering the third prompt, they would not experience the same psychological pressures when engaged in speech. One of the conditions of Gatbonton and Segalowitz's (1988) creative automatization was mimicking psychological pressures of real-world communication. Experiencing these pressures might help further build their confidence when engaging in real-world communication. In the Rikkyo University (2021), the writers stated, "The fluency questions are simple to allow students to focus on improving fluency skills, rather than worrying about content or vocabulary." However, 45.16% of participants felt that recalling ideas from reading passages



was difficult, and 29.03% stated that they did not have enough of their own ideas to speak for the entire time. However, textbook writers proposed that the three questions before and the three after the reading could be suitable additional prompts for students (Rikkyo University, 2021).

### Utilization of EDC’s Formulaic Utterances Within 3/2/1

Apart from Jun’s change in the content of his responses and a reliance on reading aloud as a strategy to increase utterance fluency, we investigated the utilization of EDC’s formulaic utterances (Appendix E). From Table 4, Jun used the EDC discussion skill of reasons within his 3/2/1 monologues; however, if he were to incorporate additional formulaic utterances, this integration might increase his spoken fluency (Ogawa, 2021). We analyzed the prompts to count instances of formulaic utterance use embedded within the prompts (Table 5).

**Table 5**  
*Formulaic Utterance Use Within 3/2/1 Prompts*

Formulaic Utterances	Textbook Examples of Formulaic Utterances	Example 3/2/1 Prompts from the Textbook	No. of Instances
Follow-up Questions	What...? Who...? Is...? How...? Would...? If...? Which...? Do...? Did...?	Which social media do you use? How often do you use social media? Do you think learning a foreign language is important?	23
Reasons	Why do you think so?	Why? Why or why not?	6
Examples	For example?	What examples of pop culture do you like?	2
Opinions	What do you think about...?	n/a	0
Connecting Ideas	Do you agree/disagree with?	n/a	0
Viewpoints	How about from the viewpoint of...?	n/a	0
Advantages/Disadvantages	What is one advantage/disadvantage of...?	n/a	0
Source of Information	How do you know about that?	n/a	0

From Table 5, follow-up questions were the most prevalent 3/2/1 prompt-type with 23 instances. Unlike the other utterances, follow-up questions lacked any corresponding responses. For instance, responses to opinions were as follows: “In my opinion...” and “I think...” and for connecting the formulaic responses: “I disagree with... He said..., but I think...” Follow-up questions, although prevalent in the EDC textbook’s 3/2/1 activity, might not elicit the target linguistic forms (i.e., formulaic utterances) presented throughout EDC because there were no corresponding responses. In Ogawa’s (2021) study, she had peers provide feedback on the amount of formulaic utterance use, and participants did increase the

use of the formulaic utterances as a result. In addition to peer feedback, another possibility would be to embed the formulaic utterances within the prompts and have a rating system within the mobile application to rate the degree to which participants used the formulaic utterances. These features might encourage students to utilize various target formulaic utterances. Connecting ideas, viewpoints, disadvantages/advantages, and sources of information were excluded from the current textbook’s 3/2/1 prompts. One potential solution would be to include a diversity of formulaic utterances. An example of connecting and giving reasons would be as follows: “Aki says that foreigners should visit Tokyo rather than Kyoto. Do you agree with her? Why or why not?” (Appendix H).

### Recycling EDC’s Prompts Across the Semester

Table 6 below revealed transcripts from Jun’s first, second, and third prompts. The first and second prompts for Week 6 were as follows: “Would you like to go abroad? (e.g., to travel, to study, to work) Why or why not?” and “If you went abroad, would you go to an English-speaking country or somewhere different? Why?” The third prompt was “Share three interesting facts or ideas from the reading.”

**Table 6**

*Trimmed Transcripts of Jun’s First, Second, and Third Prompts*

Prompts	3-Minute	2-Minute	1-Minute
First	<i>I would not like to go abroad</i> because I’m afraid of plane. I don’t like high. <i>I would not like to go abroad.</i>	<i>I would like to go abroad to travel</i> because I want to see the pyramids and other things.	<i>I would like to go abroad to travel</i> because I want to see pyramid.
Second	<i>If I went abroad, I would go to an English-speaking country</i> because I can speak Japanese and English, but I want to communicate, so <i>I would go to an English-speaking country.</i> I’m interested in...	<i>If I went abroad, I would go to an English-speaking country</i> because I can speak Japanese and English <i>only</i> , but I want to communicate to abroad people, foreign people, <i>so I want to go to an English-speaking country.</i> I’m interested in...	<i>If I went abroad,</i> I wouldn’t go to <i>somewhere different speaking country</i> because I want to go to Egypt. I’m interested in ...

Prompts	3-Minute	2-Minute	1-Minute
Third	One of the best ways to learn a foreign language is to experience the culture and language together by studying or living abroad. <i>Second</i> , living in a dormitory didn't have to use a foreign language to communicate and did not improve their language skill much. <i>Finally</i> , eating food from home country either from restaurants abroad or sent from home can also help people overcome homesickness.	One of the best ways to learn a foreign language is to experience the culture and languages together by studying or living abroad. <i>Second</i> , living in a dormitory didn't have to use a foreign language to communicate and did not improve their language skills much. <i>Finally</i> , eating food from home country either from restaurants abroad or sent from home can also help people overcome homesickness.	One of the best ways to learn a foreign language is to experience the culture and languages together by studying or living abroad. Living in a dormitory didn't have to use a foreign language to communicate and did not improve their language skill much. Eating...
First Prompt Revisited	<b><i>If I went abroad, I would like to go to abroad to travel because I</i></b> like castle. I want to go to...		

As seen in Table 6, for the first prompts, Jun changed his answers again from not wanting to go abroad to wanting to visit Egypt. The secondary author stated that Jun changed his opinion of not wanting to go abroad because after the first round, the author suggested to him that if he did not like flying, he could instead go by boat. The idea of going by boat changed Jun's opinion. Later, however, he stated that he did not want to go to Egypt, but he felt that the answer would be easier to say than that he did not want to go abroad because of his fear of heights. He also changed from wanting to go to an English-speaking country to communicate in English to not wanting to go to an English-speaking country because he wanted to go to Egypt. One possibility is that he was persuaded by his classmates to change his answers.

In addition to reading aloud, Jun also relied on the strategy of slightly altering the prompts to integrate them into his monologues. These slight changes were bolded and italicized in Table 5. From the trimmed transcriptions of his three speeches, 122 syllables (23.02%) were derived from slightly manipulating the prompt, 272 syllables (51.32%) came from reading aloud from the textbook reading passage, and 136 syllables (25.66%) were self-generated (i.e., without the help of the textbook or prompts). Slightly manipulating the prompt would be classified under Gatbonton and Segalowitz's (1988) creative automatization. To further support the automatization of these slightly manipulated linguistic items, recycling the prompts might be helpful. Topics also vary in degrees of familiarity to participants; so, instead of the mobile application comparing performances between topics, comparing performances within topics might represent participants' utterance fluency gains.

Bui et al. (2019) investigated the spaced learning effects of task repetition on accuracy, complexity, and fluency. They found that the time between initial and repeated performances seemed to play a moderating role in the impact of task repetition on speed fluency. Specifically, the benefit of increased speed fluency was most pronounced when there were immediate or shorter intervals between performances. However, for the measure of repair fluency, scores were higher for one-week intervals between performances. From the results of their study, Bui et al. (2019) explained that as the interval between practice sessions increases, lexico-grammatical items become less readily accessible for retrieval. With longer intervals, it can be presumed that students would need to begin from the beginning and recreate the process of conceptualization and formulation. Consequently, this would result in noticeably reduced fluency in their spoken output. From some participant performances, we believe that the effects of spaced learning might help automatize these prompt responses (Appendix H).

### ***Limitations***

This pilot study was subject to certain limitations. Firstly, the participant pool was restricted to a limited number of individuals; moreover, these participants were from the authors' own classes. The study also relied on descriptive statistics rather than inferential because of the small number of participants, and the sample was not multi-site, thereby eliminating the generalizability of these findings beyond our participants. Although these observations were based on our participants, the prevalence of these practices such as reading aloud for the third prompt remained uncertain, especially in the context of using the mobile application. Another limitation pertained to the survey employed in the study, which was not formally validated. Therefore, there might have been alternative methodologically rigorous approaches to generate reasons for reading aloud that could have been implemented.

### **Conclusion**

Monitoring progress matters as the language-based rationale for EDC is to build their spoken fluency, and with detailed feedback, instructors can make better pedagogical decisions. Hence, we developed a mobile application to measure in-class L2 spoken fluency of monologic speeches during the 3/2/1 activity, a modified version of the 4/3/2 activity by Maurice (1983). We investigated one participant's monologic utterance fluency for six weeks to exemplify the problems that occurred with monitoring his spoken fluency for formative feedback purposes. The possible issues concerned a) a lack of repetition of topics, b) no strategy training during pre-planning time, and c) an under-utilization of formulaic utterances. Potential solutions included recycling previous prompts, removing the third prompt of sharing three ideas and facts, and creating additional prompts embedded with target formulaic utterances to encourage use. Avenues for future research include the effects of such changes on mitigating issues experienced by mobile application users.

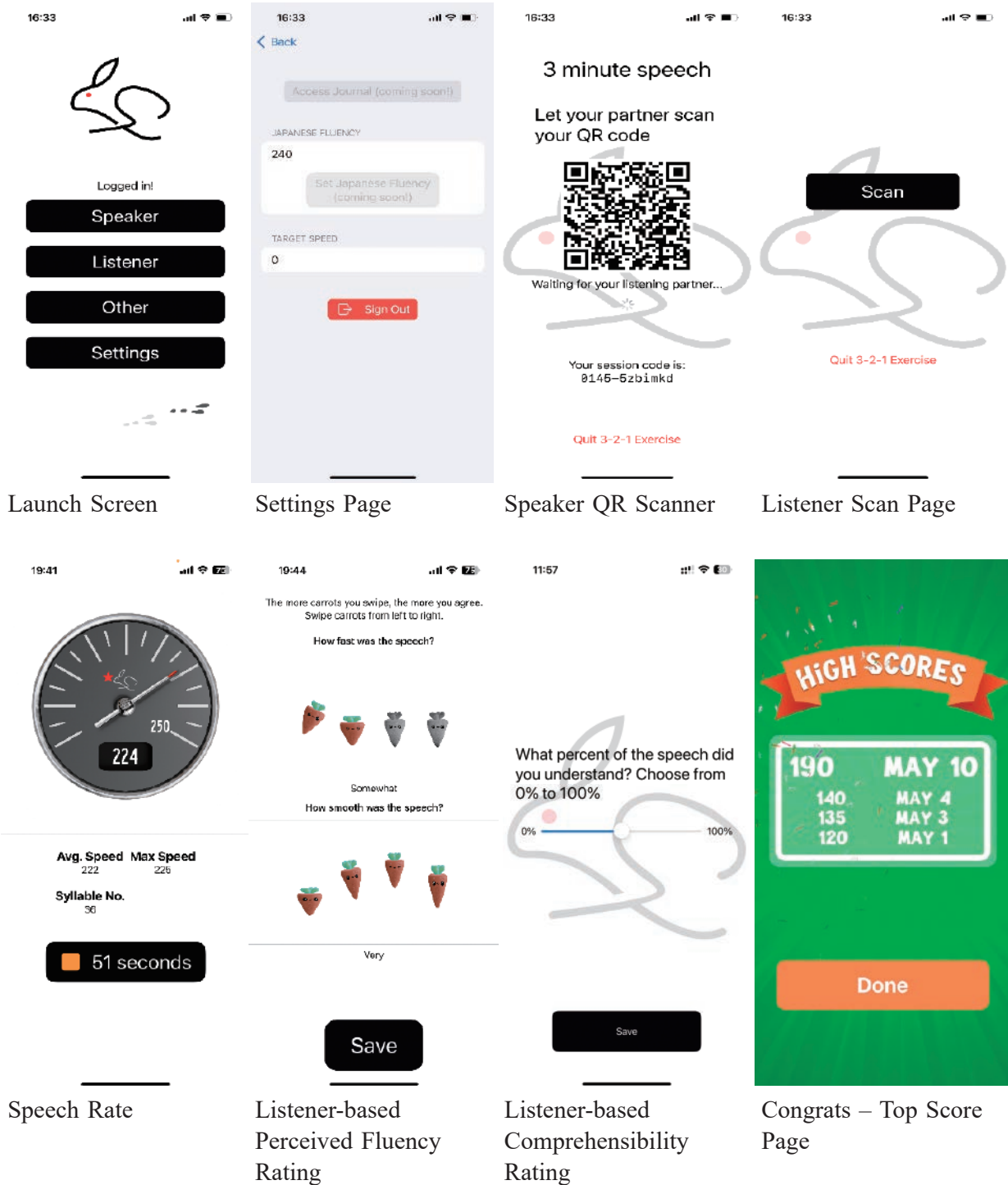
## References

- Ahmadian, M. J., & Tavakoli, M. (2011). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Language Teaching Research*, 15(1), 35–59.  
<https://doi.org/10.1177/1362168810383329>
- Bain, M., Huh, J., Han, T., & Zisserman, A. (2023). WhisperX: Time-accurate speech transcription of long-form audio. *arXiv*. <https://doi.org/10.48550/arXiv.2303.00747>
- Bjork, R. A., & Bjork, E. L. (2020). Desirable difficulties in theory and practice. *Journal of Applied Research in Memory and Cognition*, 9(4), 475–479.  
<https://doi.org/10.1016/j.jarmac.2020.09.003>
- Boers, F. (2014). A reappraisal of the 4/3/2 activity. *RELC Journal*, 45(3), 221–235.  
<https://doi.org/10.1177/0033688214546964>
- Boersma, P., & Van Heuven, V. (2001). Praat, a system for doing phonetics by computer. *Glott International*, 5(9), 341–347.
- Brooks, C., Carroll, A., Gillies, R. M., & Hattie, J. (2019). A matrix of feedback for learning. *Australian Journal of Teacher Education*, 44(4). <https://doi.org/10.14221/ajte.2018v44n4.2>
- Bui, G., Ahmadian, M. J., & Hunter, A. (2019). Spacing effects on repeated L2 task performance. *System*, 81, 1–13. <https://doi.org/10.1016/j.system.2018.12.006>
- De Jong, N., & Perfetti, C. A. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. *Language Learning*, 61(2), 533–568. <https://doi.org/10.1111/j.1467-9922.2010.00620.x>
- DeKeyser, R. M. (2017). Knowledge and skill in ISLA. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 15–32). Routledge.
- Gatbonton, E., & Segalowitz, N. (1988). Creative automatization: Principles for promoting fluency within a communicative framework. *TESOL Quarterly*, 22(3), 473–492.  
<https://doi.org/10.2307/3587290>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Hurling, S. (2012). Introduction to EDC. *New Directions in Teaching and Learning English Discussion*, 1(1), 2–10.
- Levelt, W. J. M. (1993). *Speaking: From intention to articulation*. MIT Press.  
<https://doi.org/10.7551/mitpress/6393.001.0001>
- Maurice, K. (1983). The fluency workshop. *TESOL Newsletter*, 17, 429.
- Ogawa, C. (2019). *A one-semester form-focused intervention on the development of speaking proficiency*. [Doctoral dissertation, Temple University]. Temple University Libraries.  
<https://doi.org/10.34944/dspace/2025>
- Ogawa, C. (2021). Revised 4/3/2 task: Fluency training with formulaic language in the EFL classroom. *The Journal of Asia TEFL*, 18(4), 1108–1127.  
<https://doi.org/10.18823/asiatefl.2021.18.4.3.1108>
- Ogawa, C. (2022). CAF indices and human ratings of oral performances in an opinion-based monologue task. *Language Testing in Asia*, 12(4).  
<https://doi.org/10.1186/s40468-022-00154-9>

- Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2022). Robust speech recognition via large-scale weak supervision. *arXiv*.  
<https://doi.org/10.48550/arXiv.2212.04356>
- Rikkyo University. (2021). *Instructor Handbook: English Discussion Class*.
- Skehan, P. (2014). *Processing perspectives on task performance*. John Benjamins Publishing Company. <https://doi.org/10.1075/tblt.5>
- Sturges, J., Arthurson, D., Itoi, K., Koike, A., & Reid, S. (2023). *What's your opinion? Interactive skills for effective discussion – Book II*. (3rd ed.). DTP Publishing.
- 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., & Wright, C. (2020). *Second language speech fluency: From research to practice*. Cambridge University Press. <https://doi.org/10.1017/9781108589109>
- Thai, C., & Boers, F. (2016). Repeating a monologue under increasing time pressure: Effects on fluency, complexity, and accuracy. *TESOL Quarterly*, 50(2), 369–393.  
<https://doi.org/10.1002/tesq.232>
- Tran, M. N., & Saito, K. (2021). Effects of the 4/3/2 activity revisited: Extending Boers (2014) and Thai & Boers (2016). *Language Teaching Research*.  
<https://doi.org/10.1177/1362168821994136>

## Appendices

### Appendix A: Spoken Fluency Mobile Application



### Appendix B: 3/2/1 Prompts in the EDC Textbook

Week/ Lesson	3/2/1 Prompts
2	<ul style="list-style-type: none"> <li>▪ Who are your best friends? What do you talk about?</li> <li>▪ Who do you talk to when you have a problem? Why?</li> <li>▪ Share three interesting ideas from the reading.</li> </ul>
3	<ul style="list-style-type: none"> <li>▪ Did you study hard to get into university?</li> <li>▪ Why did you decide to go to this university?</li> <li>▪ Share three interesting ideas or facts from the reading.</li> </ul>
4	<ul style="list-style-type: none"> <li>▪ Before coming to university, did you think it would be easy or difficult to make friends at university?</li> <li>▪ What do you want to do after you graduate from university? (e.g., job, family, travel)</li> <li>▪ Share three interesting ideas or facts from the reading.</li> </ul>
5	<ul style="list-style-type: none"> <li>▪ What kind of part-time job would you like to try?</li> <li>▪ In what ways are you independent? (e.g., living alone, doing housework, making important decisions)</li> <li>▪ Share three interesting ideas or facts from the reading.</li> </ul>
6	<ul style="list-style-type: none"> <li>▪ Would you like to go abroad? (e.g., to travel, to study, to work) Why or why not?</li> <li>▪ If you went abroad, would you go to an English-speaking country or somewhere different? Why?</li> <li>▪ Share three interesting ideas or facts from the reading.</li> </ul>
7	<ul style="list-style-type: none"> <li>▪ What examples of traditional culture do you like? Why?</li> <li>▪ What examples of pop culture do you like? Why?</li> <li>▪ Share three interesting ideas or facts from the reading.</li> </ul>
8	<ul style="list-style-type: none"> <li>▪ What customs do you follow? (e.g., on holidays, at festivals, in daily life)</li> <li>▪ What customs from other countries have you experienced? (e.g., Halloween, birthday cake, Chinese New Year)</li> <li>▪ Share three interesting ideas or facts from the reading.</li> </ul>
9	<ul style="list-style-type: none"> <li>▪ Do you think learning a foreign language is important?</li> <li>▪ What foreign languages have you studied? How did you study those languages?</li> <li>▪ Share three interesting ideas or facts from the reading.</li> </ul>
10	<ul style="list-style-type: none"> <li>▪ Do you want to use a foreign language at work? Why or why not?</li> <li>▪ How do you balance studying with other activities? (e.g., going to club activities, working part-time, enjoying free time)</li> <li>▪ Share three interesting ideas or facts from the reading.</li> </ul>
11	<ul style="list-style-type: none"> <li>▪ Which social media do you use or not use? Why?</li> <li>▪ How often do you use social media?</li> <li>▪ Share three interesting ideas or facts from the reading.</li> </ul>
12	<ul style="list-style-type: none"> <li>▪ What public rules do you usually follow? (e.g., at school, on trains, in restaurants, on the street)</li> <li>▪ What are some common bad manners in public places? (e.g., on trains, in cafes, at university)</li> <li>▪ Share three interesting ideas or facts from the reading.</li> </ul>



Week/ Lesson	3/2/1 Prompts
13	<ul style="list-style-type: none"> <li>▪ Is poverty a problem in Japan?</li> <li>▪ What can younger people do to help elderly people?</li> <li>▪ Share three interesting ideas or facts from the reading.</li> </ul>
14	<ul style="list-style-type: none"> <li>▪ What is your best memory from your first semester at university?</li> <li>▪ What was difficult about your first semester at university?</li> <li>▪ What skills and personal qualities have you developed at university?</li> </ul>

Note. Adapted from *What's Your Opinion? Interactive Skills for Effective Discussion—Book II*, by Sturges et al., 2023.

### Appendix C: Copyright Notices

Audacity® software is copyright © 1999–2023 Audacity Team.

Web site: <https://audacityteam.org/>

It is free software distributed under the terms of the GNU General Public License.

The name Audacity® is a registered trademark.

### Appendix D: Questions and Prompts for Survey on Reading Aloud

1	When you were using the spoken fluency mobile application today, did you read aloud from passages in the textbook? 今日、Spoken Fluencyモバイルアプリケーションを使用する際、教科書の文章を音読したのですか? Yes, no, I don't remember
2	When using the mobile application, how often do you read aloud from reading passages in the textbook? モバイルアプリケーションを使用する際、教科書の文章を音読する頻度はどのくらいですか? Never, rarely, sometimes, often, always
3	If you read aloud from the textbook while using the mobile application, what is your reasoning for doing so? (You can reply in Japanese or English) モバイルアプリケーションを使用しながら教科書を音読する場合、その理由は何ですか? (日本語でも英語でも構いません。) (short answer response)
4	<p>Below is a list of possible reasons for reading aloud while using the mobile application during the 3/2/1 activity. If you read aloud, please check all that apply. 以下は、3/2/1の活動でモバイルアプリケーションを使用しながら音読をする場合に考えられる理由です。音読をされる方は、該当するものをすべてチェックしてください。</p> <ul style="list-style-type: none"> <li>▪ I can improve my spoken fluency (speed of speech) by reading aloud. 私は、音読することでスピーキングの流暢さ（話すスピード）を向上させることができます。</li> <li>▪ I am lacking enough content to speak for the entire given time. 与えられた時間のすべてを話すには、内容が不足している。</li> <li>▪ The third question asks me about interesting facts and ideas from the reading. I cannot recall the ideas and facts from the reading passages. 3つ目の質問は、読書から得た興味深い事実やアイデアについて尋ねるものです。リーディングパッセージからアイデアやファクトを思い出すことができない。</li> <li>▪ Speaking English is difficult for me. 英語を話すことは私にとって難しいことです。</li> <li>▪ I enjoy reading aloud more than speaking. 話すことよりも音読の方が楽しいです。</li> <li>▪ I can improve my mobile application score by reading aloud. 音読をすることで、流暢な話し方のモバイルアプリケーションのスコアを向上させることができますね。</li> <li>▪ There is no rule against reading aloud from the textbook. 教科書を音読してはいけないという決まりはありません。</li> <li>▪ Other:</li> </ul>

### Appendix E: Discussion and Comprehension Skills

Unit Goal	Lesson No.	Discussion Skill	Examples
	1	Introduction to English Discussion Class	
Sharing and Supporting Opinions	2	Opinions	In my opinion, ... I think... What do you think?
	3	Supporting Opinions	One reason/example is... Why do you think so?
	4	Follow-up Questions	What...? Which...? How...? Do you...? Can you...?
	5	Comprehension Skills	Do you understand? I see. I understand.
Organizing a Discussion	6	Connecting Ideas	I agree/disagree. You said... What do you think of my idea?
	7	Joining a Discussion	Can I start? Can I say something? Would anyone like to ask a question?
	8	Changing Topics	What shall we discuss first/next? Is there anything more to add? So, we agree/disagree about...
	9	Paraphrasing	In other words...? So, do you mean...? I mean...
Challenging and Evaluating Ideas	10	Different Viewpoints	From (X's) point of view... How about (X's) point of view?
	11	Balancing Opinions	One advantage/disadvantage of... What's one advantage/disadvantage?
	12	Sources of Information	According to... I read/heard/saw/learned... Where did you read/see/hear that?
	13	Clarification	Can you explain? What do you mean? Could you repeat?
	14	Review	

Note. Adapted from *What's Your Opinion? Interactive Skills for Effective Discussion—Book II*, by Sturges et al., 2023.

### Appendix F: Example Output Data From Automated Analysis

Transcript	Transcript (Raw)	Pauses	Utterance Fluency
My best friend is <b>redacted</b> He is my high school student My high school friends She knows a lot of things and intelligent [1] Why are they your friends	My best friend is <b>redacted</b> He is my high school student My high school friends She knows a lot of things and intelligent Why are they your friends I	1: Unfilled Pause - [12.4–14.8] [Length: 2.4] 2: Unfilled Pause - [16.4–19.2] [Length: 2.8] 3: Unfilled Pause -	pause count 22 total pause length 71.4 articulation rate

Transcript	Transcript (Raw)	Pauses	Utterance Fluency
<p>[2] I think they know a lot of things and they make they make me happy and [3] I feel I feel exciting when they when I talk to them [4] I talk to I [5] I talk to my high school friends when I have a problem This is because they know they know [6] me [7] they know a lot of things about me and and [8] they [9] they can they can give me best advice for me [10] And [11] about about [12] about university universal universal course I will I will talk to I [13] will talk to my college friends when I have a problem about college [14] about college and [15] This is girls They are they are my [16] they are same They are same [17] They are same [18] They are same So they know a lot of things about my courage [19] But my high school student is different High school student is not my is not same courage [20] we can we can not [21] to talk about my problem [22] I like she's cute and she's intelligent and and</p>	<p>think they know a lot of things and they make they make me happy and I feel I feel exciting when they when I talk to them I talk to I I talk to my high school friends when I have a problem This is because they know they know me they know a lot of things about me and and they they can they can give me best advice for me And about about about university universal universal course I will I will talk to I will talk to my college friends when I have a problem about college about college and This is girls They are they are my they are same They are same They are same They are same So they know a lot of things about my courage But my high school student is different High school student is not my is not same courage we can we can not to talk about my problem I like she's cute and she's intelligent and</p>	<p>[25.7–28.6] [Length: 3.0]                      4: Unfilled Pause - [33.8–41.2] [Length: 7.4]                      5: Unfilled Pause - [42.1–43.1] [Length: 1.0]                      6: Unfilled Pause - [52.2–53.9] [Length: 1.7]                      7: Unfilled Pause - [54.4–56.5] [Length: 2.2]                      8: Unfilled Pause - [60.6–66.4] [Length: 5.8]                      9: Unfilled Pause - [66.7–68.5] [Length: 1.8]                      10: Unfilled Pause - [72.7–74.0] [Length: 1.3]                      11: Unfilled Pause - [74.2–76.2] [Length: 2.0]                      12: Unfilled Pause - [77.4–78.5] [Length: 1.1]                      13: Unfilled Pause - [84.8–86.0] [Length: 1.2]                      14: Unfilled Pause - [92.5–93.6] [Length: 1.1]                      15: Unfilled Pause - [94.7–98.8] [Length: 4.1]                      16: Unfilled Pause - [101.1–102.4] [Length: 1.2]                      17: Unfilled Pause - [104.8–106.0] [Length: 1.2]                      18: Unfilled Pause - [106.5–109.7] [Length: 3.2]</p>	<p>4.7                      mean length of run                      10.4                      syllables (untrimmed per minute)                      241                      start time in seconds                      3.8                      end time in seconds                      157.4                      speech rate (trimmed per minute)                      94.2                      Untrimmed total syllable count                      153.6</p>

Transcript	Transcript (Raw)	Pauses	Utterance Fluency
		19: Unfilled Pause - [114.0–115.8] [Length: 1.8] 20: Unfilled Pause - [123.9–126.1] [Length: 2.2] 21: Unfilled Pause - [127.6–128.8] [Length: 1.2] 22: Unfilled Pause - [131.7–153.3] [Length: 21.6]	

**Appendix G: Untrimmed Speech Rate (n = 11) Average of All 3/2/1 Rounds for Six Weeks According to Automated Transcription Method**

Users	Speech Rates					
	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7
324	63	103	111	99	107	118
325	102	124	153	135	165	172
326	108	130	132	108	124	137
327	91	77	90	94	115	117
328	74	91	112	113	115	113
329	117	122	138	141	157	159
330	101	101	125	130	157	152
331	92	98	101	89	127	108
332	67	102	130	100	116	151
333	89	122	136	112	135	144
334	67	110	142	123	133	105

**Appendix H: New 3/2/1 Prompts**

Week	Prompts
2	<ul style="list-style-type: none"> <li>▪ Who are your best friends?</li> <li>▪ What do you talk about with your best friends? (e.g., movies, music, classes, club activities, relationships)</li> <li>▪ When you have a problem do you usually talk to your friends or family? Why?</li> <li>▪ In your opinion, what is the importance of having friends?</li> </ul>
3	<ul style="list-style-type: none"> <li>▪ Who are your best friends?</li> <li>▪ What do you talk about with your best friends?</li> <li>▪ Did you study hard to get into this university? Why or why not?</li> <li>▪ Why did you decide to go to this university?</li> <li>▪ Do you think high school is stressful?</li> <li>▪ What are examples of stressful experiences for high school students?</li> </ul>

Week	Prompts
4	<ul style="list-style-type: none"> <li>▪ Did you study hard to get into this university? Why or why not?</li> <li>▪ Why did you decide to go to this university?</li> <li>▪ Did you think it would be easy or difficult to make friends at university?</li> <li>▪ Why did you think it would be easy or difficult?</li> <li>▪ What do you want to do after you graduate from university? (e.g., job, family, travel)</li> </ul>
5	<ul style="list-style-type: none"> <li>▪ Did you think it would be easy or difficult to make friends at university?</li> <li>▪ Why did you think it would be easy or difficult?</li> <li>▪ What kind of part-time job would you like to try?</li> <li>▪ In what ways are you independent? (e.g., living alone, doing housework, making important decisions)</li> <li>▪ Do you think it is important to be independent? Why or why not?</li> </ul>
6	<ul style="list-style-type: none"> <li>▪ What kind of part-time job would you like to try?</li> <li>▪ In what ways are you independent? (e.g., living alone, doing housework, making important decisions)</li> <li>▪ Would you like to go abroad? (e.g., to travel, to study, to work) Why or why not?</li> <li>▪ If you went abroad, would you go to an English-speaking country or a non-English-speaking country? Why?</li> <li>▪ What country would you visit? Why?</li> </ul>
7	<ul style="list-style-type: none"> <li>▪ Would you like to go abroad? (e.g., to travel, to study, to work) Why or why not?</li> <li>▪ If you went abroad, would you go to an English-speaking country or a non-English-speaking country? Why?</li> <li>▪ What are your favorite examples of traditional culture? Why? (e.g., music, tea ceremony, Obon festival, judo)</li> <li>▪ What are your favorite examples of pop culture? Why? (e.g., music, anime, manga)</li> <li>▪ Do you prefer traditional culture or pop culture? Why?</li> </ul>
8	<ul style="list-style-type: none"> <li>▪ What are your favorite examples of traditional culture? Why? (e.g., music, tea ceremony, Obon festival, judo)</li> <li>▪ What are your favorite examples of pop culture? Why? (e.g., music, anime, manga)</li> <li>▪ What Japanese customs do you follow? (e.g., on holidays, at festivals, in daily life)</li> <li>▪ What foreign customs have you experienced? (e.g., Halloween, birthday cake, Chinese New Year)</li> <li>▪ Do you prefer Japanese customs or foreign customs? Why?</li> </ul>
9	<ul style="list-style-type: none"> <li>▪ What examples of Japanese customs do you follow? (e.g., on holidays, at festivals, in daily life)</li> <li>▪ What examples of foreign customs have you experienced? (e.g., Halloween, birthday cake, Chinese New Year)</li> <li>▪ Do you prefer Japanese customs or foreign customs? Why?</li> <li>▪ What foreign languages have you studied? Why?</li> <li>▪ Do you think learning a foreign language is important? Why?</li> <li>▪ In your opinion, what is the best way to study a foreign language? (e.g., watching tv, reading, listening to music, making foreign friends) Why?</li> </ul>

Week	Prompts
10	<ul style="list-style-type: none"> <li>▪ What foreign languages have you studied? Why?</li> <li>▪ Do you think learning a foreign language is important? Why?</li> <li>▪ Jun believes everyone should use a foreign language at work. Do you agree?</li> <li>▪ Why or why not?</li> <li>▪ From the viewpoint of university students, what are good ways to relax during the semester? (e.g., going to club activities, hot springs, drinking tea, hanging out with friends)</li> <li>▪ What are the advantages of relaxing during the semester?</li> <li>▪ What are the disadvantages of relaxing during the semester?</li> </ul>
11	<ul style="list-style-type: none"> <li>▪ Jun believes everyone should use a foreign language at work. Do you agree?</li> <li>▪ Why or why not?</li> <li>▪ From the viewpoint of university students, what are good ways to relax during the semester? (e.g., going to club activities, hiking, going to hot springs, drinking tea)</li> <li>▪ What are some examples of social media?</li> <li>▪ Ryo thinks the best social media for students is Instagram. Do you agree?</li> <li>▪ Why or why not?</li> <li>▪ What is one disadvantage of social media?</li> </ul>
12	<ul style="list-style-type: none"> <li>▪ What are some examples of social media?</li> <li>▪ Ryo thinks the best social media for students is Instagram. Do you agree?</li> <li>▪ Why or why not?</li> <li>▪ What is one disadvantage of social media?</li> <li>▪ What do you think about following public rules? (e.g., school rules, train rules)</li> <li>▪ Why?</li> <li>▪ From the viewpoint of teachers, what are the advantages of following rules? (e.g., doing homework)</li> </ul>
13	<ul style="list-style-type: none"> <li>▪ What do you think about following public rules? (e.g., school rules, train rules)</li> <li>▪ Why?</li> <li>▪ From the viewpoint of teachers, what are the advantages of following rules? (e.g., doing homework)</li> <li>▪ Aki thinks that poverty is a problem in Japan. Do you agree?</li> <li>▪ How do you know about that?</li> <li>▪ What can younger people do to help elderly people? (e.g., do household work, make conversation, offer to buy groceries)</li> </ul>
14	<ul style="list-style-type: none"> <li>▪ Aki thinks that poverty is a problem in Japan. Do you agree?</li> <li>▪ How do you know about that?</li> <li>▪ What can younger people do to help elderly people? (e.g., do household work, make conversation, offer to buy groceries)</li> <li>▪ How do you know about that?</li> <li>▪ Who are your best friends?</li> <li>▪ What do you talk about with your best friends? (e.g., movies, music, classes, club activities, relationships)</li> <li>▪ When you have a problem do you usually talk to your friends or family? Why?</li> </ul>

Note. Adapted from *What's Your Opinion? Interactive Skills for Effective Discussion—Book II*, by Sturges et al., 2023.

# **Content, Language, and Rock n' Roll: Finding the Harmony Between Content and Language in a CLIL Course**

**Kevin Thomas**

---

## **Abstract**

This paper details how a Content and Language Integrated Learning (CLIL) course was designed and the research conducted on two groups of students to explore their attitudes to decisions made during its creation. CLIL was chosen as the teaching approach to a Lecture and Discussion course at Rikkyo University. The decision was based on the course's fit in wider language provision at the university and the assumed content goals and language needs of students. The CLIL course design outlines the Language Triptych, and the 4Cs: Content, Cognition, Communication, and Culture were considered when designing the course along with different approaches to CLIL, namely hard and soft. A harder approach to CLIL was taken, meaning the course was more balanced more toward content than language support. Two studies were carried out over two subsequent years, which found that students had overwhelmingly positive attitudes toward the harder approach to CLIL. They showed appreciation for the balance between content and language, which allowed for the opportunity to focus on specialized content of interest. They also felt positively toward the flexible approach to language support allowed by the approach. Although the research results were seen as indicating student satisfaction with the course, it was decided that the research method was flawed, and subsequent studies must consider a more exploratory interview-based approach.

**Keywords:** *CLIL course, British music, Course design, 4Cs*

## **Introduction**

Before starting work at Rikkyo University, teachers are asked to provide their preference of classes to teach. The Lecture and Discussion class is one such class. This is an advanced level class and consists of two 100-minute classes per week for 14 weeks. It is an elective course, meaning students can opt for it. The course is unavailable for first-year students, and students choosing this course will have taken mandatory English language classes and may possibly still be taking language-focused classes.

The content and curriculum of this class is to be decided by the teacher. As a new

teacher, the author of this paper took this opportunity to propose a course based on his passion for British politics, culture, and music. The author was assigned the course and started to think about materials to use. An abundance of materials were available to the author such as CDs, streaming subscriptions, music magazines, books, newspaper articles, memorabilia, and YouTube playlists of music-related documentaries, videos, interviews, and news stories.

The material considered for use was intended for native speakers with a large amount of cultural knowledge assumed. It could be challenge for anyone unacquainted with British culture and more challenging for people with English as a second language. Hence, it had to be decided if the material should be adapted or presented in its original form. This led to an investigation of Content and Language Integrated Learning (CLIL). The author had some awareness of CLIL due to its increasing popularity as a teaching approach both internationally and in Japan. CLIL has been defined as an approach to teaching in which the delivery of content and language teaching are combined. The teacher may decide to emphasize content or language teaching, but both must be present for it to be considered CLIL (Coyle et al., 2010). CLIL was summarized by the author in a review of a book on CLIL as “not a pedagogy in itself but as an approach that allows the course designer or teacher (provider) to choose between the best language supportive methodologies while delivering content in the desired second language” (Thomas, 2021). This means there are different varieties of CLIL used to fit different contexts. It seemed like a suitable approach to teach content on British music and society while supporting students in their learning.

CLIL as a concept arose in Europe “during the mid-1990s to adhere to the European Union’s (EU) multilingual policies” (Tsuchiya & Pérez Murillo, 2015, p. 26). At the time, it was the EU’s objective for its citizens to speak two or more languages. CLIL at this time was mainly conducted in schools and included content teachers teaching subjects such as math or science in a second language. These teachers were usually non-native and untrained language teachers (Tsuchiya & Pérez Murillo, 2015). The practice of CLIL spread to European universities, and in the mid-2000s, inspired by these European countries, Japanese universities started providing CLIL classes (Pinner, 2013). Since 2009, there has been increasing awareness of CLIL at Japanese universities (MacGregor, 2016), and the government has encouraged universities to include English medium education (Morizumi, 2015). In fact, at the authors current institution, there has been an ongoing development of courses supported by CLIL theory and methodology (Yamamoto & Nitta, 2021). However, according to Brown and Bradford (2017), there has been no common approach to how CLIL is implemented, and there has been confusion between CLIL and other English medium teaching approaches, namely English medium instruction (EMI) and content based instruction (CBI). There has also been reports of student resistance to CLIL due to its unfamiliarity and perceived lack of usefulness (Smith & Ssali, 2019).

## **Literature Review**

### **CLIL, EMI, and CBI**

CLIL is an approach to teaching that integrates the provision of content knowledge with



language support. It is not simply providing content in English or as a science teacher wrote “a teacher blabbing about physics in English is not CLIL because CLIL attends to the learners’ ability to use language” (Teresa Ting, 2011). However, as Mok (2021) argues, CLIL classes should not be simply language classes with content themes “where content is still the means to the end (language)” (p. 55).

According to Brown and Bradford (2017), CLIL has been confused by institutions and teachers with EMI and CBI. This has led to instructors not providing the balance between content provision and language support integral to CLIL. EMI refers to content being taught in English, wherein content provision is connected to the learners major and content knowledge is the basis of assessment. As Bradford (2018) explains, “the goal of EMI is subject-content knowledge, not the acquisition of language” (p. 65). In EMI, there is no need to provide language support to help learners understand content or encourage retention; therefore, it clearly is not CLIL. Alternatively, CBI has been described as a teaching method wherein content is the medium for teaching language. The teacher teaches a subject and related content, but the learners understanding and learning of language is the focus of the teacher. CBI curricula can be organized sequentially, with tasks that require higher level language comprehension placed later in the curricula. It seems safe to argue neither EMI nor CBI can be seen as interchangeable with CLIL, which is supposed to integrate content and language support. CLIL classes need to support students in understanding content unlike EMI, but it should not let language dictate what is taught and when, as in CBI.

### **The Language Triptych**

To guide teachers in creating CLIL courses, Coyle et al. (2010) created the Language Triptych framework, which can be used to help select the language presented and activities used in classes to enable students to tackle content and retain language. The Triptych is represented as a triangle, with the three points being *Language of learning*, *Language through learning*, and *Language for learning* with CLIL linguistic progression being in the center (Martin del Pozo, 2016). The first point of the triangle, *language of learning*, refers to the language required to understand the meaning of the content. Instructors should consider the functional and notional levels of difficulty demanded by the content. In other words, can learners grasp the ideas being taught? Are they relatable to learners of a certain age, in a certain country, and in a certain context? The second point of the triangle, *language for learning*, concerns the language that learners need to use in the process of learning, such as asking questions and discussions. The third point, *language through learning*, focuses on how activities included in lessons can encourage language retention. By utilizing the Language Triptych, the teacher does not change the content to be presented or reorganize it in order of language difficulty. Rather, the teacher can use the framework to reflect on what language support students can be given to help them benefit most from lessons in terms of both content and language.

### **The 4Cs**

In addition to the Language Triptych, Coyle et al. (2010) proposed the concept of the 4Cs, which can be used as a framework for designing CLIL lessons, enabling course

designers to incorporate theory in addition to needed functional language in their lesson plans (Iyobe & Li, 2013). The 4Cs stand for *content*, *cognition*, *communication*, and *culture*. Used to guide teachers in the combination of content and language support, the 4Cs allow them to identify whether a gap exists between a student's language ability and the cognitive demands of the given content. The first C, *content*, refers to the need in CLIL classes for content learning, which is the knowledge and skills provided about the theme of the class. *Content* does not just refer to the content presented; rather, a successful CLIL course should enable students to research the theme further and create their own content knowledge (Meyer, 2010). Coyle et al. (2010) suggest that language and *communication* should be viewed as inseparable. In a CLIL class, students need to articulate observations on the meaning of content, which of course will require essential communicative language. The second C, *cognition*, in other words, thinking and understanding, is considered the most important (MacGregor, 2016). *Cognition* refers to the ability of learners to apply their newly acquired content knowledge to contexts outside the class (e.g., with colleagues or friends). The final C, *culture*, refers to students understanding and interacting with the cultural community in which the content is based. For example, if one was to teach a course on American movies, students would need to know the sociopolitical background against which certain movies were made if they were to understand the aim of the person making the movie beyond the monetary benefit.

### **CLIL in Japan: Hard vs Soft**

In Japan, English has been seen as an important element of education with the aim of making the country and its citizenship globally competitive and able to participate in the global society. However, English proficiency in Japan has remained comparatively low (Leontjev & deBoer, 2022).

The adoption of CLIL in Europe inspired the Japanese government, universities, and educators to consider it as possibly useful in meeting the nation's language learning goals (Lockley, 2015). Lockley (2015) argues that CLIL fosters international posture, which was described by Yashima (2002) as an attitude that includes "interest in foreign or international affairs, willingness to go overseas to stay or work, readiness to interact with intercultural partners, and, one hopes, openness or a non-ethnocentric attitude toward different cultures" (p. 57). This aspect of CLIL, which is compatible with Japan's English language goals, was acknowledged by Ohmori (2014) who suggested that it encourages cultural understanding. She suggests as learners are made to grapple with meaning in its original cultural context and unchanged for Japanese consumption, they are given an insight of the culture usual language classes fail to provide.

Sophia University and Saitama University, who started piloting CLIL courses in 2008 and 2010 respectively, were leaders in CLIL education in Japan (Ohmori, 2014). Inspired by these examples, other universities attempted to provide CLIL classes, but these courses were seen to take very different approaches (Brown & Bradford, 2017). One such institution is Rikkyo University, which has started providing CLIL classes in an attempt to bridge its first-year mandatory English language classes and EMI classes available to second to fourth year students (Yamamoto & Nitta, 2021). The CLIL courses at Rikkyo University interweave content with explicit instruction on study skills. This inclusion categorizes the course at

Rikkyo as a softer variety of CLIL.

Ohmori (2014) identifies hard and soft CLIL as varieties of approach to CLIL with differing emphasis on content and language. Hard CLIL concentrates on content, with the teacher providing no language support. It could be argued that such an approach would be undistinguishable from EMI and inaccessible for only very proficient students. In contrast, soft CLIL is weighted toward language support, with content being the theme or delivery method of language and not in its original cultural context. This may help learner comprehension but would lose the cultural context desired by CLIL practitioners.

CLIL courses in Japan have tended toward soft CLIL (Leontjev & deBoer, 2022) due to misunderstanding of CLIL, students being less proficient in English and needing more help than their European counterparts, and the provision of CLIL being situated in language learning departments. Ohmori (2014) suggests that the CLIL courses at Saitama University tended to be weighted toward language learning to be considered as CLIL. This was due to the courses being organized by the language department and teachers having insufficient content knowledge to teach subjects effectively.

### Curriculum Design and Assessment

When designing a CLIL curriculum, it seems reasonable to suggest a softer or harder stance rather than a completely soft or hard approach that would better accommodate the language and content elements of CLIL. Using a softer or harder method would necessitate significantly different approaches to curriculum design. However, the 4Cs and Language Triptych frameworks can be applied to either form of CLIL. There are also recognized features of all CLIL curriculum.

CLIL courses should utilize authentic materials, i.e., material in its original context free from adaptations, although support should be given based on the learners' language proficiency (MacGregor, 2016). CLIL lessons should necessitate students to use higher order thinking skills (Margana & Widyanoro, 2017). Higher order thinking entails students critically interacting with authentic material and evaluating and discussing it (Tanaka, 2019). This interaction with authentic material corresponds with the *cognition* aspect of the 4Cs. CLIL courses should instill global human resources in students, allowing students to participate in the global society by enabling collection of information in English, thinking in English, and sharing ideas in English (MacGregor, 2016). Margana and Widyanoro (2017) argue that having students search for material in its original context is essential to CLIL, which complements the *culture* element of the 4Cs.

If a softer approach is chosen when designing the curriculum, teachers must imagine what challenges students may encounter with authentic materials and plan explicit support for them. This means formally providing students with the language *of* and *for* learning identified in the Language Triptych. Soft CLIL curricula should clearly state planned interactions (Leontjev & deBoer, 2022). For example, a CLIL debate course curricula at Rikkyo University includes both language support and study skills in its curriculum. These interventions are planned for specified times in the course to support student activities. Language skills include teaching phrases for making propositions and rebutting, whereas study skills include instructions on how to take notes (Nitta & Yamamoto, 2020). Another

example of language support in a softer CLIL class is the teaching of tenses before showing a video to aid comprehension in a course for people in the hospitality industry (González González et al., 2022).

When choosing a hard approach, language support needs to be given but does not need to be included in the curricula. The elements of CLIL identified by the 4Cs and Language Triptych should be present but need not be formally stated. The teacher is aware that language is being taught, but this is not explicitly communicated to the learner. Language is taught inductively, presented in its original context with students able to learn from examples (Mok, 2021). This entails grammar or lexis being taught and then students should be asked to notice it in its original context.

Although language support does not need to be stated in the curricula, language comprehension can be supported both premeditatively and spontaneously. Teachers can predict when students in a class may have problems and teach at this “episode” (Mok, 2021). This could be done by priming students through the scaffolding of lexis (Valiukiene, 2017), which means providing students with language they can use to decode a text for meaning. This lexis is usually slightly above the students’ anticipated level. For example, before students watch a video or read a text, the teacher could prepare them by voicing and visually presenting lexis contained in the material. This could also be done for background and cultural knowledge. Another way of priming students for possibly challenging material is to activate their schemata (Korosidou & Griva, 2014). This means a teacher should conduct an activity around or discuss what students already know on a topic before adding to it. This utilization of background knowledge is argued to make students aware that they have some knowledge on a topic, allowing them to more effectively grapple with new knowledge.

Not all language support needs to be anticipated, support can be given when problems arise (Leontjev & deBoer, 2022). This could take the form of corrective feedback. Mok (2021) uses the notions of driver and co-driver to suggest how feedback can be given. He uses the metaphor of how a co-driver can set a destination, the driver chooses their own route and the co-driver corrects the drive along the journey if needed. This could be done by repeating incorrect grammar and vocabulary, giving necessary phrases, or focusing on phonology.

As there has been debate about the breakdown between how curricula should be weighted in terms of content or language in a curriculum, there has been similar debate concerning assessment (Otto & Estrada, 2019). There is a broad agreement that weighting should follow the curriculum, but there is an argument over how the two can be combined in a rubric available to learners.

## **The Study**

The study took place over two years with two groups of students taking a 14-week course in the second semesters of 2021–2022 and 2022–2023.

## Study 1

### *Context*

The class studied was a Lecture and Discussion class, which consists of two 100-minute classes over 14 weeks. This class is an elective class. The class is not available to first-year students, and students are required to have an advanced level of English. The university stipulates that in this class, students should be given a university style academic lecture and given opportunities to discuss given and related contents. Alongside gaining content knowledge, it is explicit in the online curriculum available to students that they will practice study skills such as taking notes and conducting research. British music and society was selected as the content for this class due to the authors interest and assumed interest of the students. The title of the course was “British Music 1960 – Present: How music reflects and inspires social change.”

It was chosen to split the lecture and discussion elements between classes, so there would be one 100-minute lecture class per week and one 100-minute discussion. One music movement would be the focus for each of the first eleven weeks. During the lecture, the teacher would give a slideshow presentation detailing the narrative of the movement, with authentic materials embedded in the lecture. During the discussion, following a teacher-led model discussion, students would present something they had researched connected to the theme of the previous lecture and lead a discussion on it (Rikkyo University, 2021).

A CLIL approach to the class was taken as it was believed the content and language elements of CLIL would be present throughout the curriculum and in each class. As for content, authentic materials would be presented in English by someone involved in the culture concerned, as a music enthusiast in this case. As for language, the class is provided by the language department, so it would seem safe to suggest students had some language goals by choosing to take the class. Although the students possessed an advanced level of English, it was anticipated that they would at times need language support to navigate challenging notions and material. It was also stated to students in the curriculum that they would have opportunities to practice study skills related to language learning.

A hard approach to CLIL was chosen. The students would have already taken mandatory language classes, such as discussions, debates, and presentations, and possessed an advanced level of English. The students taking the course would have also chosen it over more language-based elective courses, suggesting that they were seeking something different. Brown (2015) suggests that students choose CLIL courses due to the challenge it presents and the opportunity to interact with authentic materials.

Authentic materials embedded in lectures included music videos, clips from documentaries, TV news features, and quotations from newspapers, music magazines, and books. Magazine quotations included excerpts of music reviews, opinion pieces, and articles on artists and movements, whereas book excerpts included biographies, autobiographies, and genre explorations. Between the lecture and discussion classes, students were given a choice of materials to further aid their research for the discussions; these included Apple and Spotify playlists, books and magazines from my personal library, and links to online newspaper articles. Students would also be encouraged to find their own materials related to the lecture, which could be focused on a similar movement in another country. For the discussion classes,

sometimes the teacher would bring their own realia or memorabilia to instigate the discussion.

Language support would be both planned and spontaneous. Narrative slides before focus on authentic materials would include lexis and background knowledge key to aid comprehension, giving students a scaffold of the language to aid comprehension. For example, students would be provided with lexis and information about the West Indies, empire, immigration, integration, and assimilation, before a documentary video clip on the experiences of British reggae artists as this language and knowledge would be essential for comprehension the video. The teacher would also attempt to elicit students' background knowledge and schemata on cultural context of musical movements such as the British economic situation at the time of punk rock. Language support would also be provided spontaneously when needed.

The class would be assessed mainly on content. Students would be assessed on the understanding they showed of the materials they selected for discussions. The final assessment would be a presentation with students assessed on breadth of research and understanding of the topic. Language would only be considered if it hindered expression of meaning.

The curriculum created followed the Language Triptych and 4Cs framework. *Language of learning* would be given in the scaffolding of language in lecture slide presentations and the modeling of discussions. *Language for learning* would also be modeled by the teacher when leading and participating in discussions. Students taking notes in lectures and needing to retain language for use in discussions and assessments would correspond with *language from learning*. As for the 4Cs, the class would be built on *content* and feature authentic materials. *Communication* would be the basis of discussion classes. *Cognition* would be present in discussion involving giving opinions, cultural comparisons, and lyric interpretation. The course is based on *culture*, and the students would have to make comparisons to life and music in Japan during the relevant eras.

### ***Participants***

The members of the study group were three males from the third year. During orientation, they all appeared to possess a similar upper-intermediate to advanced level of English. Two of them displayed a wide knowledge of British music and identified this as the reason they chose the course, whereas one said he did not know much about music but was interested in the sociological and political elements of the class. The students in the study will be referred to as Students 1, 2, and 3.

### ***Objectives***

The main objective of the study was to investigate the student's opinions concerning the hard approach to CLIL. To discover this, whether the students' reasons for taking the course were content or language related needed to be investigated. If they were mainly language related, a softer approach in the future would need to be considered. Whether the content weighted breakdown of the course was perceived to be conducive to learning and enjoyment of the course would also be a research objective. If students felt a content-focused approach was impeding their ability to learn, it seems a softer approach should be taken. Enjoyment is arguably an important element of any class, especially an elective class that students

choose to take due to personal interest. Therefore, it was important to know how the hard approach to CLIL affected enjoyment. Moreover, students' opinions about a harder CLIL approach changed over the course would be explored. Resistance to CLIL has been reported due to its unfamiliarity to Japanese students (Smith & Ssali, 2019). Therefore, it would be necessary to ascertain if such resistance is present and whether it changed as students become accustomed to the approach.

### ***Data Collection***

Three questionnaires were given to the participants, a pre-course after the first lesson (Appendix A), mid-course after Week 7 (Appendix B), and post-course after Week 14 (Appendix C). Questionnaires were translated into Japanese by a bilingual associate and both Japanese and English were used in the document; only English has been included in the appendices due to word count limitations. Permission for the research was granted by the university ethics committee, and a consent form for each document was prepared according to the institution's regulations. The documents were given and received via the course's Google Classroom. The questionnaire comprised six questions with a seventh space for other comments.

The research tool for the questionnaire was chosen as they can be used to collect attitudinal data (Dornyei & Taguchi, 2010), and as the research was conducted during the pandemic, it was seen as a safe and reliable way to collect information. The majority of information gathered was qualitative in that the questionnaire comprised open-ended questions but did contain a quantitative element in one question. A mainly qualitative method was chosen as it would hopefully produce more data, and the possibility of a small number of participants was anticipated.

The first and second questions in the pre-course questionnaire differed from that of the mid- and post-course questionnaires. The other questions remained the same to assess any change in attitudes with the wording being changed to reflect the stage of the course. The first question in the pre-course questionnaire related to why students chose the course to judge whether they had content- or language-related goals, whereas the first question on the later questionnaires considered enjoyment of the course. The second question on the pre-course questionnaire concerned a more detailed questioning of student's language and content goals, whereas the later ones were concerned with difficulty. The third question involved how much language support students felt they should receive to ascertain whether their attitudes complemented the chosen harder approach to CLIL. The fourth question included students desired balance of content and language, asking them to explain and quantify their preferred balance. This was again to gauge how students' opinions aligned with the chosen approach of the course design. The fifth and sixth questions were concerned with assessment and the course's cohesion with other studies and was outside the remit of this course but was included for further research.

### ***Results***

Quantitative data collected was put into Excel and converted into graphic data to allow results and patterns of change over the course to be more easily observed. Japanese responses

were anonymized and translated by a bilingual assistant. Qualitative data was read several times and then color-coded according to the theme. For example, when exploring students' reasons for joining the course, language- and content-related aims were highlighted in different colors. These were then collated to explore similarities and differences in perspectives, so pertinent observations could be communicated when writing up the results.

**Pre-Course Questionnaire.** Students 1 and 2 indicated that their primary motivation was not only interest in class content but also a wish to develop listening and speaking skills. Student 3 was purely interested in the content. Students 1 and 2 believed that the teacher should give explicit language support to students. Explanation of the vocabulary and “teach ways of listening” was requested by Student 1, whereas Student 2 requested “a key vocab list” to be given. Student 3 desired no help. Students 1–3 desired a breakdown of 75% / 25%, 50% / 50%, and 80% / 20% between content and language, respectively.

**Mid-Course Questionnaire.** Student 2 was unable to submit his questionnaire. Students 1 and 2 both indicated they were enjoying the course, whereas Student 3 said “it was hard to listen” to lectures “for over an hour.” Students 1 and 2 were satisfied in terms of language support, indicating “enough help provided” and “given enough help.” Both were unchanged in their desired class time breakdown between language support and content with Student 1 stating “if you want to learn language, there is a dedicated class.” Student 1 said listening was difficult, but he was “able to grasp” the meaning of the content.

**Post-Course Questionnaire.** All students indicated they enjoyed the class. Student 1 indicated “British music was the reason I chose the class, so I enjoyed it.” The opinion that the class has a relaxed atmosphere and there was ample speaking time was shared by all respondents. Further, all respondents felt that the difficulty was appropriate. The three students felt they were given enough language help. Student 1 said help was always given if needed, whereas Student 2 said “help given was flexible. Teacher gave important vocabulary when necessary. When the class becomes huge, flexibility may be more difficult.” Students 1 and 3 were unchanged in their ideas about the ideal balance between content and language with Student 3 indicating the balance “was appropriate for the students in the class.” Student 2 said an appropriate balance for a future course should be 70% content to 30% language, changing from his initially stated equal split. All students indicated they found the lectures hard to concentrate on for the duration of the lesson.

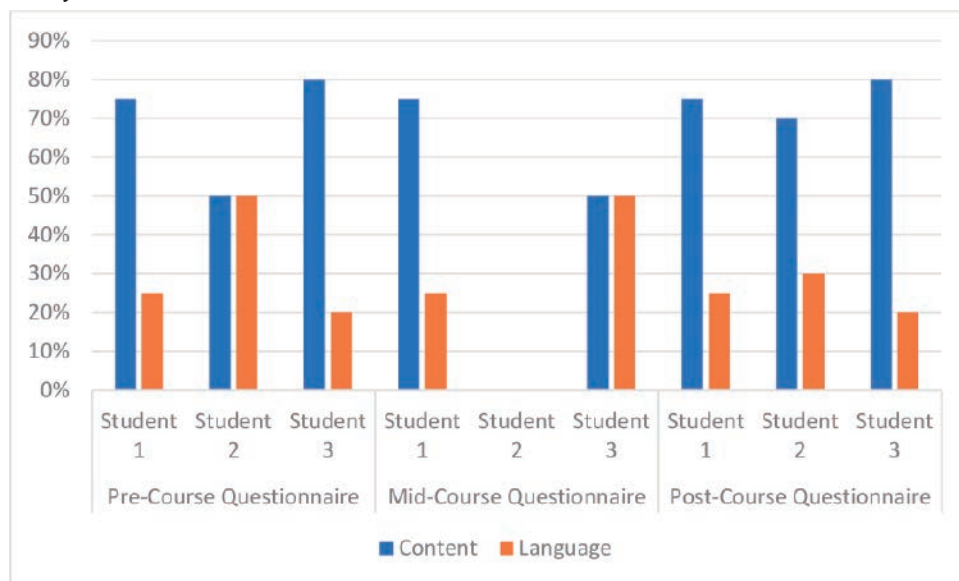
### ***Discussion***

Overall, the results suggest the harder CLIL approach was justified. Content was given as the primary goal for joining the course with a lesser emphasis on skills practice. This was indicated throughout the course with students expressing appreciation for the content knowledge provided. Student 2 who initially favored an equal weighting, moved toward prioritizing content.



**Figure 1**

*Study 1 Students' Desired Class Time Focus*



Difficulty was seen as appropriate with listening indicated as a challenge by Student 1, but the content was still graspable. Although explicit language support such as vocabulary lists and listening skills strategy focus were not provided, students felt they received sufficient help. Student 2 indicated the “flexibility” of the approach to language support allowed aid to be given when needed. This corresponds to the harder CLIL approach of dealing with language problems when they arise. However, he questioned whether such support could be given if dealing with a larger group. All students said they found it difficult to listen and concentrate for an entire lesson.

These concerns inspired me to conduct the research again with the following year’s students. If the group was bigger, I would be able to assess whether the flexible harder CLIL approach to language support was seen as appropriate for larger classes. I would also attempt to deal with the issue of students finding concentration difficult during lectures.

## Study 2

### *Context*

In 2022, I taught the Lecture and Discussion course to a second group of students. The curriculum remained the same and the same harder approach to CLIL was taken. The only intervention was three 7-minute discussion intervals were added after authentic content had been viewed. For example, after a lecture on British hip-hop artists of the 80s and 90s and viewing a selection of music videos, students were asked the following: If they liked any of the music? If it has a different sound and image from US hip-hop and rap? How British social conditions influenced the subject matter? The inclusion of these intervals meant cutting some lecture points and shortening videos.

The intervals as well as giving students a break as requested by the Study 1 students could have benefits for learning. Discussions after content presentation have been suggested to help students process and retain language (Puspitasari, 2016), which corresponds with the

*language through learning* point of the Language Triptych. Discussion can also help integrate students' schemata with newly taught content and encourage comradery. Student 2 of Study 1 identified the bond between students as contributing to learning as students felt free to experiment with language.

### ***Participants***

The students were four females and three males from the third and fourth years who had participated in the same mandatory English language programs as the previous study group. All students said they had an interest in British music during the orientation. Students were all of an advanced level apart from one who seemed considerably lower. They all appeared keen to take the course. The participants will be referred to as Students 1–7.

### ***Objectives***

Objectives remained the same apart from additionally wanting to investigate if including discussion intervals would result in students not indicating difficulty maintaining concentration during lectures.

### ***Results***

Similar to the first group, the quantitative data was converted into graphic form for analysis. The qualitative data was highlighted and put into themes to ascertain any pattern. Perspectives given in earlier questionnaires were contrasted with to in later ones to see if changes in opinions were present.

**Pre-Course Questionnaire.** Student 4 did not submit the questionnaire. All respondents indicated their main reason for joining the course was content. For example, Student 1 said, "I love UK rock," whereas Student 3 indicated an interest in British bands. Three students indicated an interest in the social background of music. Student 7 said one reason to select the course was as the lecturer was British, they would meet someone who had experienced content first-hand. The desired class time breakdown between focus on content and language support differed with Student 3 desiring 100% content as "I will learn English by myself" to Student 2 wanting a 50/50 split. Although the balance varied, all students, except Student 2, thought balance should weigh toward content with remaining student percentages varying from 60% to 80%. Six students said they should receive help when needed, whereas Student 2 said key vocabulary should be presented pre-task. Student 7 said the teacher should "help in terms of terminology but don't interrupt, it could be discouraging."

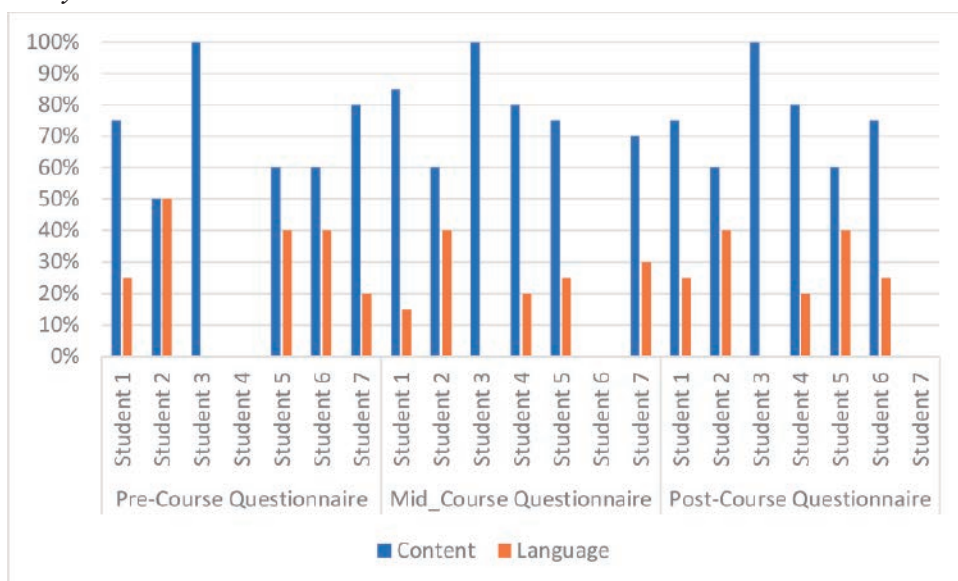
**Mid-Course Questionnaire.** Student 6 did not complete the questionnaire. All respondents said they were enjoying the class and ascribed this to content. For example, Student 1 said they were "deepening their knowledge of music," and Student 5 said they enjoyed the playlist and "made new discoveries every week." Student 3 said part of their enjoyment was due to being able to speak without being criticized. Student 1 said others seemed to have difficulties, and Student 3 said the content could be challenging, all indicating the area of difficulty as social context. However, they all said they were receiving enough help; Student 7 said, "slang was explained." This time all students said class time should be weighted toward content with percentages ranging from 60% to 100%.

**Post-Course Questionnaire.** Student 7 did not respond. All respondents said they enjoyed the course very much, three going as far as to say it was the best course they had taken at the university. All indicated content knowledge as key to enjoyment. Two students said their increased knowledge of context helped them understand songs better. Student 2 said conversations were also key to enjoyment, whereas four students said the atmosphere of the class was enjoyable with Student 4 stating “everyone was involved and enjoyed.” Four students indicated that the content was difficult at times but all felt they received enough help. Student 2 said “the teacher helped by joining discussions.” Student 4 indicated that “cooperation with other students” helped them cope with difficulties. All students said if the course was repeated, class time should incline toward content with percentages ranging from 60% to 100%.

**Discussion**

All students indicated their primary reason for joining the course was content related, and content was stated to be key to enjoyment throughout the course. When starting the course, all but one student thought that content should have more weightage of the class time. The one student who desired an equal breakdown changed their weightage toward content later in the course.

**Figure 2**  
*Study 2 Students’ Desired Class Time Focus*



All students felt they were given enough help, even though language was not taught explicitly. In fact, the flexible approach was noted and supported by four respondents. The content was felt to be challenging at times according to three respondents but it did not impede enjoyment. Satisfaction with the course was high and unanimous.

The results indicate the students felt positive concerning a harder approach to CLIL, and the course design met their learning goals and was enjoyable. There was also no indication that it was difficult for students to concentrate in lectures. However, this lack of indication

highlights a limitation of this study. The lack of clarity could be due to the decision to use a questionnaire. Dornyei and Taguchi (2010) argue that respondents often only superficially interact with questionnaires giving brief and surface level answers. If something is not asked directly, respondents are unlikely to address it. Future studies on CLIL classes can use interviews, which may be more effective in exploring student attitudes and experiences.

The question asking students to break down desired class time balance between language support and content seems flawed. Students are not teachers and may be unable to judge what is content and language support. Language support was embedded in slides presented before focusing on authentic material. However, students could understandably see the slides as pure content. If students are unable to understand how language content can be presented, it seems a desired breakdown would be superficial. An interview could further explore the dividing of class time between content and language.

Four respondents said some notions presented were difficult to understand but this did not impede enjoyment. However, it is unclear whether the lack of understanding impede comprehension of authentic materials. It seems a more exploratory research tool may be required to uncover more information.

## Conclusion

This paper has detailed the journey from being presented by Rikkyo University with a course to teach, choosing an approach to teaching, designing curriculum to gauging students' attitudes to decisions taken on course design. CLIL was chosen as a teaching method as although students were proficient, they would need language support to better comprehend authentic materials. A harder approach to CLIL was taken as it was assumed students would have chosen this elective class due to the content. They also would have already taken and have the opportunity to take more classes related to language learning.

Two courses were studied over two subsequent years. Results from both groups suggest the choice to choose harder CLIL was justified. Student goals were content based, and the content was the basis of their enjoyment and perceived success of the class. Students overwhelmingly believed class time should be weighted toward content focus over language support. Those who felt a more equal division would be preferable and explicit language support should be given moved toward a more content-focused balance and less teacher intervention over the course.

An intervention was made in the second study to address student claims that maintaining concentration in lectures was difficult. This assertion was not made by the second group. Rather than indicating positive evidence that the problem had been rectified, the study pointed to a flaw in the research design. It seems reasonable to suggest the more exploratory research method of interviews be used in any further research.

Although the course design decisions suited the groups in question, as a student indicated, such a flexible approach would not be effective with larger groups and certainly not groups with a lower level of English proficiency. This could be an area for further study, along with assessment design and the relation between language scaffolding in teacher presentation and comprehension of authentic materials.

Overall, I feel the investigation described has been a success both in terms of my understanding of CLIL course design and its suitability for the study groups. For university educators in Japan, CLIL is now an inescapable fact of life. An understanding of the Language Triptych, the 4Cs, and different approaches to CLIL will allow language educators to better present the content they want to highlight. Moreover, by using these tools, teachers can better satisfy students' language needs and learning goals, while providing enjoyable and informative classes.

## References

- Bradford, A. (2018). Undergraduate English-medium instruction in Japan: Policy and implementation. *Humanities and Social Sciences Bulletin Meiji University*, 64, 65–76.
- Brown, H. (2015). Factors influencing the choice of CLIL classes at university in Japan. *ELT World Online*, 1–22.
- Brown, H., & Bradford, A. (2017). EMI, CLIL, & CBI: Differing approaches and goals. In P. Clements, A. Krause, & H. Brown (Eds.), *Transformation in language education* (pp. 328–334). JALT.
- Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and language integrated learning*. Cambridge University Press.
- Dornyei, Z., & Taguchi, T. (2010). *Questionnaires in second language research: Construction, administration, and processing* (2nd ed.). Routledge.  
<https://doi.org/10.4324/9780203864739>
- González González, G. M. E., Diaz Robayo, D. R., & Leon Mora, E. C. (2022). A CLIL curriculum design for future professionals of hospitality and tourism management. *Latin American Journal of Content & Language Integrated Learning*, 15(1), e1513.  
<https://doi.org/10.5294/laclil.2022.15.1.3>
- Iyobe, B., & Li, J. (2013). CLIL to what degree: A trial in English medium education at a Japanese university – Is it CLIL or not? In P. Robertson, & J. Adamson (Eds.), *The Asian EFL Journal Special Edition CLIL in Asian Contexts: Emerging Trends*, 15(4), (pp. 372–381).
- Korosidou, E., & Griva, E. (2014). CLIL approach in primary education: Learning about Byzantine art and culture through a foreign language. *Study in English Language Teaching*, 2(2), 240–257. <https://doi.org/10.22158/selt.v2n2p240>
- Leontjev, D., & deBoer, M. A. (2022). Multimodal mediational means in assessment of processes: An argument for a hard-CLIL approach. *International Journal of Bilingual Education and Bilingualism*, 25(4), 1275–1291.  
<https://doi.org/10.1080/13670050.2020.1754329>
- Lockley, T. (2015). Promoting international posture through history as content and language integrated learning (CLIL) in the Japanese context. *Studies in Second Language Learning and Teaching*, 5(1), 87–108. <https://doi.org/10.14746/ssllt.2015.5.1.5>
- MacGregor, L. (2016). CLIL focus: An interview with Professor Makoto Ikeda. *The Language Teacher*, 40(1), 18–21. <https://doi.org/10.37546/JALTTLT40.1-4>
- Margana, M., & Widyanoro, A. (2017). Developing English textbooks oriented to higher

- order thinking skills for students of vocational high schools in Yogyakarta. *Journal of Language Teaching and Research*, 8(1), 26–38. <https://doi.org/10.17507/jltr.0801.04>
- Martin del Pozo, M. A. (2016). An approach to CLIL teacher language awareness using the Language Triptych. *Pulso. Revista De Educación*, 39, 141–157. <https://doi.org/10.58265/pulso.5082>
- Meyer, O. (2010). Towards quality CLIL: Successful planning and teaching strategies. *Pulso. Revista De Educación*, 33, 11–29. <https://doi.org/10.58265/pulso.5002>
- Mok, J. (2021). Ontological and pedagogical reconsideration of CLIL: Thoughts on concept, course design, and assessment. *Journal of Foreign Language Education and Research*, 2, 52–67.
- Morizumi, F. (2015). EMI in Japan: Current status and its implications. *Educational Studies, International Christian University*, 57, 119–128. <http://doi.org/10.34577/00003421>
- Nitta, R., & Yamamoto, Y. (2020). Reconceptualizing CLIL from transformative pedagogy perspective: Pilot debate study in English language curriculum. *Journal of Foreign Language Education and Research*, 1, 47–62.
- Ohmori, A. (2014). Exploring the potential of CLIL in English language teaching in Japanese universities: An innovation for the development of effective teaching and global awareness. *The Journal of Rikkyo University Language Center*, 32, 39–51.
- Otto, A., & Estrada, J. L. (2019). Towards an understanding of CLIL assessment practices in a European context: Main assessment tools and the role of language in content subjects. *CLIL Journal of Innovation and Research in Plurilingual and Pluricultural Education*, 2(1), 31–42. <https://doi.org/10.5565/rev/clil.11>
- Pinner, R. (2013). Authenticity of purpose: CLIL as a way to bring meaning and motivation into EFL contexts. *Asian EFL Journal*, 15(4), 49–69.
- Puspitasari, E. (2016). Classroom activities in content and language integrated learning. *Journal of Foreign Language Teaching and Learning*, 1(2), 1–13. <https://doi.org/10.18196/ftl.129>
- Rikkyo University. (2021). *Lecture and Discussion H*. [https://sy.rikkyo.ac.jp/web/preview.php?nendo=2021&t\\_mode=sp&template=t1&no\\_id=2106748&gakubu\\_id=14](https://sy.rikkyo.ac.jp/web/preview.php?nendo=2021&t_mode=sp&template=t1&no_id=2106748&gakubu_id=14)
- Smith, C., & Ssali, V. (2019). From Foundation course to text book: Introducing CLIL in a Japanese university. *The Journal of the Japan CLIL Pedagogy Association (J-CLIL)*, 1, 85–99.
- Tanaka, K. (2019). Content and language integrated learning (CLIL): Compatibility of a European model of education to Japanese higher education. *Meiji Gakuin Review. International & Regional Studies*, 54, 61–81.
- Teresa Ting, Y. L. (2011). CLIL... not only not immersion but also more than the sum of its parts. *ELT Journal*, 65(3), 314–317. <https://doi.org/10.1093/elt/ccr026>
- Thomas, K. (2021). Review of the book CLIL: Content and language integrated learning by D. Coyle., P. Hood., & D. Marsh. *Journal of Foreign Language Education and Research*, 2, 195–196.
- Tsuchiya, K., & Pérez Murillo, M. D. (2015). Comparing the language policies and the students' perceptions of CLIL in tertiary education in Spain and Japan. *Latin American Journal of Content and Language Integrated Learning*, 8(1), 25–35.

<https://doi.org/10.5294/laclil.2015.8.1.3>

- Valiukiene, R. (2017). How complex is CLIL vocabulary in the eyes of a subject teacher? *Knowledge – International Journal*, 19(3), 1021–1027.
- Yamamoto, Y., & Nitta, R. (2021). Action-orientated approach to curriculum development in CLIL courses: A theoretical and methodological framework. *Journal of Foreign Language Education and Research*, 2, 122–135.
- Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal*, 86(1), 54–66.  
<https://doi.org/10.1111/1540-4781.00136>

## Appendices

### Appendix A: Pre-Course Survey

1. Why did you choose to study this course?
2. What do you hope to gain from the course? What language skills (if any)? What content knowledge (if any)?
3. The materials in the class are authentic (meant for people with English as a first language). Do you think students should be given help by the teacher in class to prepare for understanding and participating in listening, reading and discussion activities.  
  
If so, what help should be given? E.g., key vocab given pre-task, focus on listening skills
4. In class time, what breakdown do you think there should be between focus on language (for comprehension and participation) and content (the subject being studied)? E.g., 40% language, 60% content.
5. Do you think students should be assessed on their language ability, content knowledge learned or a mix of the two? If a mix of the two, what do you think the breakdown should be?
6. How do you think the course will complement your other studies at Rikkyo University?
7. Any other comments?

### Appendix B: Mid-Course Survey

1. Are you enjoying the course so far? Why? Why not?
2. How would you assess the difficulty of the course? Why? Please explain.
3. Do you think you are being given enough help to comprehend listening and reading tasks and participate in discussions? If not, what kind of help would you like?
4. Would you like the teacher to concentrate more on language or content? What breakdown do you think there should be? E.g., 40% language, 60% content?
5. Do you think you are being prepared for the assessment? Please explain.
6. How do you think the course complements your other studies at Rikkyo University?
7. Any other comments?



### **Appendix C: Post-Course Survey**

1. Did you enjoy the course? How difficult was it? Please explain.
2. Did you feel the course was meaningful? What did you gain from it?
3. Do you think you are were given enough help to comprehend listening and reading tasks and participate in discussions? If not, what kind should have been given?
4. If the course was run again, do you think the teacher should concentrate more on language or content? What breakdown do you think there should be? E.g., 40% language, 60% content?
5. Did you think the assessment reflected the use of class time? Were you prepared enough? Please explain?
6. Do you think the course complemented your other studies at Rikkyo University?
7. Any other comments?

# To What Extent Does Form-Based Priming Account for the Mnemonic Effect of Phonological Patterns? A Preliminary Investigation

Michael Green

---

## Abstract

There is a growing body of evidence that suggests that phonological patterns in multi-word units may act as mnemonic devices and thus aid in the acquisition of L2 vocabulary. Some researchers state that learners' attention should be deliberately drawn to such patterns, though there are conflicting findings in the literature on this point. One explanation that has been posited for the mnemonic influence involves a form-based priming effect. If it can be shown that phonological patterns automatically aid lexical processing, the intervention to raise awareness of such patterns may not be necessary. To shed light on this issue, the experiment reported herein is a preliminary investigation into the extent to which form-based priming accounts for the mnemonic effects observed in the literature. A lexical decision task (LDT) was conducted with 24 English L1 speakers to ascertain if alliteration and assonance facilitate lexical processing. The data from the LDT do not support any strong claims that perceptual priming is the determining factor for the processing advantage, which indicates the need for explicit attention to the patterns. Limitations and further avenues for investigation are also discussed.

**Keywords:** *Form-based priming, Phonological patterns, Vocabulary acquisition*

## Introduction

It seems almost axiomatic that phonological patterns such as rhyme, alliteration<sup>1</sup>, and assonance<sup>2</sup>, can act as aids to memory; for instance, rhyme can be used to learn historical events (*In fourteen hundred and ninety-two, Columbus sailed the ocean blue*) or the number of days in the month (*Thirty days hath September, April, June and November...*). There is also a longstanding and widespread use of phonological patterns in advertising, slogans, and brand names (*Coca Cola, Kit Kat, and Tim Tams*), fictional characters (*Donald Duck and Wonder Woman*), nursery rhymes, and of course, poetry and prose. Research by Jusczyk et

---

1 Operationalized herein as the repetition, in two consecutive words, of the same consonant sound in an initial stressed syllable, e.g., *tall tree*.

2 The repetition of a vowel or diphthong sound in the prominent syllable of two consecutive words, e.g., *main gate*.

al. (1999) shows that even nine-month-old infants are already sensitive to alliteration in consonant-vowel-consonant syllables, which suggests that patterns of similarity are pertinent to building L1 vocabulary. Phonological patterns also emerge in early childhood wordplay, reflecting the ludic or playful function of language (Crystal, 1997), and in Peters’ (1983) seminal account of first language acquisition, phonological patterns are posited as heuristic devices that young children employ to segment or compare “units” of meaning in the intermittent stream of speech. The literature on oral traditions and the genres of epics and ballads (for a comprehensive survey, see Rubin, 1995) proposes that phonological patterns, together with meaning and imagery, are constraints that cue memories and restrict choices, enabling thousands of lines of songs, stories, and poems to be memorized and transmitted for centuries.

In the field of second language acquisition, Nation (2001, 2014) notes that English multi-word units (MWUs), such as binomials, collocations, compounds, and idioms, often alliterate or assonate and advises L2 learners to pay deliberate attention to these patterns, advice that has been repeated more recently by Szudarski (2017). This advice is partially predicated on empirical work from a single group of researchers (Table 1). This series of small-scale, classroom-based quasi-experiments all explore whether English L2 learners can identify or recall MWUs with phonological patterns better than equivalent<sup>3</sup> sequences with no such pattern. In general, the experiments involve teacher-led dictations of target MWUs followed by a series of free- and cued-recall tests over various time intervals (immediate tests and delayed tests up to two weeks). The designs are generally iterative, involving a partial replication of a previous experiment, and collectively, these studies build the case that phonological patterns in MWUs have a mnemonic effect on L2 learners to some degree.

**Table 1**  
*An Overview of the Findings in the Experiments of Boers et al. on Phonological Patterns*

	<b>Study (<i>phonological pattern</i>)</b>	<b>Awareness-Raising?</b>	<b>Statistically Significant Finding?</b>
1	Lindstromberg & Boers (2008a) ( <i>alliteration</i> )	Exp 1	Yes
		Exp 3	Yes
2	Lindstromberg & Boers (2008b) ( <i>assonance</i> )	Yes	Yes
3	Boers et al. (2012) ( <i>alliteration</i> )		Yes
4	Boers et al. (2014a) ( <i>alliteration</i> )	Exp 1	Yes
		Exp 2	Yes
5	Boers et al. (2014b) ( <i>assonance</i> )	Exp 1	
		Exp 2	Yes
6	Boers et al. (2014c) ( <i>consonance</i> )		Yes <sup>1</sup>
7	Boers, Eyckmans & Lindstromberg (2014) ( <i>alliteration and consonance</i> )	Yes	

<sup>3</sup> Experimental items are often balanced in terms of length, frequency, collocational strength (with reference to MI or t-scores), concreteness and imageability, and L1 cognate status.

Study ( <i>phonological pattern</i> )		Awareness-Raising?	Statistically Significant Finding?
8	Lindstromberg & Eyckmans (2014) ( <i>assonance</i> )	Exp 1	Yes
		Exp 2	Yes
9	Boers et al. (2014) ( <i>alliteration</i> )		Yes
10	Eyckmans et al. (2016) ( <i>alliteration</i> )	Yes	Yes
11	Eyckmans & Lindstromberg (2017) ( <i>alliteration and assonance</i> )	Yes	Yes
12	Lindstromberg & Eyckmans (2017) ( <i>assonance</i> )	Yes	

<sup>1</sup> A statistically significant finding, but not in the predicted direction.

An issue that arises is whether participants need to have the phonological pattern highlighted to benefit from any mnemonic effect. As can be seen from the *Awareness-Raising* column in Table 1, there is inconsistent evidence on this point; statistically significant results have been found when there is no attention-raising (e.g., Studies 3, 4, and 9), and non-significant results found when the patterns have been brought to the attention of the participants (e.g., Studies 1(3), 7, and 12).

One possible explanation for the facilitation of lexical retrieval of MWUs displaying phonological patterns is the short-term memory process of priming, an explanation that the researchers have posited in the past (e.g., Boers et al., 2014a; Eyckmans et al., 2016; Eyckmans & Lindstromberg, 2017; Lindstromberg & Boers, 2008a; Lindstromberg & Eyckmans, 2017). If priming is the principal mechanism behind which the mnemonic effect operates, no teacher intervention would be necessary in theory as priming is held to be an implicit memory process and is thus often described as unintentional or unconscious (Tulving & Schacter, 1990). To further tease apart this issue and to see if an awareness-raising component is indeed warranted, a lexical decision task (LDT) was performed to assess the plausibility of priming as an explanatory factor for the mnemonic advantage of alliterating or assonating MWUs, as seen in the experiments outlined in Table 1.

## Literature Review

The term “priming” appears to do a lot of heavy lifting in the literature: it can refer to a *component* of implicit long-term memory (Baddeley, 2001) or a research *paradigm* or *tool* for examining lexical processing (McDonough & Trofimovich, 2009). Moreover, Shallice and Cooper (2011) variously employ the term “priming” to cover a functional imaging *procedure*, a short-term memory *process*, a general *characteristic* of cognitive subsystems, and a *property* of processing. Distinctions between these disparate usages are not always clearly articulated. Such terminological diffusion may have arisen partly because priming has been the focus of research across several disciplines, notably neuropsychology, cognitive psychology, psycholinguistics, and, increasingly, second language acquisition studies.

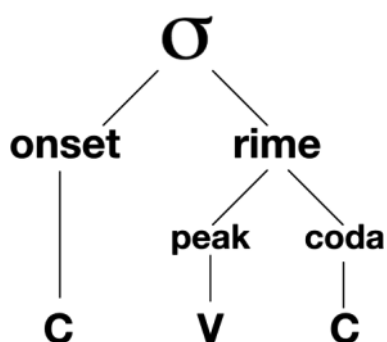
Broadly speaking, “priming” is used here as a term to describe how prior experience

with language can unconsciously influence the processing of subsequent language. One reason phonological patterns such as alliteration and assonance are thought to be mnemonic devices is that form-based aspects of a given word privilege the processing of a following word that shares the same form-based aspects, whether they are phonological, orthographic, or phonetic.

The literature on form or perceptual priming is extensive, so when looking for evidence that alliteration or assonance can aid lexical processing, it may help first determine how such phonological patterns are operationalized within the priming paradigm. In effect, the question of interest is whether hearing a word beginning with or containing a particular sound will make it easier to process a subsequent word with the same sound in the same position. The form-based similarity of a prime and its target is frequently manipulated in terms of matching or overlapping particular segments of the two words and/or manipulating the number of segments that match. This segmentation is often based on a model of syllabic division (Figure 1).

**Figure 1**

*Segmentation of a Consonant-Vowel-Consonant Monosyllable*



Thus, a facilitatory priming effect for alliteration would involve test items with matching phonological consonantal onsets and no other relationship—morphological, syntactic, or semantic—as they could themselves account for any observed priming effect. Likewise, a priming effect for assonance would be based only on the single aspect of a matching stressed vowel sound (the peak or nucleus).

However, in Zwitserlood’s (1996) overview of the form priming experimental paradigm, the stimuli overlap is described as “word-initial or rhyme” (p. 590), with no scope for a priming effect based on assonance alone. Rhyme is usually operationalized as phonological similarity of both the syllabic peak and coda (the rime), or in other words, similarity starting from the stressed vowel sound to the word offset. As we shall see below, perceptual priming based on assonance alone has received very little attention. As the findings for word-initial similarity are inconsistent, those based on rhyme will be considered first. Indeed, much of the most robust evidence is found in studies where the prime-target relationship is that of rhyme. Thus, participants will identify the target letter string *BEAN* more accurately and faster after being exposed to the prime *mean* than after the prime *pink*. This systematic

priming effect for rhyme has been found in both monosyllabic and bisyllabic prime-target pairs, irrespective of frequency, for both words and pseudowords, and independently of the task performed: LDTs (Norris et al., 2002), shadowing (Dumay et al., 2001), and identification in noise (Slowiacek et al., 1987).

In contrast to the evidence that rhyme can prime, even under different experimental designs, the situation for a priming effect for word-initial similarity is more complex and inconsistent, with facilitatory, inhibitory, and null effects being reported. Although the same dependent variable of response latency is usually the focus across experimental designs, task differences can introduce different variables, so it may be helpful to focus on just one task, the LDT, to glean insight from the results. When looking at word-initial overlap that forms an alliterative relationship between the prime and the target, a set of “rather messy results” (Dumay et al., 2001, p. 121) can be found in the literature, as seen in five representative studies in Table 2.

**Table 2**

*Examples of LDT Studies that Include Alliterative Word-Initial Overlap*

Study	Examples of <i>primes: TARGETS</i>	Effect <sup>1</sup>
1 Slowiacek & Pisoni (1986) (Experiment 1)	<i>black, bland, bleed, burnt, /blæt/, /blim/, /brɛm/; BLACK</i>	No effect <sup>2</sup>
2 Radeau et al. (1989) (Experiment 2)	<i>palais, poulet, rouler: PARURE</i> (French L1 speakers)	No effect <sup>2</sup>
3 Goldinger et al. (1992) (Experiment 3)	<i>bang: BONE</i>	Facilitation <sup>3</sup>
4 Monsell & Hirsh (1998) (Experiment 1)	<i>broom: BRUISE</i>	Inhibition
5 McQueen & Sereno (2005)	<i>zeep: ZOON</i> (Dutch L1 speakers)	No effect

<sup>1</sup> Slowiacek and Hamburger (1992) propose facilitation is due to activation/excitation at a prelexical phoneme level, and inhibition is the result of competition between words at the lexical level.

<sup>2</sup> A facilitatory effect, in terms of faster reaction times, was found only when the prime and the target were identical (a repetition effect).

<sup>3</sup> An auditory priming technique was used, and a facilitatory effect was found only when the targets were presented in white noise.

From the examples given in Table 2, it is clear that the nature of some of the prime-target relationships goes beyond a simple alliterative pattern and encompasses a matching vowel nucleus (e.g., broom – BRUISE in Monsell & Hirsh, 1998).

In Studies 1 and 2, no priming effect was found when primes shared one, two, or three phonemes with the target. The use of real words as primes further confounds the issue as the relationship between prime and target is no longer purely form-based: for example, in Study 1, if *burnt* had shown a priming effect for *BLACK*, it might have been due to semantic

association rather than form-based similarity.<sup>4</sup>

It might seem that the best way to avoid semantic priming would be to use pseudowords. However, these too can create difficulties, and the construction of the pseudowords may also be a factor in the lack of clear results. For example, in Study 1, some of the pseudoword targets are pseudohomophones (a nonword that sounds like a real word), such as /stik/ /skot/ /bæns/ and /slæk/. Jiang (2012) advises against the use of such items as several studies have shown that pseudohomophones generally take longer to reject than pseudowords, and these could affect the mean response times (RTs) in the different conditions.

The experiments in Table 2 also incorporated different timing elements in the designs, specifically the interstimulus intervals, the interval between the offset of the prime and the onset of the target, and differences in stimulus onset asynchrony, the interval between the onset of the prime and the onset of the target. Lengthening these two variables is often claimed to increase strategy use (Radeau et al., 1989). Although none of these was itself uncontrolled within its study, they are all potential variables determining the opportunity for an observable effect, and between them, these variations dilute the strength of evidence for initial consonant priming.

The failure to find a facilitatory effect with reaction time measures could be due to differences in methodologies, dependent measures, and/or stimuli, and as reiterated by Goldinger et al. (1992), null results do not support any definitive conclusions. Rastle and Brysbaert's (2006) exhaustive meta-analysis of the relevant priming literature concluded that phonological primes *do* have an effect, and some experimental results cannot be attributed to task-based biases or strategy use. However, the evidence for a priming effect based solely on alliteration is ambiguous. Furthermore, as mentioned previously, there is scant data to make any reliable assumptions about an equivalent effect for an assonating pattern.

## Methodology

### Experiment Design

One challenge facing the researcher is that although priming experiments have a long history in some form or another, “research labs accumulate a lot of informal knowledge about how to run particular [priming] experiments, which is rarely published” (Rastle & Brysbaert, 2006, p. 185). As a result, it can be difficult to ascertain exactly how such experiments are designed and implemented from previously published studies alone. Fortunately, for one of the standard methods used in priming research, namely reaction time studies, there exists a body of more prescriptive work with more explicit guidelines on the protocols (Jiang, 2012; McDonough & Trofimovich, 2009). Though there are a wide range of experimental tasks for measuring and quantifying priming effects, the most widely used for investigating the mental lexicon is the LDT. In a standard LDT, two stimuli are presented successively on a computer screen, first the “prime” and then the “target.” The experimental task requires the participant to respond only to the target by either pressing the “Yes” button on a keyboard if they think

---

4 Aside from the *burnt-BLACK* example given in the paper, the authors only list the target items in the appendix, not the items used as primes; hence, it is not possible to ascertain the overall number of potential semantic associations between the test items.

the target is a real word or the “No” button if they decide the target is not a real word. Participants are asked to respond as fast and as accurately as possible to each target. The reaction times, mediated by the participant’s motor responses, are thought to provide indirect evidence for underlying cognitive processes. Priming is said to occur if the prime facilitates the response to the target in terms of faster reaction times (measured in milliseconds) and increased accuracy.

Facilitation can occur because of a conceptual or semantic relationship between the prime and the target: for example, participants react faster to the visual string *NURSE* if they have just been presented with *doctor*. However, as this experiment seeks to establish if there is a facilitatory effect due solely to a perceptual or form-based relationship between the prime and the target, semantic variables need to be controlled for (discussed presently).

As priming is thought to be an implicit process that precludes the use of strategies in language processing, the initial visual stimuli can be hidden or “masked.” In the priming research paradigm, masking is often accomplished in two ways. Firstly, the prime stimulus is presented so briefly on screen (e.g., 50–60 ms) that the participant is unaware of it. Secondly, a series of characters, often a string of hash keys as long as the target string (#####), can be used as a forward mask preceding the prime, and the target word-string used as a backward mask (Jiang, 2012). These masking techniques are thought to block retinal after-images and pixel overlap, overwriting any visuo-sensory representations. Such steps help ensure that the participants cannot employ conscious strategies such as attempting to guess upcoming targets, trying to find the relationship between the primes and targets, or tuning in to experimental “bias” such as word length for primes and targets.

Before outlining the procedures adopted for this study, a further challenge needs to be addressed. There is growing empirical evidence that the bilingual lexicon is non-selective (Brysbaert, 2003; Nakayama et al., 2012). This premise means that the automatic activation of a phonological representation by a visual word stimulus is not limited to a representation specific to the language being read. This leads to the conclusion that if a participant has knowledge of more than one language—and if, as supposed, the first stages of word recognition are indeed language independent—it is possible to prime a target word in the L2 by a homophonic stimulus of the L1. There is an increasing amount of support for this notion in the literature; for example, for French-Dutch bilinguals, a French target (such as *OUI*) can be primed by a phonologically similar Dutch word (*wie*) (example from Experiment 1 in Brysbaert et al., 1999). There is also evidence of cross-language cognate priming across disparate writing systems. For example, Nakayama et al. (2012) found significant priming effects with Japanese-English bilinguals, where a masked stimuli such as ガイド /<sub>μ</sub>ga<sub>μ</sub>I<sub>μ</sub> dɔ̃/ (guide) primed the English target *GUIDE*.

These findings raise four intriguing questions. Firstly, is this priming effect uni- or bi-directional? That is, would the L2 stimulus *guide* prime the L1 target ガイド (/<sub>μ</sub>ga<sub>μ</sub>I<sub>μ</sub> dɔ̃/) for Japanese-English bilinguals? Secondly, what is the role of proficiency? Are the same priming effects found with less-proficient bilinguals such as adult L2 learners? Thirdly, does the priming effect only apply to alphabets and syllabaries or also to logographic scripts such as Japanese Kanji? Finally, is there a cross-language priming effect when the primes and targets are not cognates and therefore have no semantic relationship? For example, with



Japanese-English bilinguals would the L2 prime *guy* facilitate the processing of the L1 *ガイ* ト /<sub>μ</sub>ga <sub>μ</sub> <sub>μ</sub> dɔ/ and vice versa?

These issues would seem to suggest that it is thus paramount to rigorously control for pseudohomophone and cognate status in a cross-language LDT. However, numerous other lexical characteristics can also affect the processing of stimuli in such tasks and thus affect the participant's reaction time and accuracy. According to Jiang (2012), the following properties have all been found to affect lexical processing:

frequency, familiarity, word length (in terms of letter, syllables, phonemes or morphemes), neighbourhood density, neighbourhood frequency, concreteness, imageability, age of acquisition, spelling-sound regularity, affixation, polysemy, bigram frequency, number of associates, lexicality, nonword legality, pseudohomophone and cognate status (pp. 80–82).

A further challenge is that inconsistent results are not uncommon in reaction time studies, and the exact effect of these variables is not without controversy (Jiang, 2012). Furthermore, different theoretical positions can complexify the explanation of apparently simple effects. For example, word frequency is often considered one of the most robust predictors of word recognition performance. It has long been known that participants respond more quickly and accurately to high-frequency than low-frequency words across virtually all lexical processing tasks (Whaley, 1978, cited in Yap & Balota, 2015). Although there are various potential models that explain the word frequency effect, most reflect a version of frequency affecting baseline activation. For instance, according to the interactive activation model (McClelland & Rumelhart, 1981), high-frequency words are responded to faster because they have a higher resting activation level (or lower threshold) and thus require less stimulus information to be recognized.

Considering the various complications of LDTs, two key decisions were taken to attempt to minimize uncontrolled complexity. Firstly, to avoid the interaction of a different language as L1 with the language of the experiment, it was decided to conduct this experiment on English L1 speakers. Secondly, it was decided to use pseudowords, that is, strings of letters that conform to legal English orthographic and phonological constraints,<sup>5</sup> for example, *nirk*. By using pseudoword primes, it is thus possible to restrict the prime-target relationship to form-based variables only and avoid semantic priming effects. Of course, even English L1 speakers might encounter priming from another language they know, but as they will perceive the experiment to engage with their L1, it should minimize the effect to background noise. Secondly, it is only the masked primes that are pseudowords, whereas the targets are the necessary mixture of real words and pseudowords. As a result, this experiment should come across as fully about the participants' L1.

## Participants

Data was collected from 24 English L1 speakers, 20 from the US, 3 from the UK, and

<sup>5</sup> The term “pseudoword” is used rather than “nonword” as the latter can refer to an illegal letter string, such as *mgfa*.

1 from Canada (average age 24 years and 8 months, 15 females and 9 males). The participants were all university students in an exchange program at a private university in Japan. A self-report questionnaire indicated a low-intermediate knowledge of Japanese. The participants had normal (or corrected-to-normal) vision and no history of language impairment. The questionnaire also asked if the participants were left- or right-handed. The participants provided informed consent and received a raffle ticket to win a pair of headphones in exchange for their participation. Data was collected over several weeks of university term time. Although 24 participants may appear to be a small sample, Jiang (2012) notes that many reaction time studies are done with similar sample sizes as increasing the number of participants has little effect on the RT data.

### **Compiling the Stimuli Items for the LDT**

In an LDT, two equally sized sets of stimuli are needed for the target items: a set of real words and a set of pseudowords. The inclusion of pseudowords is a necessary part of the design because it prevents the participants from simply pressing the “Yes” button to every single target and developing a bias of responses.

For the set of primes, pseudowords were used to avoid any semantic priming effects. Furthermore, in a task aiming to establish the impact of a feature on how quickly words are recognized, there need to be subsets of words with and without the feature. In this experiment, the set of real word stimuli needed to consist of words used to test the effects of both alliteration and assonance and a control set that had neither feature. The following section outlines how the real words were chosen for the LDT and how the pseudowords were generated for the primes and the “No” response targets.

### **Real Word Stimuli (Targets)**

As adjective-noun collocations and noun-noun compounds were the most commonly used stimuli in the work of Boers et al. (Table 1), it was decided that nouns would also be used as the real word stimuli in the LDT. In an attempt to control for possible frequency, familiarity and length effects, a pool of words was assembled, which comprised 290 monosyllabic nouns of 4–5 letters in length, from the first two frequency sub-lists of the British National Corpus/Corpus of Contemporary American English (BNC/COCA) corpus (Nation, n.d.). Although it is unlikely that any corpus matches the linguistic experience of any particular speaker exactly, this approach was deemed valid in that corpora are representative of the types of input the English L1 participants are likely to have encountered.

The items were then cross-referenced with the Medical Research Council (MRC) database (Coltheart, 1981) for concreteness and imageability ratings. This database contains up to 26 linguistic and psycholinguistic attributes for over 150,000 words collated from previously published sources. The words in the database have subjective ratings on an integer scale from 100 to 700. There were 24 nouns in the initial pool that had no ratings in the MRC database, so these were eliminated from the pool (e.g., *bike, dish, farm*). The mean and standard deviations were calculated for concreteness and imageability of the remaining 266 nouns. This dataset was then trimmed of any nouns that had ratings of more than  $\pm 2$  SD from the mean. This resulted in the removal of 11 more nouns (e.g., *bunch, place, thing*),

with the aim of producing a range of concreteness and imageability as narrow as possible. The remaining group of 255 potential stimuli had mean concreteness ratings of 576.9 (SD 36.7) and mean imageability ratings of 572.2 (SD 36.1). In comparison, the mean concreteness and imageability ratings in the MRC database are 438 and 450, respectively, suggesting the target stimuli for the LTD were slightly more concrete/imageable than the database average.

The pseudowords in an LDT play two different roles: as primes and as targets for the “No” responses.

### **Pseudoword Stimuli (Primes)**

The nature of an LDT, in which participants have to make rapid decisions as to whether the visual letter string is a real word, means that the characteristics of the pseudowords become an essential part of the experimental design. Not all pseudowords are the same. There is evidence that the types of pseudowords used in a priming experiment have a strong effect on reaction time performance. Keuleers and Brysbaert (2010) state that “the more dissimilar the nonwords are to the words [in the task], the faster are the lexical decision times and the smaller is the impact of word features such as word frequency, age of acquisition, and spelling-sound consistency” (p. 627). Clearly, features such as frequency and age of acquisition apply only to real words, though the form of the pseudoword can affect the speed of all decision processes in an LDT. In Gibbs and Van Orden (1998), for example, mean reaction times to reject nonwords were the shortest when the nonwords were illegal letter strings (e.g., *ldfa*: 496 ms), longer when the stimuli were legal letter strings (e.g., *dilt*: 558 ms), and longer still when the nonwords were pseudohomophones sounding like real words (e.g., *durt*: 698 ms).

To construct matching pseudowords for the 255 nouns, the Wuggy pseudoword generator was used (Keuleers & Brysbaert, 2010) to create monosyllabic stimuli of equivalent length (4–5 letters) with sound-spelling consistency. The Wuggy algorithm produced 10 candidate pseudowords for each real word noun. As words that are orthographically similar to many other words are recognized faster (Yarkoni et al., 2008), neighborhood size and density plays an important role when developing stimuli lists. The Wuggy algorithm uses the Levenshtein edit distance of orthographic similarity (OLD20), where similarity includes neighbors generated by insertion, deletion, substitution, and transposition of letters to produce perceptually similar words. By referring to the OLD20 values in the Wuggy output, it was possible to select the pseudowords that remained as close as possible to real words, that is, with OLD20 values as close to 1.0 as possible. An OLD20 value of 1.8 was arbitrarily set as a cut-off point for pseudoword candidates.

The ideal pseudoword candidate would not cause any inadvertent triggering of semantic or associative representations in a participant’s mental lexicon. This proved challenging to control for as a considerable proportion of the 2,550 pseudowords could be construed as a proper name, brand name, acronym, pseudohomophone, or informal, archaic, or variant spelling of a common noun (examples to follow). Moreover, closer scrutiny of the Wuggy output unearthed several words attested in online dictionaries (such as *blog*, *mage*, and *bling*).

The pseudowords were cross-referenced with the COCA database (<http://www.americancorpus.org>). Any items that appeared as proper names with more than 10 instances

in the corpus were discounted (e.g., *crowe*, *imes*, *gide*). Similarly, any items that occurred more than 10 times as acronyms (*pacs*, *facs*, *mact*) were also deleted. Pseudowords were also discounted if they appeared in the corpus as brand names (such as *fage*), abbreviations or variant spellings (e.g., *hols*, *dept*, *lite*, *nite*), or slang (*crip*, *shart*).

The Google search engine was then used to check if the remaining pseudowords elicited any dictionary references, especially in Merriam Webster, as participants were most likely to be from a North American background given the educational context; this led to two more deletions (*thang*, *fleed*). Finally, all pseudohomophones were deleted (e.g., *bocs*, *coaks*, *ceal*, *bild*, *tode*, *wead*).

The experimental design meant that for each of the 255 noun targets, there needed to be one pseudoword prime in each condition: alliteration, assonance, and no phonological pattern. However, after removing all pseudoword candidates with OLD20 values higher than 1.8, pseudohomophones, and those attested in corpora or dictionaries, some nouns did not have equivalent pseudowords in all three conditions; hence, these nouns were removed as test items. Their pseudoword equivalents that met the criteria were retained to use as filler items (see below).

This process of elimination resulted in a final pool of 102 target nouns, each with one alliterating pseudoword prime, one assonating pseudoword prime, and one pseudoword prime with no phonological/orthographic overlap, an example of which can be seen in Table 3.

**Table 3**

*Example of a Stimulus Item Across Three Conditions*

Prime	Target	Condition
<i>mouch</i>	<i>MILK</i>	Alliteration
<i>skift</i>	<i>MILK</i>	Assonance
<i>drate</i>	<i>MILK</i>	Control

One limitation of the Wuggy algorithm is that it does not generate the pronunciation for the pseudowords. To check that the prime-target pairs did alliterate and assonate, three English L1 speakers were asked to read aloud the list of items to check the researcher's intuitions of phonological similarity. No discrepancies were found in this regard.

### **Pseudoword Stimuli (Targets)**

To avoid response bias in the LDT, there had to be an equal number of "Yes" and "No" responses. Without such filler items, participants would only need to press the "Yes" key to produce a correct response. Furthermore, such filler items help disguise the critical stimuli, so it is less likely a participant will notice and consequently develop a processing strategy while performing the task. A total of 102 pseudoword targets, which would hopefully elicit a "No" response on the decision task, were selected from the remainder of the Wuggy output. For the 102 pseudoword primes, the mean OLD20 value was 1.34 (min. 1.0, max. 1.7, SD 0.22) and the mean value for the 102 pseudoword fillers was also 1.34 (min. 1.0, max. 1.75, SD 0.24).

In sum, there were 102 noun targets, designed to elicit a “Yes” response in the LDT. The following variables were controlled for frequency, familiarity, word length, neighborhood density, concreteness, and imageability. Each noun target had three pseudoword primes, one in each condition. There were an additional 102 pseudoword targets designed to elicit a “No” response. The primes for these filler items were chosen from the remainder of the Wuggy output. The pseudowords complied with the phonotactic constraints of English and were controlled for in terms of length and orthographic similarity.

### **Setting up the LDT Program**

The LDT was done using the free stimulus presentation software DMDX<sup>6</sup> (Forster & Forster, 2003). An item file was created in a Rich Text Format, which basically instructed the software on what items to present and how to present them. This allowed the 102 nouns plus 102 pseudowords to be combined into 204 trials, which were then divided into 17 blocks. Each block contained two alliterating prime-target pairs, two assonating prime-target pairs, two prime-target pairs with no form overlap, and six pseudoword filler items. Items were scrambled within each block and blocks were also scrambled for each participant. This pseudo-randomization avoided long successions of words or pseudowords appearing by chance. It also ensured extraneous serial effects (such as practice or fatigue) were more evenly distributed across conditions, and no two subjects were likely to receive the same sequence of items. Moreover, it avoided the introduction of systematic errors of measurement arising, when difficult items (on which an error is likely) can affect the reaction time for following items. The three conditions were counterbalanced across three presentation lists in a Latin square design, such that in each list, one third of the stimuli (34 items) appeared in each condition. This allowed for direct comparisons across conditions and avoided list effects. Participants were randomly assigned to one of these lists and responded to an equal number of trials in all three conditions, but never responded to the same target more than once. Two versions of the item file were created to allow for different keys being allocated to “Yes” and “No” responses, one for right-hand dominant participants and one for left-hand dominant participants. The experiment was run on a PC with Microsoft Windows XP (1920×1080 resolution, 60 Hz refresh rate). TimeDX, a component of the DMDX software suite, was used to verify that the hardware features of the computer were satisfactory.

### **Procedure**

Participants were given written instructions (adapted from those used in Ellis et al., 2008) in which they were told that they were participating in a vocabulary experiment and would see a letter string on the computer screen; they were then required to press the key labeled “Yes” if they thought it was an English word and the key labeled “No” if they thought otherwise. Participants were asked to respond as quickly and accurately as possible. Participants were informed neither of the presence of the masked stimuli nor the three critical conditions.

The participants were tested individually and sat approximately 50 cm in front of the

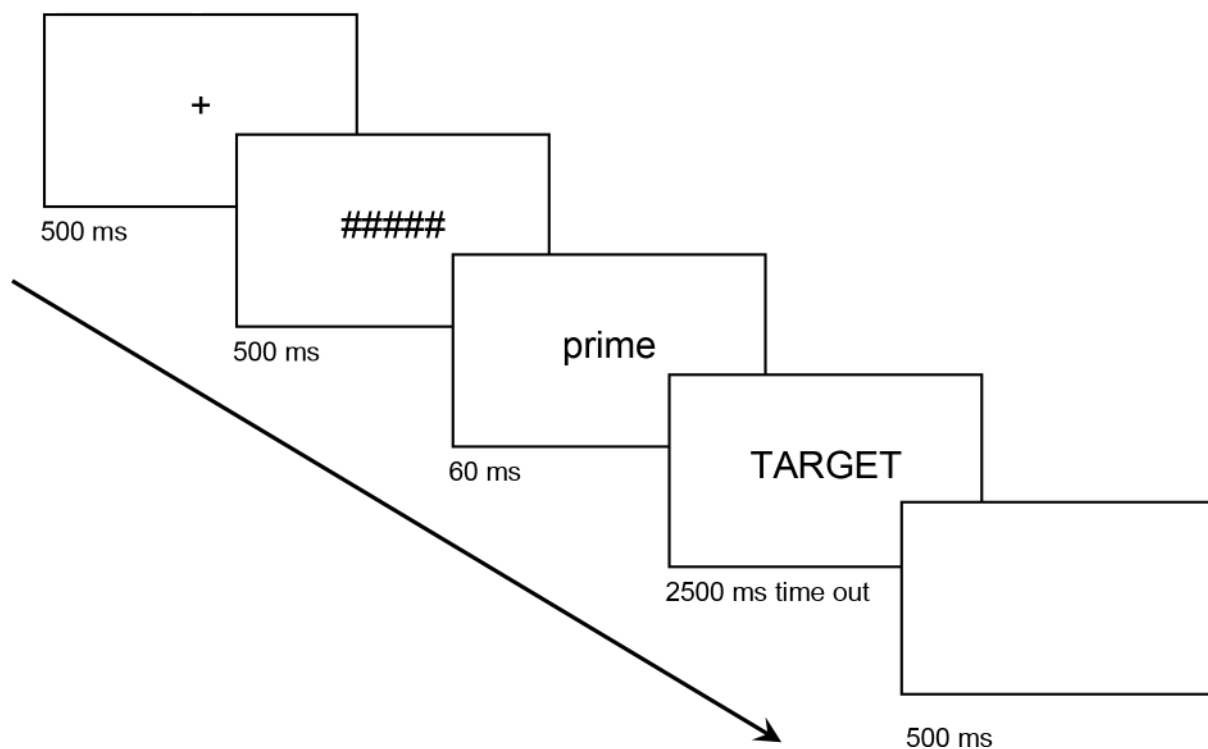
---

6 Available at <http://www.u.arizona.edu/~kforster/dmdx/download.htm>

eye-level computer screen, wearing noise-cancelling headphones in a quiet and dimly lit room. The index finger of their dominant hand was assigned to the “Yes” key and their other index finger to the “No” key. The experiment began with 10 practice trials, followed by the 204 experimental trials. There was a programmed break in the middle of the experiment for participants to rest, and the experiment resumed when participants pressed the space key. Trials were randomized for each participant, and the experimental phase took approximately 17 minutes. The stimuli were presented in the center of the computer screen in black Arial font, size 14, on a light gray background using the DMDX software. An explanatory schema of a trial can be seen in Figure 2.

Each trial began with the presentation of a fixation point (+) for 500 ms to direct the attention of the participant to the appropriate location on the screen. A forward mask (a meaningless string of five hash marks) appeared for 500 ms centered at the same location, followed by the presentation of the prime for 60 ms, an amount of time deemed safe for the item to be processed but without the participant becoming aware of it. Then, the target appeared on the screen, which served as a backward mask for the prime. The target remained on the screen until the participant’s judgment or a 2500 ms timeout, followed by a 500 ms inter-trial interval (a blank screen). RTs were measured from target onset until the participants’ response on the keyboard. After the experiment was completed, participants were asked to complete a short exit questionnaire to self-report on how well they followed task instructions and whether they were aware of the primes.

**Figure 2**  
*Sample Trial Schema*



## Results

### Data Analysis

The raw data from the response latencies and associated accuracy were saved as a data file and entered into the *Analyze* program, another component of DMDX, for which an input specification file was written. This text file instructs the program how to treat the raw data and which item goes to what condition. The script parameters also specify how outliers and incorrect responses are to be processed, as described below.

Outliers, defined as responses less than 200 ms or more than 3 standard deviations above the participant's mean, were replaced by the mean value for that participant. This appears to be standard practice in an LDT (McDonough & Trofimovich, 2009). It eliminates the need to perform a log transformation on the subsequent dataset and ensures a normal distribution, thereby safeguarding from the unwanted effects of outlier response latencies.

The script parameters instructed the program to exclude incorrect responses from the analysis. They also instructed the program to identify and automatically remove items that generated error rates of 40% or more, together with the counterpart items in the other two conditions so that test materials remained matched across conditions. The file also instructed the program to reject data from participants with an error rate of 20% or over.

### Results

Inspection of the completed data showed that no items generated error rates of 40% or more, so no items were eliminated on these grounds. No participants had error rates of 20% or over. The mean error rate per participant was 3.79% (SD = 2.40), resulting in the removal of 186 incorrect responses. A total of 55 responses qualified as outliers and were replaced by the mean value for that participant (the mean number of outliers per participant was 2.29, SD = 1.23). Furthermore, in the exit questionnaire, none of the participants reported seeing the primes, and all the participants indicated that they had carried out the task exactly as instructed.

A paired-samples t-test was conducted to compare the reaction times in the real word and pseudoword conditions. There was a significant difference in the scores for real word (M = 566.14 ms, SD = 33.20) and nonword (M = 671.8 ms, SD = 76.32) conditions;  $t(23) = 6.7$ ,  $p < 0.001$  (two-tailed). These results suggest that the experiment design and choice of pseudoword matches were sound in that words and pseudowords elicited the expected responses, which is in line with Gibbs and Van Orden (1998).

To recap, the research question is whether phonologically patterned prime-target items will result in faster reaction times than prime-targets in the baseline condition with no phonological overlap. Descriptive results from the LDT can be seen in Table 4.

**Table 4**  
*Descriptive Results*

	Mean Reaction Time (RT) <sup>1</sup>	SD
Alliteration	571.01 ms	34.45
Assonance	585.46 ms	39.07
No Phonological Pattern	575.05 ms	39.92

<sup>1</sup> in milliseconds (ms)

Table 4 shows that alliterating prime-target items produced the fastest mean reaction times and assonating prime-target items the slowest mean reaction times. A repeated-measures ANOVA was used to analyse the results with RTs as the dependent variable and alliteration, assonance, or no pattern as the categorical independent variable. Preliminary examination of the data satisfied the assumptions of homoscedasticity and normality of distribution, and Mauchly's test of sphericity showed that the assumption of sphericity was not violated ( $\chi^2(2) = 2.93, p = 0.86$ ). There was a significant main effect of phonological pattern on the reaction times ( $F_{2,46} = 7.04, p = 0.002, \eta_p^2 = 0.23, \text{power} = 0.91$ ).

Pairwise post hoc comparisons using t-tests and Bonferroni corrected levels of significance showed that participants' RTs were significantly faster in the alliteration condition compared to the assonating condition (mean difference = 14.45,  $t = 3.69, p = 0.004$ ) with a medium effect size (Cohen's  $d = 0.76$ ), according to Plonsky and Oswald's (2014) interpretation. However, there was no statistically significant difference between reaction times in the alliteration condition and reaction times in the baseline condition (mean difference = -4.03,  $t = -1.06, p = 0.90, d = 0.22$ ), nor was there a statistically significant difference between the assonance reaction times and the no phonological pattern reaction times (mean difference = 10.41,  $t = 2.49, p = 0.06, d = 0.51$ ).

Thus, to answer the research question, phonologically patterned prime-target items did not lead to faster reaction times than prime-targets in the baseline condition with no phonological overlap.

## Discussion

The experiment sought to test the hypothesis raised by Boers et al. that MWUs displaying alliteration and assonance have a processing advantage compared to equivalent MWUs with no such phonological or orthographic overlap. Previous priming experiments have found reliable evidence that rhyme produces a priming effect, though the evidence is inconclusive for primes and targets that alliterate, and there is scant data on assonance alone. The results reported here do not contribute to evidence of a facilitatory priming effect with a dataset that used masked pseudoword primes and noun targets. The absence of a statistically significant result could be a consequence of a methodological shortcoming or it could be evidence of the null hypothesis, that is, alliteration and assonance (as operationalized in this experiment) do not facilitate lexical processing.

The experiment used a standard LDT with prime-target items comprising pseudoword



primes and monosyllabic high-frequency nouns in three conditions. No individual items generated large error rates, and the participants' mean error rate (5.6%) is in line with other published masked LTDs (e.g., a 6% mean error rate in Perea et al., 2015). The number of participants (N = 24) was deemed appropriate (Jiang, 2012) as was the number of items.<sup>7</sup> Thus, the 24 participants responded to 34 prime-target items in each of the three conditions, generating 816 observations per condition. However, in light of the fact that any mnemonic advantage for alliteration and assonance has tended to be quite modest in the experiments in Table 1, this number of observations per condition may have been insufficient to properly investigate a small effect. Indeed, for a repeated-measures reaction time study, Brysbaert and Stevens (2018) recommend a minimum of 1,600 observations per condition. This suggests that future experiments with sufficient power to detect a small effect would require approximately twice the number of items per condition or twice the number of participants. Recruiting a larger body of participants could perhaps be done more easily via a browser-based application like the *Gorilla Experiment Builder* (Anwyl-Irvine et al. 2020), though there are potential drawbacks to Internet-based reaction time data gathering (see Woods et al., 2015). To conclude, one possible explanation for the lack of a statistically significant finding is that although alliteration and assonance do facilitate lexical processing, the experiment was underpowered and thus was unable to capture the effect.

Another possible reason for not finding a priming effect for phonological patterns relates to the thorny issue of what counts as a “similar sound.” In this experiment, alliteration and assonance were operationalized as sounds that match at the phoneme level. This may not be a safe assumption. At this level, allophonic differences created by the phonological environment are not considered. Thus, for example, the /ɪ/ in *skift* and *milk* would be considered the same, even though the latter vowel would be darker before the velarized /l/. Future studies could investigate whether similar phonemes can also alliterate or assonate by virtue of sharing sub-phonemic features.

In sum, the data from the LDT study do not support strong claims that the mnemonic effect of alliteration and assonance is due to a form-based priming effect. This suggests that teachers should draw their students' attention to phonological patterns to aid L2 vocabulary acquisition.

## References

- Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, 52(1), 388–407. <https://doi.org/10.3758/s13428-019-01237-x>
- Baddeley, A. (2001). The concept of episodic memory. *Philosophical Transactions of the Royal Society B*, 356(1413), 1345–1350. <https://doi.org/10.1098/rstb.2001.0957>
- Boers, F., Eyckmans, J., & Lindstromberg, S. (2014). The effect of a discrimination task on L2 learners' recall of collocations and compounds. *International Journal of Applied*

<sup>7</sup> Jiang (2012) states “It is desirable to have ten to twenty items in a condition whenever it is possible” (p. 48). This experiment had 34 items in each condition.

- Linguistics*, 24(3), 357–369. <https://doi.org/10.1111/ijal.12033>
- Boers, F., Lindstromberg, S., & Eyckmans, J. (2012). Are alliterative word combinations comparatively easy to remember for adult learners? *RELC Journal*, 43(1), 127–135. <https://doi.org/10.1177/0033688212439997>
- Boers, F., Lindstromberg, S., & Eyckmans, J. (2014a). Is alliteration mnemonic without awareness-raising? *Language Awareness*, 23(4), 291–303. <https://doi.org/10.1080/09658416.2013.774008>
- Boers, F., Lindstromberg, S., & Eyckmans, J. (2014b). When does assonance make L2 lexical phrases memorable? *The European Journal of Applied Linguistics and TEFL*, 3(1), 93–107.
- Boers, F., Lindstromberg, S., & Eyckmans, J. (2014c). Some explanations for the slow acquisition of L2 collocations. *Vigo International Journal of Applied Linguistics*, 11(1), 41–62.
- Boers, F., Lindstromberg, S., & Webb, S. (2014). Further evidence of the comparative memorability of alliterative expressions in second language learning. *RELC Journal*, 45(1), 85–99. <https://doi.org/10.1177/0033688214522714>
- Brysbart, M. (2003). Bilingual visual word recognition: Evidence from masked phonological priming. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: The state-of-the-art* (pp. 323–343). Psychology Press.
- Brysbart, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, 1(1), 1–20. <https://doi.org/10.5334/joc.10>
- Brysbart, M., Van Dyck, G., & Van de Poel, M. (1999). Visual word recognition in bilinguals: Evidence from masked phonological priming. *Journal of Experimental Psychology: Human Perception and Performance*, 25(1), 137–148. <https://doi.org/10.1037/0096-1523.25.1.137>
- Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, 33(4), 497–505. <https://doi.org/10.1080/14640748108400805>
- Crystal, D. (1997). *The Cambridge encyclopedia of language* (2nd ed.). Cambridge University Press.
- Dumay, N., Benraïss, A., Barriol, B., Colin, C., Radeau, M., & Besson, M. (2001). Behavioral and electrophysiological study of phonological priming between bisyllabic spoken words. *Journal of Cognitive Neuroscience*, 13(1), 121–143. <https://doi.org/10.1162/089892901564117>
- Ellis, N. C., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly*, 42(3), 375–396. <https://doi.org/10.1002/j.1545-7249.2008.tb00137.x>
- Eyckmans, J., Boers, F., & Lindstromberg, S. (2016). The impact of imposing processing strategies on L2 learners' deliberate study of lexical phrases. *System*, 56, 127–139. <https://doi.org/10.1016/j.system.2015.12.001>
- Eyckmans, J., & Lindstromberg, S. (2017). The power of sound in L2 idiom learning. *Language Teaching Research*, 21(3), 341–361. <https://doi.org/10.1177/1362168816655831>
- Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments and Computers*, 35(1), 116–124.

- <https://doi.org/10.3758/BF03195503>
- Gibbs, P., & Van Orden, G. C. (1998). Pathway selection's utility for control of word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 24(4), 1162–1187. <https://doi.org/10.1037/0096-1523.24.4.1162>
- Goldinger, S. D., Luce, P. A., Pisoni, D. B., & Marcario, J. K. (1992). Form-based priming in spoken word recognition: The roles of competition and bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(6), 1211–1238. <https://doi.org/10.1037/0278-7393.18.6.1211>
- Jiang, N. (2012). *Conducting reaction time research in second language studies*. Routledge.
- Jusczyk, P. W., Goodman, M. B., & Baumann, A. (1999). Nine-month-olds' attention to sound similarities in syllables. *Journal of Memory and Language*, 40(1), 62–82. <https://doi.org/10.1006/jmla.1998.2605>
- Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, 42(3), 627–633. <https://doi.org/10.3758/BRM.42.3.627>
- Lindstromberg, S., & Boers, F. (2008a). The mnemonic effect of noticing alliteration in lexical chunks. *Applied Linguistics*, 29(2), 200–222. <https://doi.org/10.1093/applin/amn007>
- Lindstromberg, S., & Boers, F. (2008b). Phonemic repetition and the learning of lexical chunks: The power of assonance. *System*, 36(3), 423–436. <https://doi.org/10.1016/j.system.2008.01.002>
- Lindstromberg, S., & Eyckmans, J. (2014). How big is the positive effect of assonance on the recall of L2 collocations? *International Journal of Applied Linguistics*, 165(1), 19–45. <https://doi.org/10.1075/itl.165.1.02lin>
- Lindstromberg, S., & Eyckmans, J. (2017). The particular need for replication in the quantitative study of SLA: A case study of the mnemonic effect of assonance in collocations. *Journal of the European Second Language Association*, 1(1), 126–136. <https://doi.org/10.22599/jesla.26>
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, 88(5), 375–407. <https://doi.org/10.1037/0033-295X.88.5.375>
- McDonough, K., & Trofimovich, P. (2009). *Using priming methods in second language research*. Routledge.
- McQueen, J. M., & Sereno, J. (2005). Cleaving automatic processes from strategic biases in phonological priming. *Memory & Cognition*, 33(7), 1185–1209. <https://doi.org/10.3758/BF03193222>
- Monsell, S., & Hirsh, K. W. (1998). Competitor priming in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(6), 1495–1520. <https://doi.org/10.1037/0278-7393.24.6.1495>
- Nakayama, M., Sears, C. R., Hino, Y., & Lupker, S. J. (2012). Cross-script phonological priming for Japanese-English bilinguals: Evidence for integrated phonological representations. *Language and Cognitive Processes*, 27(10), 1563–1583. <https://doi.org/10.1080/01690965.2011.606669>
- Nation, I. S. P. (n.d.). The BNC/COCA word family lists. [https://www.wgtn.ac.nz/\\_\\_data/assets/pdf\\_file/0005/1857641/about-bnc-coca-vocabulary-list.pdf](https://www.wgtn.ac.nz/__data/assets/pdf_file/0005/1857641/about-bnc-coca-vocabulary-list.pdf)

- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Nation, I. S. P. (2014). *What do you need to know to learn a foreign language?* Victoria University of Wellington. [https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/paul-nations-publications/publications/documents/foreign-language\\_1125.pdf](https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/paul-nations-publications/publications/documents/foreign-language_1125.pdf)
- Norris, D., McQueen, J. M., & Cutler, A. (2002). Bias effects in facilitatory phonological priming. *Memory & Cognition*, 30(3), 399–411. <https://doi.org/10.3758/BF03194940>
- Perea, M., Vergara-Martinez, M., & Gomez, P. (2015). Resolving the locus of case alternation effects in visual word recognition: Evidence from masked priming. *Cognition*, 142, 39–43. <https://doi.org/10.1016/j.cognition.2015.05.007>
- Peters, A. M. (1983). *The units of language acquisition* (Vol. 1). CUP Archive.
- 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>
- Radeau, M., Morais, J., & Dewier, A. (1989). Phonological priming in spoken word recognition: Task effects. *Memory & Cognition*, 17(5), 525–535. <https://doi.org/10.3758/BF03197074>
- Rastle, K., & Brysbaert, M. (2006). Masked phonological priming effects in English: Are they real? Do they matter? *Cognitive Psychology*, 53(2), 97–145. <https://doi.org/10.1016/j.cogpsych.2006.01.002>
- Rubin, D. C. (1995). *Memory in oral traditions: The cognitive psychology of epic, ballads and counting-out rhymes*. Oxford University Press.
- Shallice, T., & Cooper, R. P. (2011). *The organisation of mind*. Oxford University Press.
- Slowiaczek, L. M., & Hamburger, M. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(6), 1239–1250. <https://doi.org/10.1037/0278-7393.18.6.1239>
- Slowiaczek, L. M., Nusbaum, H. C., & Pisoni, D. B. (1987). Phonological priming in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(1), 64–75. <https://doi.org/10.1037/0278-7393.13.1.64>
- Slowiaczek, L. M., & Pisoni, D. B. (1986). Effects of phonological similarity on priming in auditory lexical decision. *Memory & Cognition*, 14(3), 230–237. <https://doi.org/10.3758/BF03197698>
- Szudarski, P. (2017). Learning and teaching L2 collocations: Insights from research. *TESL Canada Journal*, 34(3), 205–216. <https://doi.org/10.18806/tesl.v34i3.1280>
- Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science*, 247(4940), 301–306. <https://doi.org/10.1126/science.2296719>
- Woods, A. T., Velasco, C., Levitan, C. A., Wan, X., & Spence, C. (2015). Conducting perception research over the internet: A tutorial review. *PeerJ*, 3, e1058. <https://doi.org/10.7717/peerj.1058>
- Yap, M. J., & Balota, D. A. (2015). Visual word recognition. In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 26–43). Oxford University Press.
- Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart’s *N*: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, 15(5), 971–979. <https://doi.org/10.3758/PBR.15.5.971>
- Zwitserlood, P. (1996). Form priming. *Language and Cognitive Processes*, 11(6), 589–596. <https://doi.org/10.1080/016909696386980>

【研究論文】

# 実例でみる「SV了〈我〉時間」構文

李 菲

## 要旨

動詞が表す行為の動作者である「我（私）」が目的語の位置に現れる「SV了〈我〉時間」構文は、従来の研究では「倒置使役文」とよばれ、「時間の損失」を表す無生物主語の使役文としてとらえられてきた。しかし、これまでの分析は作例に基づくものが多く、当該構文が実際の発話の場面でどのように用いられ、どんな表現機能をもつかについての指摘が少ない。本稿は実例を通して当該構文の用法について考察した。SNSや動画サイトなどで実例検索を行った結果、「SV了〈我〉時間」構文は専ら短文投稿サイト（Weibo）での「つぶやき」表現として用いられていることがわかった。この構文は、主語のない形（「V了〈我〉時間」）で現れることが多く、さらに動詞が他動詞か自動詞かによって大きく二種類に分かれる。他動詞の文は、「話し手が動詞の表す行為を遂行するのに、長時間かかった」ことを表し、「思わぬところで、手こずってしまった」という意味が含まれている。一方、自動詞の文は、「ある事物が原因となって、動詞が表す話し手の行為や状態が長時間続いた」ことを表す。他動詞文と自動詞文は「行為の持続」を含意する点で意味的なつながりをもつ。

キーワード：無生物主語，使役文，構文，誇張表現

## 1 「時間の損失」を表す倒置使役文

中国語には、次のような変わった構造をもつ構文が存在する。まず、逐語訳をヒントに次の文の意味を考えてみよう。

- (1) 这只破熊 画 了 我 一天 (微博)  
 こんなクマ 描く ASP<sup>1</sup> 私 一日

中国語は英語と同じSVOの言語として知られる。通常のSVO、「主語＋動詞＋目的語」の語順で(1)の意味を考えると、「こんなクマが、私の絵を一日描いていた」と訳せる。しかし、実際には(1)は「私はこんなクマの絵を描くのに、一日を費やしてしまった」という意味である。主語の位置にある「クマ」が描く主体ではなく、描かれる対象であり、逆に目的語の位置にある「私」がクマの絵を描く主体である。

1 ASPは、アスペクトマーカを指す。

これは、次のような通常の SVO、「S（動作主）+ V + O（動作対象）」の文と異なる。

- (2) 我 画 了 一天 熊。  
私 描く ASP 一日 クマ  
「私がクマの絵を一日描いた。」

(2) が通常の文ならば、(1) は動作対象と動作主の語順が逆さになった文である。このことから、(1) の構造をした文は従来、「倒置致使句」（倒置使役文）とよばれている<sup>2</sup>。「倒置」とは、「動作主と動作対象の位置が逆さになる、動作対象が文頭に前置される」ことを指す。英語や日本語にみられる「倒置文」「倒置法の表現」を思い起こさせるが、「倒置」の意味は異なる。

(1) は構造の面で通常の文と異なるだけでなく、意味の面でも特異性が見られる。(1) の文が伝えたいのは、「クマの絵一枚を描くのに、一日もかかってしまった」ことである。この文には、話し手の「思わぬことで手こずってしまった」や「一日を台無しにしてしまって、もったいない」といった苛立ちや悔しさが込められている<sup>3</sup>。しかし、こういったニュアンスは、(1) の文成分の字面の意味から読みとることができない。文を構成する要素に還元できない、非合成的な意味が(1)に含まれている。このような文に対しては、全体の意味を各要素に還元せずに、文全体を一つの「構文」として包括的にとらえる必要がある。

(1) の構造をもつ文は、倒置使役文に関する先行研究の中で、「時間の損失」を表す無生物主語の使役文としてとらえられている。「時間の損失」を表す倒置使役文の例としては、次の二つの作例が先行研究で取り上げられることが多い。

- (3) 那篇文章 改 了 我 一下午。(Gu, 1997)  
その文章 修正する ASP 私 午後  
「その文章を修正するのに、午後がまるまるつぶれた。」

- (4) 那个实验 做 了 他们 整整 一晚上。(Gu, 1997)  
その実験 する ASP 彼ら 丸 一晚  
「その実験をやり遂げるのに、彼らは丸一晚を費やした。」

Gu (1997) は、(3) (4) における最も大事なメッセージは「時間の損失」にあると指摘している。すなわち、両者は「ある行為を遂行するのに、長時間を要した」ということを表す文だとしている。Gu (1997) はこの「時間の損失」という意味に注目し、(3) (4) には「花／take」（かかる、費やす、費やさせる）という「covert verb」が隠れているとしている。その論拠として、(3) (4) の動詞を「花」（かかる、

2 (1) のような、「動作対象」と「動作主」の文中での位置が反転する文は、先行研究では様々な名称でよばれている。郭 (2006) では「倒置致使句」（倒置使役文）、周 (2006) では「客体归因句」（客体帰因文）、Gu (1997)・施 (2008) などでは「Inverted Causative Structure／反転使役结构」（反転使役構造）、さらに孫・郭 (2015) では「隐性述结式」（潜在動補構造）、石村 (2020) では「致事型数量动结式」（使役型数量動補構造）とよばれ、多くの研究者から関心をよせられている文法現象である。

3 話し手の苛立ちは「这只破熊」の「破」から読みとれる。「破」は指小辞 (diminutive) としての働きをもち、話し手の「クマ」への嫌悪感を表す。「这只破熊」は、日本語で言う「このクマめ」と対応している。

費やす、費やさせる) に変えても文が成り立つ点に言及している。

(3') 那篇文章 花 了 我 一下午。  
 その文章 take ASP 私 午後  
 「その文章を修正するのに、午後がまるまるつぶれた。」

(4') 那个实验 花 了 他们 整整 一晚上。  
 その実験 take ASP 彼ら 丸 一晩  
 「その実験をやり遂げるのに、彼らは丸一晩を費やした。」

(3') (4') は無生物主語の文である。動詞「花」は英語の take と同様、無生物を主語にとることができ、「(無生物) + 花 + (人) + (時間/お金)」の形で、「(無生物) のせいで、(人) がたくさんの時間やお金を費やした、失った」ことを表す。動詞「花」は「(人) に (お金や時間) を費やさせる」という使役の意味をもち、(3') (4') は無生物を主語とする使役文である。Gu (1997) は、(3) (4) を (3') (4') の使役文に変換し、両者の類似性を示すことで、(3) (4) のような倒置構造をもつ文も使役文の一種であるとした。使役構造をもつ (3) (4) において、主語の位置にある無生物の名詞句は「時間を損失させる原因 (cause)」であるという。

(3) (4) は無生物主語をもつ使役文の一種だとする見方の妥当性はさておき、こうした文は果たして単に「時間の損失」を表しているだけなのかについて、検討する必要がある。(3) (4) は (1) と同様、話し手の「文章の修正や実験が大変だった」、「費やした時間をもったいない」といった気持ちが込められている。そうした残念な気持ちがなく、単に「ある行為に使った時間」を言う場合、(1) (3) (4) のような倒置構造の文は成立しにくく、別の表現を用いなければならない。

(5) 我 十分钟 就 写完 那篇文章 了。  
 私 十分 だけ 書き終わる その文章 SF<sup>4</sup>  
 「私はたった十分でその文章を書き終えた」

(5')? 那篇文章写了我十分钟。

もし、「10分もしないうちに、クマの絵をすぐに完成させた」なら、やはり (1) の構造の文を使うことができない。

(1')? 这只熊画了我10分钟。  
 「10分でクマの絵を完成させた」

「時間の損失」や「倒置使役文」というだけでは、(1) (3) (4) のような文の特徴をとらえきれないことがわかる。以下では、短文投稿サイトである微博 (Weibo) で検出された実例を取り上げながら、こうした文の用法と意味について考察を行う。議論の便宜上、(1) (3) (4) の文を「SV 了〈我〉時間」構文とよぶ。

4 SF は、sentence final を指す。

## 2 つぶやき表現

従来の研究において、「SV了〈我〉時間」構文がどのような場面、文脈に出現しやすいのかについての議論があまりなされていない。これに対し、筆者は2022年12月から2月までの3ヶ月間、動画配信サイトやSNSを使い、中国語の話し言葉における「SV了〈我〉時間」構文の出現率について調査を行った。まず、動画やドラマ、映画を中心とするYouTubeとU-NEXTの場合<sup>5</sup>、一日2～3時間の視聴を3ヶ月間つづけても、確認できたのは次の1例のみだった。

- (6) 渴 了 我 十几个小时 (YouTube)  
 のどが渴く ASP 私 十何時間  
 「(水を飲ませてもらえず) 十何時間ものどが渴いたままだった」

3ヶ月間、毎日生の中国語を2時間以上聴きつづけても得られたのはこの1例のみである。つまり、「SV了〈我〉時間」構文は話し言葉にはほとんど現れない、と行ってよかろう。なお、(6)は主語がなく、かつ動詞「渴く」は自動詞であるので、先行研究における(3)(4)の文とタイプが異なる。この点についてはまた後述する。

一方、短文投稿サイトの微博では約93個の実例が見つかった。中国版Twitterとよばれる微博(Weibo)は、個人の日常生活での体験に基づく「つぶやき」が中心である。微博を中心に実例が多く見つかるということは、「SV了〈我〉時間」構文が目の前の相手に直接投げかけることばというより、話し手(短文投稿者)の心の声、感情を表出する一種の「つぶやき表現」となっているのではないかと考えられる。以下の例でみるように、感情を表出しているこうしたつぶやきは、「！」や絵文字が多く用いられている<sup>6</sup>。

- (7) 匡威夹心饼干太好看了, 等了我小半个月 (絵文字)  
 「コンバースの厚底スニーカーは本当に美しい。半月弱も待っていたんだ」<sup>7</sup>
- (8) 阿凡达2!!!! 我来了 (絵文字)  
 等了我多少年了 (五つの絵文字)  
 「アバター2!!!! 来たぞ 何年待たされたことか」<sup>8</sup>
- (9) 今天上班了, 咳了我一天!  
 「今日出勤したけど、1日ずっと咳がとまらなかった」

5 YouTubeでは社会・歴史がテーマの中国語のチャンネルを視聴し、U-NEXTでは中国語のドラマを視聴した。社会や歴史に関する話題を取り上げて解説するYouTubeでは書き言葉的表現が中心であるのに対し、ドラマでは口語的な表現が多い。

6 微博の実例に含まれている絵文字については、ワード上で同一のものを見つけられなかったため、本稿では単に「絵文字」で代用する。

7 「SV了〈我〉時間」構文が現れる前後文脈を提示するため、実例を原文のまま引用する。紙幅の関係上、以下では逐語訳を割愛する。

8 微博でのつぶやきは(8)のように、二行に分かれて表記されることが多い。これは投稿者の気まぐれというより、改行する場所が話題の提示と転換と大きく関わっている可能性がある。このため、本稿では改行の例についてはなるべくもとの形で提示する。中国語のつぶやきの文にみられる改行現象について、今後談話構造の観点からの分析が待たれる。



(10) 昨天看到了一张图，笑了我一节课 (絵文字二つ)

「昨日ある絵を見た。面白くて授業中ずっと笑いがとまらなかった」

(7) ~ (10) における「！」と絵文字の数から、投稿者の強い思いを読みとることができる。「(S) V 了〈我〉時間」の文は、個人の思い、実感を「つぶやき」という形で伝える際に用いられる表現であることがわかる。こうした構文は個人の思いを伝えるための「つぶやき」なので、目的語の位置にある「我」を、二人称の「你(あなた)」や視点や感情を共有していない第三者に置き換えることができない。

(7')\* 等了你小半个月

「あなたが半月弱も待たされた」

(10')\* 笑了一个人一节课

「ある人が授業中ずっと笑いがとまらなかった」

この点は郭(2006)でも指摘されている。郭(2006)によれば、倒置使役文の例において、「我(私)」や「他(彼)」などの人称代名詞が目的語になることが多く、人を指す一般名詞は倒置使役文の目的語になれないとし、次の例を挙げている。(11)は「異物付着」を表すタイプの倒置使役文である。郭(2006)はこれを「時間の損失」タイプの文と区別せずに、同じ倒置使役文として扱っている。

(11) a 那碗面吃了他一头汗。(郭, 2006)

「その面を食べていたら、彼が額いっぱい汗をかいた」<sup>9</sup>

b\* 那碗面吃了学生一头汗。

「その面を食べていたら、学生が額いっぱい汗をかいた」

bの文が成立しないという、この現象に対し、郭(2006)は談話構造の観点から分析し、「倒置使役文は目的語の位置にある使役対象を強調するための文ではないので、旧情報を表す人称代名詞が目的語に最も相応しい」と述べている。本稿ではこの現象に対し、「つぶやき」という観点からとらえてみたい。「SV 了〈我〉時間」の実例がほとんどSNS上での「つぶやき」表現であることを考えると、(7')(10')(11) bの不成立は、これらが個人の实感を表現した「つぶやき」ではないことと大きく関係していると思われる。つまり、目的語が「我(私)」だから実感になるのであって、「你(あなた)」や「ある人」「学生」だと実感ではなく、事実を客観的に伝える叙述文の口調になってしまう。一方、「我」以外に、三人称の「他(彼)」も当該構文の目的語になれるのは((4)、(11) aの例)、話し手の第三者への共感、感情移入と関係していると思われる。話し手が第三者と同じ目線で事態をとらえ、当事者としての「彼」「彼ら」の思いに共感し、感情移入したときに、(4)、(11) aのような文が生じるのだと思われる。

とはいうものの、(4)、(11) aは(7) ~ (10)のような実例ではなく、作例

9 以下、中国語の文献から引用した例文の日本語はすべて筆者によるものである。

である点に気をつけたい。ある表現が成立するということと、その文が実際の発話の場面でどのくらい用いられるかとは無関係である。そして、その表現の文法、意味の特徴を全般的にとらえるには、実際の発話の場面における実例をみる必要がある。倒置使役文の実例を調査すると、三人称の「他（彼）」が目的語となる用例にほとんど出会えない。これは「SV了〈我〉時間」の文が話し手の心情を表していることと関係している。(7)～(10)は、話し手の、自身の境遇や体験に対する一種の嘆きである。例えば、(8)は「待ちに待った映画がようやく見られることの嬉しさ」、(9)は「咳がとまらない状態で仕事に復帰したことの大変さ」、(10)は「笑いがとまらないほどの面白さ」をそれぞれ表現している。こうした感情は自分自身のものであり、ほかの誰かの感情を、「つぶやき」という形で表現することは、通常は考えにくい<sup>10</sup>。第三者のことを我が事同然のように感じなければ、(7)～(10)のような表現が生じにくい。三人称の「他（彼）」が目的語となる実例が少ないのはこのためであろう。

### 3 「V了〈我〉時間」の意味

次は、(6)～(10)でみた実例が、倒置使役文の先行研究における作例(3)(4)との違いについて考察したい。まず、実例と作例の最も大きな違いは、主語の有無である。(6)～(10)において、倒置使役文はすべて主語をもたない「V了〈我〉時間」の形となっている。

- (6) 渴了我十几个小时  
「十何時間ものどが渴いたままだった」
- (7) 等了我小半个月  
「半月弱も待っていたんだ」
- (8) 等了我多少年了  
「何年待たされたことか」
- (9) 咳了我一天  
「1日咳がとまらなかった」
- (10) 笑了我一节课  
「授業中ずっと笑いがとまらなかった」

それから、(6)～(10)の動詞のタイプが作例の(3)(4)と異なる。(3)(4)の動詞は作成の行為を表す他動詞で、動詞の動作対象(作成するモノ)が主語になっているのに対し、(6)～(10)では動詞「等(待つ)」以外はすべて動作対象をもたない自動詞である。例えば、(6)の「渴」、(9)の「咳」は「のどが渴く」「咳をする」という身体の状態を表している。(6)～(10)において主語が現れていないのは、動詞が自動詞であるため、そもそも動作対象をもたないことと関係していると思われる。

さらに、(8)のような、「了」が二つ入った「V了〈我〉(時間)了」という構造

10 三人称「他」で書かれている小説の場合、三人称「他」は実質一人称であるので、「他」は倒置使役文の目的語になることができる。

も(3)(4)と異なる。動詞直後のアスペクト(ASP)「了」と、文末助詞の「了」(SF)を二つ並べた「V了…了」の文は、英語の現在完了の「継続」用法に似た働きがあり、「過去のことが現在も継続している」ことを表す。(8)も「行為・状態の継続」の意味を含んでおり、「待ち続けて、もう何年になるか」を意味する。先行研究は(3)(4)の文を「時間の損失」の意味でとらえているが、「V了〈我〉(時間)了」は、「損失した時間」よりも動詞が表す行為・状態がなおも続いていることに焦点をあてている。

以上は作例だけでは知りえない、実際の例文にみられる構造上の主な特徴である。従来の研究ではふれられていないこうした特徴を、以下の3点にまとめることができる。

- ①主語のない、「V了〈我〉時間」の文
- ②動詞が自動詞である文
- ③時間の継続を表す「V了…了」の文

こうした構造上の特徴は、構文がもつ意味と直結している。ここで、(6)～(10)を、動詞が他動詞(等「待つ」)か自動詞かによって二つのタイプに分け、それぞれの構文がもつ意味について考えてみたい。まず、他動詞「等(待つ)」からなる(7)(8)は、待ち望んでいるモノを手に入れるまでの長い期間に焦点をあてており、「ものすごく長い時間がかかった」ことを表している点で、「時間の損失」ととらえられる。さらに、(7)(8)は動作対象(待ち望んでいるモノ)を主語の位置に加えることができる。次の(12)は形も意味も、作例の(3)(4)と同じタイプである。

(12) 匡威夹心饼等了我小半个月。(作例)

「コンバースの厚底スニーカーを手に入れるのに、半月弱も待っていた」

したがって、(7)(8)は主語が現れていないものの、構文が含意する意味に関しては、先行研究で論じられてきた作例(3)(4)と類似していることがわかる。他動詞からなる(7)(8)は、(3)(4)に最も近い実例といえる。

(7)(8)に対し、自動詞からなる(6)(9)(10)は動作対象をもたず、ゆえに主語を加えることが難しい<sup>11</sup>。そして、意味の面でも(7)(8)と大きく異なる。「時間の損失」を含意できる(7)(8)に対し、(6)(9)(10)が表しているのは自動詞が表す「行為・状態の継続」である。「十何時間も飲み物を与えてもらえず、のどが渴いたままだった」、「一日ずっと咳がとまらなかった」、あるいは「授業中ずっと笑わずにはいられなかった」ということから、(7)(8)という意味が読みとりにくい。こうした表現は、ある行為や状態が長時間つづいたことを表すことで、そうした事態をもたらした環境や原因の事物が「常軌を逸している」ことを伝えている。

例えば、私たちは「授業中、教科書の中から偶然すごく面白い絵を発見して、そのせいで授業中笑いがとまらなかった」としよう。そしてそのことをSNSなどで、

11 自動詞からなる文で、主語をもつ実例も検出されている。ただ、後述するように、自動詞文における主語は動詞の動作対象ではなく、動詞が表す行為、状態を招いた原因としての事物である。

(10) のようにつぶやいたとする。恐らく、そのつぶやきで伝えたいのは「絵の面白さ」であり、「ああ、笑いで一コマの授業の時間を失った」ではないはずである<sup>12</sup>。このように考えると、(10) の意味は実は二重構造になっていることに気づく。すなわち、一つは「授業中笑いがとまらなかった」であり、もう一つは「笑いがとまらないほど、絵が面白かった」である。

それからもう一つ大事な点がある。そもそも、(10) における「笑了我一节课（一コマの授業の間、笑っていた）」を文字通りに解釈してよいだろうか。授業時間にもよるが、50分授業なら、「50分間ずっと笑えばなしだった」ということになる。普通の人間ならまず無理であろう。したがって、(10) は一種の誇張表現 (hyperbole) としてとらえるべきである<sup>13</sup>。「一コマの授業」という長めの時間を設定することで、「笑いがとまらなかった」ことを誇張して表現する。さらに、「笑いがとまらない」ことを通して、偶然見つけた絵の面白さを伝えたり、あるいは授業に集中できないことを悔やんだりする。このように、(10) が含意しうる意味は多重構造となっており、それを単に「時間の損失を表す」という意味でとらえてしまうと、当該構文がもつ意味機能、表現機能を見えにくくさせているといえる。

#### 【(10) が含意しうる意味】

- ・ 授業中笑いがとまらなかった
- ・ 絵が面白かった
- ・ 授業を台無しにした（大事な授業で、話し手が真面目な学生の場合）

この意味のリストの中から、どれを汲み取るかは聞き手次第である。構文の意味を「時間の損失」と単一的にとらえてしまうと、当該構文が「誇張」表現としてもっている修辞効果、表現機能が見えてこない。

以上は(10)を例に、自動詞タイプの文の意味について考察したが、(6)(9)についても同様のことがいえる。(6)では「喉が渴いた時間」が大事なのではなく、「水を飲ませてもらえない過酷な環境だった」ことが文脈からもみてとれる。(9)では咳が続いた時間を誇張することで、「病気がまだ回復しておらず、咳に悩まされている」ことを表現している。

こうした修辭的表現としての特徴は、他動詞からなる(7)(8)からも見てとれる。

「待ち望んでいたモノと対面するまでかかった時間の長さ」という意味が読みとれると同時に、「ようやく対面できることの嬉しさ」「待ちきれない気持ち」、さらには「待ち望んでいるモノへの期待や愛着」なども含意されている。「！」と「絵文字」の数がこうした気持ちを直接表している。単に「失った時間の長さ」を述べるなら、「！」と「絵文字」を用いる必要がない。よって、(7)(8)も単に「時間の損失」を表すだけの文ではないといえる。

12 ただし、真面目な学生の場合、あるいは試験に関わるような大事な授業ならば、「時間の損失」の意味で解釈することが可能である。

13 李 2021 では、「異物の付着」を表す倒置使役文「V+了+我+一身汗」は一種の誇張表現 (hyperbole) であることについて指摘している。「V+了+我+一身汗」は文字通りに訳すと、「ある行為をした結果、全身汗だくになった」という意味であるが、実際は「全身汗だく」の状態にならなくても当該構文を用いることができる。「全身汗だく」という誇張した結果事態を述べることで、話し手が「思わぬことで手こずってしまい、苦労させられた」ことを表す。「V+了+我+一身汗」は、本稿の考察対象である「SV了〈我〉時間」と同じ表現効果、表現機能をもつことがみてとれる。

## 4 主語つきの文

ここまで、主語をもたない実例について紹介した。次は、主語をもつタイプの「SV了〈我〉時間」の実例を取り上げたい。主語つきの文は主語なしの文に比べ、用例の数が少ない。3ヶ月間にわたって微博上で実例検索をした結果、主語なしの文が66例、主語つきの文は27例検出された。つまり、(6)～(10)のような主語なしの文が全体の約7割を占めているのに対し、主語つきの文は全体の約3割にとどまっている。主語つきの文は数が少ないものの、形がバリエーションに富んでいる点が興味深い。一言「主語つきの文」といっても、それらの文は、主語の名詞句の形、動詞の意味、副詞の「就」の有無などで顕著な違いがみられる。以下では、大まかに三種類に分けて、それぞれの例を見ていく。

### 4.1 数量フレーズ

一つ目は、主語の名詞句が数量フレーズとなっているタイプである。数量フレーズとは、「数字+量詞<sup>14</sup>+名詞」の構造をもつ名詞句を指す。主語となる数量フレーズの中でも、「一+量詞+名詞」の形が多く見られる。

- (13) 一张小卡等了我一星期  
「たかがカード一枚で、一週間も待たされた（一週間もかかった）」
- (14) 一个外卖等了我三个半小时（絵文字）  
「出前一つで三時間半も待たされた」
- (15) 1000个字整了我一晚上（絵文字）下班下班！  
「一千字を書くのに、一晩かかった。仕事終了！」

(13)～(15)の主語は数量フレーズの形をとっているが、数量を表しているだけではない。「一つの～」という表現は、文末の長い時間を表すフレーズと対比をなしており、両者の落差が引き立つ。主語の部分で「行為が簡単に遂行できること」を期待させておいて、文末の時間フレーズで「長時間かかった」ことを言う。そしてこの落差から、「たかが～くらいで、長い時間をとられた」という意味が生じている。これらの文から確かに「時間の損失」を感じることができるとは、それに加え、「思わぬところで手こずってしまった、予想通りに物事が進まなかった」という意味も含意されている。

(13)～(15)では、主語の名詞句が動詞の動作対象でもある。これに対し、次の例では、主語が動詞の動作対象ではなく、動詞が表す行為を引き起こす原因となっている。動詞との関係性が以上の例と異なるが、主語の形が数量フレーズである点は同じである。

- (16) 一顿烧烤  
拉了我一早上

14 「量詞」は「助数詞」を指す。

(絵文字三つ)<sup>15</sup>

「焼肉のせいでお腹を下し、朝はずっとトイレから離れられなかった」

意味の面では、この文はむしろ主語なしの(10)の文と類似している。動詞「拉」は「排泄をする」という行為を表す。動詞の意味を考えると、(16)は(13)～(15)に含意されている「時間の損失」の意味が読みとりやすく、「長時間にわたる行為・状態の継続」として理解すべきである。(10)との違いは、そうした状態を引き起こした原因が直接文の主語として言語化されている点である<sup>16</sup>。

#### 4.2 副詞「就」

次は、主語の名詞句の後に副詞「就(～だけで)」が現れている例である。倒置使役文では「就(～だけで)」が現れやすい点については、郭姝慧 2006 がすでに指摘しているが、実例を挙げていない。微博上で見つかった次の実例は、主語が「特定」の事物を指す名詞句である。この点は、数量フレーズが主語となる4.1のグループと異なるが、副詞「就(～だけで)」を伴うことで、4.1のグループと同様の意味を表すことができる。すなわち、「～をただで、長い時間を取られた」という意味を次の(17)(18)から読みとることができる。数量フレーズがはたしている役割を、副詞「就」が代わりに担っているということができる。

(17) 这堆报销单就搞了我一早上(絵文字)

「これらの領収書进行处理するだけで、朝の時間が全部つぶれた」

(18) 广告前的科普部分大纲就做了我两个多星期(的下班后啦)

「コマーシャルの前に流す製品紹介の部分の概要を作るだけで、二週間以上残業させられた」

このタイプの動詞はすべて動作対象をもつ他動詞であり、主語が動詞の動作対象でもある。(16)のように原因が主語となり、さらに直後に「就」を伴う例は見つからない。(17)(18)は(13)～(15)と同様、「時間の損失」を含意すると同時に、「思わぬところで手こずってしまった」ことを表す。

#### 4.3 原因を表す主語

最後は、動詞が自動詞で、かつ主語が現れているタイプの文である。この場合、動詞が動作対象をもたないため、(13)～(15)や(17)(18)のような他動詞の文に比べると、主語と動詞との関係性がやや複雑になるが、概ね「原因」としてみることができる。

(19) 这图笑了我一宿哈哈哈哈哈

「この絵で一晩笑いがとまらなかった。アハハハハハ」

15 この例も上の行と下の行に分かれている。改行がもつ意味機能を考慮し、元の形で提示する。

16 倒置使役文に関する先行研究では、(16)のように「原因」が主語となる例も紹介されているが、ほとんどが作例である。

この例における主語の「この絵」は、「私を一晩笑わせた」原因である。これは(10)と全く同じ事態を表しており、「笑いが一晩続いた」ことによって「絵の面白さ」を伝える点も同じであり、一種の誇張表現としてみるができる。(10)との違いは、「絵」を主語に立てるかどうかだけである。(10)は主語にせずに、前の節で「絵」の存在を登場させているのに対し、(19)は原因の「絵」を主語に立てている。事態を引き起こす原因が主語となる点で、(16)と同じである。次の例の主語も事態の原因ととらえることができる。

(20) 22 个病人忙了我一晚上<sup>17</sup>

「患者が22人もいたので、一晩ずっと忙しかった」

この文の主語である「22人の患者」は、動詞「忙(忙しく働く)」の動作対象とはいえない。「?忙 22 个病人」は成立しない。よって、「22人の患者」は話し手が一晩忙しく働いた原因として主語となり、文全体が「患者が22人もいたので、忙しかった」ことを表す。この場合、動作対象が主語となる他動詞文に含まれている「たった～しただけで、予想以上の時間がかかった」という逆説的な意味がない。「22人の患者」と、「一晩働きづめだった」こととは自然な因果関係である。

(21) 肚子痛了我一晚上…无言…

「一晩中お腹が痛かった…無理…」

この文の主語「お腹」は文全体の中で二つの役割をもつ。一つは、動詞「痛む、痛い」の主語として、痛みがある部位を表す。もう一つは、「話し手を一晩苦しませた」原因としての役割である。(19)(20)では状態が長く続いた原因は外界にあるのに対し、(21)の場合その原因が外界ではなく、自分の内部にある。つまり、「お腹」に苦しみをもたらす原因として外から「客観的に」とらえている。(21)は単に「腹痛が一晩続いた」ことだけでなく、「お腹のせいで(腹痛のせいで)、一晩苦しかった」という「被害」の意味も含まれている。この「～のせいで、(ある状態が長く続いた)」という「被害」の意味は、自動詞文の全体にみられる意味特徴である。

## 5 まとめ

ここまで、「SV了<我>(時間)」構文の実例を色々見てきた。最後は、構造、動詞の意味、構文全体の意味特徴、という三つの観点から、これらの例をもう一度整理してみたい。まず、構造の面では、主語なしの文と主語つきの文の二種類が見られ、主語なしの文が全体の7割を占める。

表1 主語の有無による分類

S V 了<我>(時間)	主語なし	V 了<我>(時間)	Ex: 等了我小半个月
	主語つき	S V 了<我>(時間)	Ex: 一张小卡等了我一星期

17 この文は、主語が数量フレーズであるが、動詞が自動詞であり、かつ文全体の意味特徴が(13)～(15)と異なるため、4.1のグループと区別して考える必要がある。

主語つきの文は、主語の形、副詞「就」の有無、主語と動詞との関係によって、三種類に分類できる。

表2 三種類の主語つき文

S V 了〈我〉(時間)	Sが数量フレーズ	Ex: 一张小卡等了我一星期
	S 就 V 了〈我〉(時間)	Ex: 这堆报销单就搞了我一早上
	Sが原因	Ex: 这图笑了我一宿

従来の研究で、「時間の損失」を表す倒置使役文として紹介されている例は、この主語つきのタイプであると思われるが、(3)(4)のような例は以上の三種類のいずれにも属さない。すなわち、主語は指示詞を含む名詞句からなり、副詞「就」を伴わず、かつ動詞が他動詞の例は今回の調査では検出されていない。意味の面では、主語つきの文は主語なしの文にはない特徴がみられる。特に、動詞が他動詞で、主語がその動作対象である場合、「たった～しただけで(長時間がかかった)」というニュアンスが含意されている。数量フレーズという形や副詞「就」がこの意味の産出に貢献している。また、こうしたニュアンスを含意させないようにすると、文の容認性が低くなる。(13)の主語から数量詞を取り除いた次のような文は成立しにくい。

(13') ?小卡等了我一星期

主語つきの文は主語なしの文より用例の数が少ないことも考え合わせると、主語なしの文に比べ、主語つきの文は文法と意味の両面で用法上の制限が大きいと思われる。

動詞の意味からみると、「SV了〈我〉(時間)」の文は、自動詞からなる文と他動詞からなる文の二種類に分類される。自動詞文はさらに、主語をもたない文と、原因が主語となる文の二種類がみられる。両者の違いは、事態を招いた原因を主語にするかどうかにある。

表3 動詞の意味特徴による分類

	動詞	主語の有無	実例	意味
SV 了〈我〉(時間)	自動詞	V 了〈我〉(時間)	Ex: 笑了我一节课	ある事物が原因で、行為や状態が長時間続いた 訳:「～のせいで、～せずにはいられなかった・～しっぱなしだった」
		SV 了〈我〉(時間) *Sは原因	Ex: 这图笑了我一宿	
	他動詞	V 了〈我〉(時間)	Ex: 等了我小半个月	思わぬところで、ある事物に時間をとられ、手こずってしまった 訳:「たった～したくらいで、～の時間もかかってしまった」
		SV 了〈我〉(時間) *Sは動作対象	Ex: 一张小卡等了我一星期	



この表から、自動詞文と他動詞文は構造面だけでなく、意味の面でも違いが大きいことが見てとれる。自動詞ならば、構文全体が「行為・状態の持続」の意味となり、他動詞ならば、全体が「時間の損失」の意味に傾く。

ただ、他動詞文にも「行為の持続」という意味が含まれている。例えば、「等了我一星期(一週間も待った)」という発話には「一週間も待ち続けた」という意味と、「(カードが届くまで)一週間かかった」の両方の意味が含まれる。「行為の持続」と「時間の損失」はいわば原因と結果のような関係にあり、隣接している概念である。自動詞文と他動詞文は「行為の持続」という意味でつながっており、自動詞文は「行為・状態の持続」を、他動詞文は「時間の損失」をそれぞれ前景化したものと考えられる。

以上の考察結果からわかるように、従来「倒置使役文」などの名称でよばれている一連の文は実はいくつかの、異なる構造と意味をもった構文の集まりである<sup>18</sup>。これらは、形や意味の面で共通点がみられるものの、「倒置使役文」という名称で一括りにして論じるには無視できない違いも多い。当該構文の実態を把握するためには、単一の意味と形にとらわれない、多義性の観点が求められる。

### 参考文献

- Gu, Yang. (1997). On defining causativity and the significance of discourse information. *Text & Talk*, 17(4), 435–455. <https://doi.org/10.1515/text.1.1997.17.4.435>
- 郭姝慧 (2006) 「倒置致使句的类型及其制约条件」。『世界汉语教学』, (2): 40–50。
- Goldberg, Adele E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- 李菲 (2021) 「中国語倒置使役文と関連構文」。『東京大学言語学論集』, (43): 103–117。
- 石村広 (2020) 「致事型数量动结式的产生机制—一致动用法的发展和变异」。『当代语言学』, (22): 1–16。
- 施春宏 (2008) 『汉语动结式的句法语义研究』。北京: 北京语言大学出版社。
- 孙天琦, 郭锐 (2015) 「论汉语的“隐性述结式”」。『语言科学』, (5): 459–472。
- 周红 (2006) 「客体致使句的认知语义分析」。『语言研究』, 26(3): 20–25。

18 Goldberg (1995) では、構文というものの性質について、「Constructions are typically associated with a family of closely related senses rather than a single, fixed abstract sense.」(p.31) と指摘している。本稿の考察結果からも、構文は単一の意味と形をもった固定した、抽象的なものというより、いくつかの、意味が緊密に関連している構文群であることが見てとれる。

# **Enhancing Receptivity in EFL Classrooms: Exploring Teacher Development Opportunities Through Follow-up Reflective Diaries**

**Anna Belobrov**

---

## **Abstract**

In recent years, reflective practices in education have received plentiful attention as a powerful learner-development tool in the general and language education fields. This action research demonstrates ample potential benefits of self-reflective activities in the language classroom from the teacher development perspective through qualitative analysis of students' study journals in a speaking English as a second language tertiary classroom. This study indicates learners' awareness of their learning progress using journal writing as a medium of self-reflection on their language skill development and overall learning progress. It provides additional insights into the language learning process, such as learners' perception of the above, which includes goal-setting ability, learners' self-analysis of learning challenges, and self-motivation. The discussion and conclusion sections highlight the embedded potential of the reflective entries that fulfill a function as a background source for curriculum planning with objectives and activities adjusted to learners' needs that, overall, can serve as stimuli for a deeper and more effective learning process.

**Keywords:** *Learner development, Teacher development, Self-reflection, Learning diaries, Pedagogical implications*

## **Introduction**

It is an active research project designed to demonstrate the benefits of self-reflective activities in the language classroom from the teacher development perspective through a qualitative analysis of students' study journals in the English as a foreign language (EFL) tertiary environment. It offers additional channels to enhance receptivity in the language classroom through the perspectives of learner reflection.

## **Research Objectives**

The main objective is to explore the ways students' reflective diaries shed light on Japanese college students' learning process by analyzing the nature of the diary entries and their contribution to the improvement of pedagogical procedures. To deepen educators'

comprehension of the immense role of learners' involvement in the teaching process through self-reflective diaries (Little, 2007), which were investigated using a mixed method approach, the following research questions were asked:

1. What is the nature of the reflective journal entries?
2. How can the knowledge gained from diary entries be applied in the planning/teaching process?

The diary entries of college Japanese students majoring in English were analyzed based on the adoption of metacognitive features from Ellis (1999) and Sinclair (1999): analysis of the subject matter, learner awareness, and learning process awareness.

Theoretical sources indicate a strong connection between the implications of the self-reflective practice and enhanced receptivity to the learning contents in the EFL setting. The recognition of its significance for learners' empowerment, encouragement, and responsibility for learning is undeniable: it can naturally activate engagement with learning material, deepen comprehension, and reinforce critical and autonomous thinking (Belobrovy, 2018).

This study investigates the additional function of reflective practices, which is enhancing the receptivity of the material by sustaining teachers' choices of approach, curricular goals, and activities. This study was conducted in a higher education context in Japan; however, the outcomes can be generalized to other such contexts.

## Literature Review

### Definitions of Reflective Activities in the Literature

Reflective activities are defined as “those intellectual and affective activities that individuals engage in to explore their experience, which leads to new understanding and appreciation” (Boud et al., 1985, p. 19). According to Dewey (1991), experience-based reflective activities can potentially serve as a critical role in learner-centered education as a self-activating learning medium imposed at the beginning but appreciated at the end.

Studies by Lew and Schmidt (2011), attempted to prove the existence of connectivity between reflective activities and enhanced academic performance by adopting the view of the above as an inaccurate process that improves over time and includes engagement in reflecting on the content of past learning experiences by the learners and leads to potentially higher academic achievements in the future.

### Research on Journal Writing as a Self-Reflection Medium

Amulya (2004) emphasizes that written reflection can be a beneficial learning tool for gaining reflective experiences from the students' perspective and serving the educators. Park (2003) includes journals in the range of reflective activities that increase student interest in and engagement with course material, encourage and empower students to take more responsibility for their own learning, and allow them to have a voice and provide valuable feedback to the teacher. Such potential can provide irreplaceable information for pedagogical practices.

### **Metacognitive Learners' Benefits**

Multiple studies throughout the years suggested a correlation between the implementation of prescriptive reflective activities and enhanced receptivity of learning material. Reflective activities can also help learners increase their metacognitive awareness of their own learning and master successful learning behaviors as a life skill (Sevilla & Gamboa, 2016). Watson (2001) identifies reflective activities as a tool imminent for a performance-enhancing learning environment where transmitted knowledge reflected upon serves as mutually gained experience and thus accumulates potential for shortening the distance between the instructor and the learners. He suggests that reflection offers the learners additional opportunities to “negotiate strategies and alter content” (p. 140).

### **Implications for the Educators: Pedagogical Benefits**

Although learner-related benefits of reflective activities are fully captured in the literature with strong evidence of their efficacy, teachers' perspectives have been slightly less documented and implicit.

Reflection is seen as a disguised pedagogical tool that offers insights into language learning studies to target learner needs (Marefat, 2002). The learners are guided to examine their own behavior during class, i.e., introspection. If conducted correctly, such information can serve as an indicator of attitudes, needs, and the learning process, which is vital in the planning procedures. Cavilla (2017) notes the benefits of the infusion of reflective activities within the curriculum as they hold the potential to reveal the true meaning of the assignment at hand and its effectiveness in the acquisition process.

So far, research on reflective practices in EFL has focused on student and teacher performance (Klimova, 2015), with little attention paid to the potential value of learners' voices in broader pedagogical functions like curriculum planning. This study exemplifies student-generated pedagogical implications to maximize the impact of instruction.

## **Methodology**

This study utilizes the qualitative method of data collection and identifies categories of analysis revealing the nature of the reflective entries with regard to pedagogical issues of teaching a course at the university level. The factors explored in the study aimed at improving teacher–student interaction, enhancing participation, and elevating the effectiveness of the language learning process.

### **Setting**

The study was conducted in a private Japanese university in an undergraduate English language course targeting listening and speaking skills. Generally, the students enrolled in the course were second-year English majors. The primary objectives of the syllabus were to increase the independent ability of students to receive (listen), produce (verbal presentation), and interact in English. To achieve the learning goal and improve the level of production and quality of exposure, the students were engaged in a series of interactive pair and group activities during the course. A total of 11 student-participants enrolled in the course who had

scored an average of 500 on the Test of English for International Communication test and had intermediate levels of reading and writing skills. All the students acquired Japanese as their first language and had a similar history of English acquisition.

### Procedure: Diary Writing

As the literature suggests positive outcomes of self-reflective activities such as diary and journal writing on the learning process (Amulya, 2004; Park, 2003), I chose reflective journals as the main source of analysis. Moreover, as the educator was the one to manage the class and simultaneously engage in the research activities in my case, journal analysis was the most accessible and feasible data source. The students were instructed to reflect in a freestyle at the end of each session. It was deliberately emphasized that all the self-reflective activities would not be graded. The participants of the study were introduced to the purposes of the research in class and willingly signed a formal participant consent form upon the approval of the institution.

### Data Collection and Units of Analysis

The method of analysis chosen for this research was inspired by the study on promoting learner autonomy in the EFL classroom by Chuk (2003), which was centered on learner reflection. They include elements of metacognitive awareness described by Ellis (1999) and Sinclair (1999) as a starting point to define, eliminate, and measure learners' awareness of the learning process and identify those involving pedagogical implications. Journal entries were transcribed and analyzed. Each entry was categorized based on the dominant theme of the content. Influenced by Chuk (2003), the categories of analysis were coded as learners' awareness (LA), learning process awareness (LPA), and subject matter awareness (SMA) as shown in Table 1.

LA refers to entries containing pedagogical implications of students' self-awareness as learners in terms of attitudes, beliefs, motivation, needs, and learning styles. SMA implied the entries accommodating students' awareness of English as a system. LPA featured entries with students' awareness of learning process components such as self-assessment, goal-setting, monitoring progress, evaluating activities, and organizing time and resources (Chuk, 2003).

**Table 1**

*Categories of Analysis*

Awareness	Examples
Learner Awareness (LA)	Attitudes, beliefs, motivation, and learning styles
Subject Matter Awareness (SMA)	English as a Subject, recognizing metalanguage and referring to learning components and skills, speaking ability, listening comprehension, and others
Learning Process Awareness (LPA)	Self-assessment, goal-setting, monitoring progress, evaluating activities, and planning time for assignments

## Findings

This section offers the answer to the first research question regarding the nature of the diary entries and introduces the findings elicited from the analysis of diary entries. The outcomes in Table 2 demonstrate the number of weekly entries dominated by one of the categories—SMA, LA, or LPA—per student over the 15-week course. As this is a small-scale study, the numerical values showcase the general tendency in a specific context and build a notion of engagement activity patterns but cannot be generalized without large-scale further research.

**Table 2**

*Distribution of LA, SMA, and LPA Categories in Students' Diaries*

Weekly Entries Dominated by Category	SMA	LA	LPA	Total:
S1	2	1	12	15
S2	1	1	13	15
S3	1	1	13	15
S4	0	0	15	15
S5	4	1	10	15
S6	0	3	12	15
S7	4	2	9	15
S8	3	0	12	15
S9	0	6	9	15
S10	1	2	12	15
S11	2	0	13	15
Total per Semester:	18 (11%)	17 (10%)	130 (79%)	165

As shown in Table 2, LA comments were the least prevalent among the participants, accounting for 17% of the total. Additionally, three participants did not share any comments concerning the LA category. The total average of SMA comments accounted for a slightly higher average of 18%. Similarly, three participants made no comments regarding SMA. The majority of entries belonged to the LPA with an average of 79%. The numerical values provide an additional layer of comprehension, add visibility, and can lead to a discussion on the connections between student-led reflection and teachers' professional development. As the numerical values are not sufficient to illustrate the major self-reflective tendencies, specific examples are given below to create sufficient evidence for the value of enhancing teaching practices discussed later in the study.

### Learner Awareness

LA referred to entries containing pedagogical implications of students' self-awareness

of language skill development and learning difficulties recognized before and discovered through the course. Several entries were identified to include LA comments, with the two following dominating themes: acknowledging attitude and recognizing success (e.g., *I did my best* and *I tried to speak well*).

### **Subject Matter Awareness**

SMA referred to the entries containing pedagogical implications related to the students' awareness of English as a system. The following themes of SMA comments were identified in the data:

1. Awareness of content and language needs  
*I only use easy words, need to know more words*
2. Awareness of delivery skill development  
*Could know how to present*
3. Awareness of skill development necessity  
*I can't understand the listening*

### **Learning Process Awareness**

LPA referred to the entries with pedagogical implications from the perspective of students' awareness of learning process components such as self-assessment, goal-setting, monitoring progress, evaluating activities, and organizing time and resources. Three dominant components were detected:

1. Monitoring progress  
*Group work is fun, should continue improve stress of words.*
2. Self-assessment  
*I can't say my opinion, but I wanted to. Not good at speed reading and not good at grammar.  
Have to make sentences but I can't. We think about leadership- difficult as I have no leadership.*
3. Evaluating activities: feedback on the level of activity engagement.  
*Fun - can talk about a lot of things. Class fun - role-playing is interesting.  
Enjoyed-share different opinions with friends. Could discuss many things.*

The next section will focus on synthesizing the analysis of the findings and their relation to the main goal of this study: creating pedagogical implications for teachers' professional development.

## **Discussion and Pedagogical Implications**

The limited literature sources indicated that self-reflective practices can be approached as a powerful tool with educational implications by making general suggestions to incorporate the above into pedagogical routines (Cavilla, 2017). This study supports the outcomes and

expands general recommendations into the EFL learning field by offering concrete enhancement strategies based on active research findings.

The results described in the above section indicate that within the specific learning environment, the students demonstrated a high level of enthusiasm and openness to share their earnest feedback toward the content of the activities and foresee their possible learning benefits from the tasks. The pedagogical enhancement opportunities derived from the knowledge accumulated in the study are as follows.

First, as the findings show that the participants were fully aware of their strengths and weaknesses as learners and appeared to be in full control of the learning process in terms of goal-setting, self-motivation, and planning, the first enhancement suggestion for the educators is to allow a high degree of student autonomy by positioning themselves as facilitators. In this regard, the students might benefit from a guided introduction to self-assessment and peer reviewing.

Second, concerning SMA, it appears that the learners described themselves as already possessing a steadily functioning language processing system, with a certain awareness of its components as well as their own language development needs. As these appear pre-existing, educators can focus more on their enhancement and further development by introducing new collaborative project- and task-based activities.

The third recommendation fell under the category of LA in the diary entries. Given the small number of entries as well as the limited range of comments, it appears that this is the weakest awareness of the three among the learners. There is a clear indication that the learners have limited awareness of their personal role in the learning process. This further means that even though they are fully aware of its process and subject, they fail to recognize their language ability and express low confidence and motivation. This may also be the reason that majority of LA entries featured effort rather than success, underlying the necessity of raising students' self-esteem in the language acquisition process. Therefore, educators can consider contributing to the adjustment of learners' goals according to their feedback by adopting language confidence-building strategies in lesson planning by introducing fluency-oriented activities rather than accuracy.

Based on these discoveries, the suggested role of the educator here is to find ways to greatly enhance LA functions through a focus on open-ended and communicative lesson design. A low LA can be supported not just by shifting the focus of lessons toward fluency activities but also, as mentioned earlier, by allowing a higher level of autonomy, adopting a facilitating role, and enhancing mutual trust between the learner and educator. In combination with the high existing levels of SMA and LPA, proposed applications from the data and further continuous self-reflection and analysis of students' comments could bring about new perspectives, as well as increased language gains for the learners.

## **Conclusions**

This study was designed to contribute to the discussion on the extended benefits of learner reflection to language lesson planning, an area of high priority (Amulya, 2004; Park, 2003) but is under-researched. The results of this study contributed to a deeper understanding



of the nature of reflective diaries and identified the pedagogical implications associated with them. The diary entries revealed awareness in the areas of learners' motivation, language acquisition, and learning process management. The case is methodologically replicable and can lead to pragmatic outcomes that can offer additional insights into the mechanisms of educational enhancement of the language teaching process in other higher education institutions. Future studies can be extended to the influence of technology on the categories adopted in the study or the evaluation of curricular innovations.

## References

- Amulya, J. (2004). *What is reflective practice?* Center for Reflective Community Practice the Massachusetts Institute of Technology.
- Belobrovny, A. (2018). *Theories on self-reflection in education*. The Asian Conference on Education 2018 Official Conference Proceedings, Tokyo, Japan.  
[https://papers.iafor.org/wp-content/uploads/papers/ace2018/ACE2018\\_42511.pdf](https://papers.iafor.org/wp-content/uploads/papers/ace2018/ACE2018_42511.pdf)
- Boud, D., Keogh, R., & Walker, D. (1985). *Reflection: Turning experience into learning*. Kogan Page.
- Cavilla, D. (2017). The effects of student reflection on academic performance and motivation. *SAGE Open*, 7(3). <https://doi.org/10.1177/2158244017733790>
- Chuk, J. Y. P. (2003). Promoting learner autonomy in the EFL classroom: The exploratory practice way. *Proceedings of the Independent Learning Conference 2003*, Australia, 1–17.
- Dewey, J. (1991). *How we think*. Prometheus Books.
- Ellis, G. (1999). Developing children's metacognitive awareness. In C. Kennedy (Ed.), *Innovation and best practice* (pp. 108–120). Longman.
- Klimova, B. (2015). The role of feedback in EFL classes. *Procedia - Social and Behavioral Sciences*, 199, 172–177. <https://doi.org/10.1016/j.sbspro.2015.07.502>
- Lew, M. D. N., & Schmidt, H. G. (2011). Self-reflection and academic performance: Is there a relationship? *Advances in Health Sciences Education*, 16(4), 529–545.  
<https://doi.org/10.1007/s10459-011-9298-z>
- Little, D. (2007). Language learner autonomy: Some fundamental considerations revisited. *Innovation in Language Learning and Teaching*, 1(1), 14–29.  
<https://doi.org/10.2167/illt040.0>
- Marefat, F. (2002). The impact of diary analysis on teaching/learning writing. *RELC Journal*, 33(1), 101–121. <https://doi.org/10.1177/003368820203300106>
- Park, C. (2003). Engaging students in the learning process: The learning journal. *Journal of Geography in Higher Education*, 27(2), 183–199. <https://doi.org/10.1080/03098260305675>
- Sevilla, H. M., & Gamboa, R. M. (2016). Student self-evaluation and autonomy development in EFL learning. *Revistade de Lenguas Modernas*, 25, 199–222.  
<https://doi.org/10.15517/rlm.v0i25.27695>
- Sinclair, B. (1999). More than an act of faith? Evaluating learner autonomy. In C. Kennedy (Ed.), *Innovation and best practice in British ELT* (pp. 96–107). Addison Wesley Longman.

Watson, J. (2001). Social constructivism in the classroom. *Support for Learning*, 16(3), 140–147. <https://doi.org/10.1111/1467-9604.00206>

# **Incorporating the Elements of Fiction Into a CLIL-Based Extensive Reading Course**

**Michael Peragine**

**Christopher Mattson**

---

## **Abstract**

Within a course at a Japanese university based wholly on extensive reading, the authors of this paper implemented the pedagogical framework of Content and Language Integrated Learning (CLIL) to improve its effectiveness (Peragine & Mattson, 2023). A course outline was designed by incorporating literary devices or elements of fiction as the main content of the course, and students utilized this content within in-class activities to analyze and explicate their self-selected graded readers. The CLIL approach is briefly explained, along with the “content vs. language” balance educators must strike within such courses, and the nature of that balance within this course. An outline of literary concepts is given, along with a summary of each lesson’s activities. Feedback from students is summarized, in which they explained which literary elements, activities, and aspects of the course were perceived to be the most useful. The paper ends with conclusions drawn from students’ feedback regarding topics of interest and of the positive benefits for students of this CLIL-centered method, including a better appreciation of literature as well as improvements in writing about and discussing literature.

**Keywords:** *Pleasure reading, Extensive reading, Graded readers, EFL, CLIL*

## **Introduction**

Extensive reading often plays a supporting role in L2 reading-based courses. However, within the Independent Module at Rikkyo University, there exists a unique elective course, Pleasure Reading, based entirely on extensive reading as a means of helping students improve their reading fluency through the use of level-appropriate materials that they select for themselves. By empowering students to choose what they read, it is hoped that it will make for a more pleasurable experience, thereby fostering reading as a habit. To facilitate the objectives of the course, students engage in activities such as sustained silent reading, small-group book discussions, and written and/or oral assignments that both summarize and analyze students’ self-selected reading materials.

In a previous paper, the authors set out to apply the principles of Content and Language

Integrated Learning (CLIL) to the Pleasure Reading course, providing a detailed explanation of the pedagogical methodology for a lesson and handout regarding the topic of literary genres (Peragine & Mattson, 2023). Using this CLIL-based framework, the authors' created a curriculum aimed at enhancing the overall pleasure aspect of reading by introducing tools for literary analysis so that students could better interpret the books they were reading. This resulted in more academic discussions and an enhanced feeling of satisfaction by students at having not only completed a book in a second language but also having gained the ability to provide in-depth explanations as to what made each text noteworthy.

A brief overview and definition of CLIL are appropriate at this point, but laying out a singular, canonical definition is difficult. The history of CLIL can be traced back to 1990s Europe, where it was used to facilitate multilingualism, address the language–content imbalance in foreign language teaching, and offer students a more integrated approach to learning (Marsh, 2002). CLIL has since evolved and expanded with various definitions and models proposed. Llinares et al. (2012) state there are differences in approaches to CLIL depending upon myriad factors, including language of instruction, teacher objectives, starting age, teaching materials, language objectives, and research. Similarly, Mehisto et al. (2008) believe CLIL to be an umbrella term covering a dozen or more educational approaches. Despite the overlapping definitions, Marsh (2002) describes CLIL as “any dual-focused educational context in which an additional language is used as a medium in the teaching and learning of non-language content” (p. 15), and Coyle et al. (2010) further elaborate that it is “a dual-focused educational approach in which an additional language is used for the learning and teaching of both content and language” (p. 1).

A core tenet of CLIL coursework is the “four Cs – content, cognition, communication, and cultural awareness” (Coyle, 1999, p. 53). The pedagogical focus of this paper is on content, which serves as the foundation for the other three aspects. Coyle (1999) suggests that “it is through progression in the knowledge, skills and understanding of the content, by engagement in associated cognitive processing, interaction in the communicative context, and a deepening awareness of cultural self and otherness, that learning takes place” (p. 53). Although these four aspects are deeply intertwined, Coyle (2005) points out that content, or “thematic learning,” is still “at the heart of the learning process” (p. 5). Therefore, it is beneficial for CLIL educators to conduct a “curriculum subject audit” to identify the specific content-related skills, knowledge, and comprehension that need to be imparted (Coyle, 1999, p. 60). Equating content with progression serves as a reminder that strategic planning is necessary for both scaffolded learning within individual lessons as well as to enhance the logical progression of the topics to be presented throughout the course.

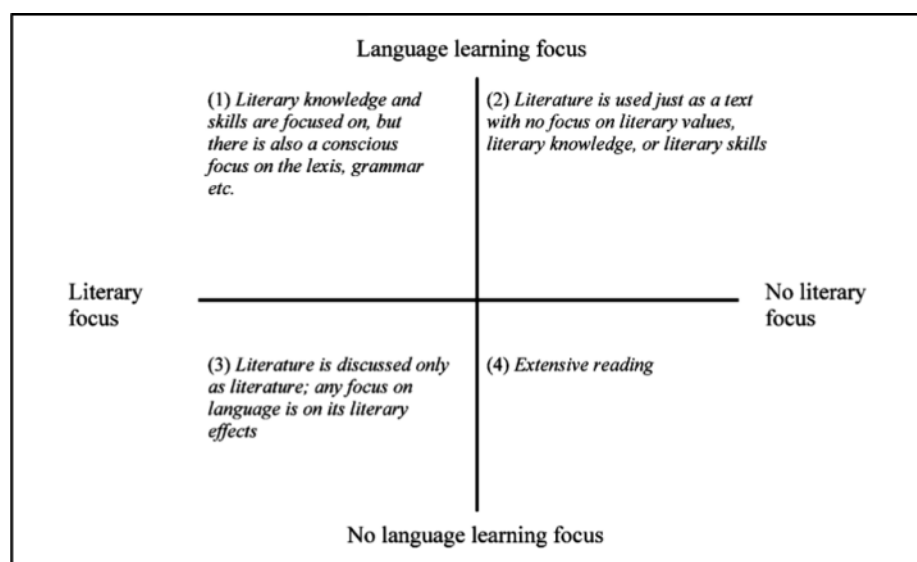
The primary content for this Pleasure Reading course was the elements of fiction and literary devices, and students were expected to use their understanding of these concepts to participate in discussion-based and written activities to help them break down their self-selected texts in a more meaningful way. As such, lesson themes progressed in terms of difficulty, starting with more conceptually approachable topics such as genres, setting, and characters before moving onward to more challenging topics such as plot, point of view, conflict, and symbolism. Student assessment was based on each student's understanding, application, and analysis of their graded readers through the lens of these literary devices.

As elements of fiction and literary devices constituted the main content of the course, they became the foundation for all discussions, presentations, and written assignments.

Another characteristic of this course was the focus on literary content rather than simply on language learning, the latter of which is more conventional at Japanese universities. According to Paran (2008), “There has been a move towards integrating language and literature” (p. 466). Paran provides a simple model (Figure 1) showing the range in which a course teaching language and/or literature might fall within a spectrum: the degree to which there is a strong literary focus (horizontal axis) and language learning focus (vertical axis). Based on Paran’s model, the implementation of CLIL with a focus on the teaching of literary analysis skills brings this course from being that of extensive reading (see Quadrant 4) to a course with both a clear literary and language learning focus (see Quadrant 1), thereby distinguishing it from standard L1 Literature courses (see Quadrant 3) that interpret texts through the lens of traditional literary criticism.

**Figure 1**

*The Intersection of Literature and Language Teaching*



*Note.* Adapted from *The Role of Literature in Instructed Foreign Language Learning and Teaching: An Evidence-Based Survey*, by Paran, 2008, p.466.

Within this CLIL-based framework, students gain the confidence to engage in in-depth discussions of various works of adapted literature with an understanding of the concepts and corresponding language necessary to participate in the collective act of scholarly interpretation using their L2. Therefore, it was the authors’ beliefs that incorporating such content would enhance the activity of extensive reading, taking it from simply reading in isolation with the purpose of improving fluency to a more fully engaging form of reasoning that stimulates cognition and critical thinking that will serve students well, both in and out of the classroom.

## A Literary Analysis-Based Outline of Course Content

The first lesson began with a certain amount of housekeeping, so to speak, mainly an introduction to the course content, an explanation of the benefits of extensive reading, a reading-level test from Extensive Reading Foundation (ERF; <https://erfpt.ealps.shinshu-u.ac.jp/>), information about accessing graded readers as well as recording student progress with a reading record, and various icebreakers that focus on building a comfortable class atmosphere while allowing students to discuss materials they enjoy reading in both English and their native language. The students were told that their book selection should be based on the difficulty level: no more than three unknown words per page was the main criterion for choosing, and by keeping the difficulty relatively low, the idea was to improve reading speed, comprehension, and enjoyment. For homework, students were required to find their first self-selected book. Although students were free to choose for themselves, they were strongly encouraged to read works of fiction as this would allow them to apply what they learn in later lessons to the books they have selected. If time permitted, the teachers accompanied students to the library to help them locate suitable reading materials.

### Genres

This lesson began with the teacher checking that all students came to class with a book to read. Students were introduced to the concept of previewing, which helped students get some idea about their books before actually having read them. To hone these skills, students viewed a movie poster for a then-unreleased Pixar film, making predictions about it using wh- questions. They then viewed the movie's teaser trailer to confirm their predictions and make new ones. Finally, they use these skills to preview their graded readers, asking questions such as "What is the title of your book? Who is the author? Is he/she famous? If so, what for? When was the story written? When/Where does the story take place? What do you think the story is about? Why did you choose the book? What do you think the characters are like?" After discussing these questions, it was expected that students would be sufficiently motivated to engage in sustained silent reading (SSR) and were given 15 minutes to read their graded readers in class. Then, the students read and completed a graphic organizer to learn about the key features of four literary genres: fiction, non-fiction, poetry, and drama (Peragine & Mattson, 2023). A series of scaffolded activities followed with the aim of further providing the language for students to both discuss their favorite genres while also assessing the genre of their graded readers. If time permitted, the lesson concluded with a second session of SSR. Typically, students were provided two SSR sessions per lesson with the goal of facilitating regular reading habits.

### Setting

The second literary concept taught was setting. The time and place of a story can be integral, so students were encouraged to ask questions like "What is the time period in which the story takes place? Where does the story take place? What is the social or cultural context of the setting? What are the customs, beliefs, and values of the characters in relation to the setting? How does the setting affect the characters and events in the story? Does the setting

impact the mood of the story?” Students were shown clips from different genres of film to show how the setting can vary greatly. Although this was a reading course, film clips were often used to visually exemplify many of the literary elements. Through graphic organizers, students applied these questions to their graded readers. In a culminating activity, students were shown various illustrations and—using their knowledge of setting—collaborated to create the setting of an invented story.

### **Character**

The concept of character is arguably the most important literary aspect of fiction. Students had to consider the difference between main and supporting characters; what it means to call a character flat, round, static, or dynamic; and the distinction between protagonists and antagonists. By applying these ideas to their graded readers, students had tools that allowed them to delve more deeply into their stories and discuss and write about the characters more fully. Students used graphic organizers in class to apply these concepts to characters that most were likely familiar with, describing attributes of characters such as those in *Doraemon* and the *Harry Potter* series. Once they understood the concepts, students were shown a visual cue, such as an illustration, and they collaborated with other students to create their own characters, applying the new concepts.

### **Plot**

By focusing on plot, students could better organize and understand the underlying structure of their graded reader. Plot is an essential element of fiction and helps readers comprehend the story’s structure and organization (Leki & Carson, 1997), and this can aid students in understanding how the story unfolds and progresses, as well as the relationships between the story’s events and characters. By using graphic organizers, students applied the stages of plot—such as exposition, climax, and denouement—to well-known stories as well as their own graded readers. Students also engaged in creative writing with visual prompts and created their own short stories.

### **Point of View**

The literary point of view involves the perspective from which a story is told, including the use of first, second, and third person pronouns; the use of unreliable narrators; and various types of narrators, such as objective, limited, and omniscient. Understanding point of view can also aid students in comprehending and analyzing literary texts by influencing how readers perceive characters, events, and themes in a story. Exploring different points of view helped students critically evaluate the biases and motivations of different characters and explore how the author’s own perspective influences the story. Students engaged with “Choose Your Own Adventure” graded readers from Atama-ii Books that used the second person, a rarely used point of view that was novel for many students.

### **Theme/Conflict**

Themes and conflicts in stories can be useful for enhancing students’ language and critical thinking skills. Themes are the underlying messages or ideas that are conveyed

through a story, such as love, good vs. evil, and revenge. By analyzing themes, students gained a deeper understanding of cultural values and beliefs. Conflicts are the obstacles or challenges that the characters face in achieving their goals, and students were introduced to six types of conflict (including person vs. person, person vs. nature, etc.) and encouraged to apply that framework to their own reading and discuss how each relates to their own lives and experiences. Through analyzing conflicts, students could learn to express empathy and understanding of diverse perspectives and situations.

### **Symbolism**

Symbolism allows the reader to interpret the meaning of an object or action in a story beyond its literal interpretation. The lesson began with scaffolding, starting with easily interpreted symbols such as local signs and flags before moving onto more abstract symbols such as numbers, colors, and flowers with an emphasis on varying cultural interpretations, thereby reinforcing the role of perspective on interpretation. Literary examples include F. Scott Fitzgerald's *The Great Gatsby*, where the large bespectacled eyes on an old billboard symbolize the eyes of God, watching over everyone. Other forms of symbolic representation were introduced, namely similes, metaphors, and allegories. By learning and applying symbolism with graphic organizers, students could better understand how authors convey deeper, more layered meanings and enhance the overall impact of the story.

### **Word Choice and Language**

A lesson that came the closest to explicit language teaching involved the teaching of an author's language and vocabulary choice, and how those choices inform and reflect the other aspects such as theme and character. The word choices that an author makes create the "voice" of the book, making it (in theory, at least) distinct from other works; that voice can extend to other aspects, such as character and plot. The students were given examples of alliteration and rhythm, of how aesthetically pleasing sentences can be made with clarity and rhythm. Students were then assigned to create their own poems and sentences with an online magnetic poetry website. With a set of simple words, students worked collaboratively while applying these concepts to create and share their own poems.

## **Methodology**

As the authors utilized a range of experimental methods to teach literary analysis skills, it was important to find out directly from students their perceptions of the lessons' contents and their influence on extensive reading. Being that this is an elective course, the participants included university sophomores, juniors, and seniors with ERF reading levels ranging from 3.7 (high beginner) to 13.8 (early advanced), which corresponds to Yomiyasusa levels of 0.7 to 5.8 and Common European Framework of Reference levels of A1 to C1. At the start of the semester, students were asked for their permission regarding the collection of verbal and written feedback through the observation of in-class discussions, interviews, and reaction papers. All students in two classes ( $N = 44$ ) agreed to participate. It is worth noting that this was an informal survey of students' reactions, meaning that students were not asked to



complete a formal survey with Likert scales. Instead, students were asked near the end of the course to write freely about their opinions on a range of topics including but not limited to the following: most useful/interesting lessons, memorable activities, the class’s effect on reading habits, future goals concerning extensive reading, and suggestions to improve the class. The authors then conducted a qualitative content analysis of these responses, focusing on the frequency of students commenting on a perceived reading level improvement, increased motivation to read in English, or endorsement of the benefits of applying literary analysis to their graded readers. Moreover, references to students’ interest in lesson themes and corresponding activities were also tallied in hopes of determining which topics were favorably viewed.

## Results

The results of the qualitative content analysis were compiled along with student quotes that represented each of these ideas.

**Table 1**

*General Feedback About the Course (N = 44)*

Student Comment	Teacher A	Teacher B	Total
Improved Reading Level	4	2	6
Motivation to Read in English Increased	12	14	26
Literary Analysis was Beneficial	10	9	19

Overall, students responded positively to the prompts (Table 1). Although students rarely commented on improvements to their reading level, they frequently reported an increase in motivation due to having the freedom to choose their own reading materials, which was described by one student who said, “Due to the freedom, I was very encouraged to read books of new genres I’ve never read before.” Moreover, the ease of finding graded readers at their level was another common reason for increased motivation as aptly explained by another student:

I think this class had a great impact on my way of thinking about reading in English. This is because each student was able to choose a book that suited his or her own level, so I did not fall behind because of difficulty. Therefore, I was able to read books with the same amount of stress as when reading Japanese books.

Although extensive reading was often cited as the impetus for increased motivation, several students also reported that activities teaching literary analysis had a positive impact, which is explained in the following comments:

Through this Pleasure Reading class, we were able to not only read books in English, but also think about them through various themes such as genre, setting, character,

plot, point of view, conflict, word choice, language, symbolism, chronology, and comparison with your life. By reading through these themes, I was able to consider things from a perspective I could not have before, and made new discoveries with an awareness of things I would not have noticed before.

Before taking this class, I had just been reading. However, after taking this class, I can read using a lot of knowledge. I am now able to think about the feelings of the characters, considering the place, the time period and the surroundings in which they are in, and I am able to enjoy the features of the authors' styles of writing.

As indicated by such comments, introducing content that helped students better analyze their books fostered a deeper appreciation of the stories they were reading, which in turn helped them support their opinions more coherently both in writing and in discussions. Furthermore, by tallying the frequency of student comments referring to certain lessons and activities as being interesting or memorable (Table 2), it becomes clearer as to the topics and skills that students believe to be most beneficial.

**Table 2**

*Lesson-Specific Feedback About the Course (N = 44)*

<b>The Most Interesting Lesson/Memorable Activity</b>	<b>Teacher A</b>	<b>Teacher B</b>	<b>Total</b>
Genres	3	1	4
Setting	2	3	5
Characters	4	2	6
Plot/Creative Writing	7	4	11
Point of View/Choose Your Own Adventure	10	10	20
Theme/Conflict	2	1	3
Symbolism	11	12	23
Word Choice and Language	3	3	6

Based on responses, 23 students thought symbolism was the most significant topic, and for 20 students, the most significant topic was point of view. Symbolism was often described as something that had gone unnoticed but useful for interpreting the subtle themes within their stories, for example:

Of all the lesson's themes, I found symbolism the most interesting. I had read English books without noticing symbols, so I did not notice the author's intention, which might be key signals for predicting the ending and understanding the situation well... By learning how symbols work in literature, I was able to predict and notice what the author meant, which made my reading activity more interesting.

Students also mentioned finding practical applications for analyzing symbolism in films,

TV dramas, commercials, Japanese novels, other English Literature courses, and even for their club activities, for example: “I belong to the English Drama Club. I am making costumes for casts there. The meaning of color is important to represent the personality of the role, so I thought I should learn more about symbolism.” Students also commented on the benefit of symbolism concerning cultural perspectives, for example:

For me, the class on symbolism was the most interesting. I learned that each person has their own feelings and thoughts of symbolism. For example, when we Japanese think of cherry blossoms, we associate them with the entrance ceremony and feel that they are a symbol of renewal. However, people from other countries tend to think of other symbols when they think of cherry blossoms, so I felt the change in my own stereotypes.

Like symbolism, students also regularly referenced having gained a newfound awareness of the narrative point of view and its effect on the stories, and again expressed benefits that extended beyond the classroom, for example:

The most interesting theme was Point of View because until this class I didn't care about something such as omniscience. I thought that if you care about point of view in books and life, it will be likely to improve both reading skills and life.

I read a novel quite often but I have never paid attention to the Narrator/Point of View. It was especially interesting to know the point of view is related to the plot because of the limitation of the amount of information. If a story is told in first person, readers also are able to know only what the narrator knows. In addition to that, I have never had a perspective of whether a narrator is reliable or not. So, I have never doubted the narrator of a novel and have understood what the narrator says directly. But, I started to think deeply of what the narrator says and see if he has some kind of prejudice or bias toward things.

I find it most interesting that the point of view of storytelling gives the readers different impressions about the story. For example, first person offers sympathy for the protagonist and at the same time it limits the readers' understanding of other characters. Therefore, people can be aware of the bias for reading a story by knowing which perspective of storytelling the author uses. This knowledge lets me enjoy reading books from different perspectives.

Although both symbolism and point of view were the most frequently mentioned literary skills, it is worth highlighting that all of the aforementioned topics received comments with a similar level of depth, emphasizing the benefits of empowering students with the content and language necessary to analyze their self-selected graded readers. This idea was expressed well by one student who eloquently extols the virtues of this style of extensive reading course:

I used to just read without thinking, but through this class, I have learned to pay attention to details such as introduction, climax, story closure, atmosphere, and characterization, which has broadened my perspective and made me enjoy reading many times more than before. I never knew how much fun it is to read English at my own pace, slowly, unlike a test.

By giving students the freedom to choose their own level-appropriate books while arming them with the concepts needed to interpret these works for themselves, students can come away with a newfound appreciation for reading in a second language.

### **Discussion**

Based on the feedback from students, the authors conclude that teaching literary concepts in a CLIL-based extensive reading course was an effective method of organizing its contents. Although the authors observed a steady improvement within in-class discussions in terms of the students' ability to analyze their graded readers using the concepts and language introduced in class, it was reassuring to hear from students directly that they also found these literary tools useful. By providing students with the necessary content knowledge skills to comprehend works of adapted fiction, they could not only gain a deeper appreciation of literature but also explain in their L2 the complex thoughts and feelings they may have while reading.

Through the focused teaching of content (literary devices), students were able to view literature through a different lens and broaden their understanding. To reiterate a few examples, the lessons on symbolism and point of view were frequently mentioned as effective in getting students to change their perception of literary works. Indeed, several student comments mentioned a kind of division in their thinking of literature; before the class, they had a narrower perspective of books and reading, but it had expanded by the end of the course. They wrote of "new discoveries" and an "awareness of things" not previously noticed, as well as new ways of thinking (cognition) about literary devices. In terms of the four Cs of CLIL, students remarked on the content of the lessons they liked as well as the cognitive aspects, but there were comparatively few comments regarding communication or culture. Based on the nature of the informal survey questions, it follows that students would focus on content. As reading is a solitary activity, it also follows that communication—namely in-class discussions—may not be as memorable. A common thread from students is that the most effective aspects of the course involved self-discovery: activities that introduced new ways of thinking about books and reading. Though it may be unclear whether these discoveries were achieved through discussion or independent self-reflection, it can be said that the content taught in the Pleasure Reading course achieved its purpose of providing a progression of knowledge that both engaged students and facilitated cognition when applied to students' self-selected graded readers.

## Conclusion

Our conclusions relied heavily on qualitative data with questions that focused on the aspects of the course that students found pleasurable. Although the data gives a general indication of students' perception of the course content, a more formal survey would be necessary to draw definitive conclusions about these lessons and activities. In future iterations of the course, it would be useful to implement a more detailed survey that allows students to respond to each of the lessons and its contents individually. For instance, although many students referred to memorable activities within the symbolism and point of view lessons, their comments do not indicate any aspects of these lessons that could be improved. Conversely, a lack of comments on topics such as genres, setting, and characters may not necessarily indicate that students found these lessons to lack significance nor imply abandoning these themes entirely, for these are the foundation upon which the latter topics are built. However, efforts could be made to incorporate more dynamic and memorable activities in hopes of fostering an appreciation of such topics. Furthermore, students rarely mentioned perceived improvements to their reading level, so it may be worth re-administering the reading-level test at the end of the course to determine whether such gains occurred. By giving students the opportunity to measure their own progress in terms of reading fluency and contribute regular topic-specific feedback to instructors, such data could be used to further enhance the overall development of the course.

## References

- Coyle, D. (1999). Supporting students in content and language integrated learning contexts: Planning for effective classrooms. In J. Masih (Ed.), *Learning through a foreign language: models, methods and outcomes* (pp. 46–62). Center for Information on Language Teaching and Research. <https://files.eric.ed.gov/fulltext/ED454735.pdf>
- Coyle, D. (2005). *CLIL: Planning tools for teachers*. The University of Nottingham, School of Education.
- Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and language integrated learning*. Cambridge University Press.
- Leki, I., & Carson, J. G. (1997). “Completely different worlds”: EAP and the writing experiences of ESL students in university courses. *TESOL Quarterly*, 31(1), 39–69. <https://doi.org/10.2307/3587974>
- Llinares, A., Morton, T., & Whittaker, R. (2012). *The roles of language in CLIL*. Cambridge University Press.
- Marsh, D. (2002). CLIL/EMILE—The European dimension: Actions, trends and foresight potential. *Public Services Contract DG EAC*. European Commission.
- Mehisto, P., Marsh, D., & Frigols, M. J. (2008). *Uncovering CLIL: Content and language integrated learning in bilingual and multilingual education*. Macmillan Education.
- Paran, A. (2008). The role of literature in instructed foreign language learning and teaching: An evidence-based survey. *Language Teaching*, 41(4), 465–496. <https://doi.org/10.1017/S026144480800520X>

Peragine, M., & Mattson, C. (2023). Pleasure reading: Incorporating CLIL into an extensive reading-based course. *Journal of Multilingual Pedagogy and Practice*, 3, 107–123.

# **English for Creative Purposes: Implementing a CLIL Framework for Tertiary Creative Writing Courses in Japan**

**Paul Curran**

---

## **Abstract**

Content and Language Integrated Learning (CLIL) has become an increasingly popular approach to English language teaching at the undergraduate university level in Japan as a means of raising the proficiency of local students and attracting international ones to study together in a global, English-speaking Liberal Arts environment. Creative Writing (CW) courses at the undergraduate and postgraduate university levels have become increasingly popular in English-speaking countries over the first quarter of the 21st century. If there is an assumption that CW courses are only for the most fluent of speakers, perhaps a CLIL approach could provide space for speakers who are not as fluent. Additionally, CW could encourage students to explore and share their diverse backgrounds. Therefore, this paper proposes that Liberal Arts programs using CLIL should be encouraged to incorporate CW into their curricula to reflect global standards. The paper outlines CLIL definitions, evaluates benefits and drawbacks, discusses a possible CLIL/CW course, and concludes that CLIL/CW could make a positive combination.

**Keywords:** *Content and Language Integrated Learning (CLIL), English Language Teaching (ELT), Creative Writing (CW), Course design, Classroom practice*

## **Introduction**

Creative Writing (CW) has been consistently called one of the most popular subjects among American undergraduate Liberal Arts students (Leahy, 2012). However, there are few CW courses in English language teaching (ELT), even those that have adopted the current trend for Content and Language Integrated Learning (CLIL). This paper reports on the process of implementing a CLIL framework for a CW course in the Center for Foreign Language Education and Research at Rikkyo University in Tokyo, Japan. The paper begins by untangling the various definitions of CLIL before offering a critical evaluation of the benefits and drawbacks of CLIL. This leads to a discussion of the process of designing a one-semester undergraduate CLIL course in CW. The paper concludes that CLIL and CW are a potentially positive combination for global tertiary education in a post-millennium world.

## Critical Evaluation: CLIL

Although Scrivener (2011) calls CLIL “one of the ugliest-ever ELT acronyms” (p. 327), he offers no explanation of why he finds it ugly. However, he does explain what CLIL is, and he is not the only one who feels compelled to do so. Much of the literature in CLIL’s relatively short history begins with a section or chapter addressing “What is CLIL?” (e.g., Ball et al., 2015; Coyle et al., 2010; Ohmori, 2014). Rather than staking out a new theoretical ground, the purpose of many CLIL definitions is to clear up confusion. This repeated need to clarify CLIL perhaps came about because there are so many different manifestations or variations of CLIL. Any stable definition is further complicated by featuring a range of benefits and drawbacks, and this instability has been amplified by more recent calls to move *beyond* CLIL (Coyle & Meyer, 2021). As such, in this section, a broad definition of CLIL is presented. This is followed by a critical evaluation of the benefits and drawbacks of CLIL.

### Definitions of CLIL

One of the most widely circulated definitions of CLIL calls it a “dual-focused educational approach in which an additional language is used for learning and teaching of both content *and* language” (Coyle et al., 2010, p. 1). However, there is neither a commonly accepted definition of what constitutes an appropriate integration of content and language learning nor is there an agreement on the exact role of subject and language teachers. This confusion has led to CLIL being labelled an elusive concept and an “umbrella term, covering a dozen or more educational approaches” (Ball et al., 2015, p. 24).

CLIL originated in Europe in the 1990s as an approach to teaching content (e.g., primary or secondary school curriculum subjects such as science or math) through L2, which in many cases, but not exclusively, meant English. CLIL soon spread into tertiary education and around the globe, where it was embraced by language teachers as a form of English for specific purposes or task-based learning (Ball et al., 2015). However, Ball et al. (2015) also criticize the way CLIL has been imported into language lessons, where they say it has been “taught by language teachers, assessed as language teaching, and makes no formal contribution to the subject curriculum” (p. 1). This crossover between subject and language teachers is central to issues revolving around the definition of CLIL.

Attempts to streamline CLIL’s definition and the issue of subject and language teachers have resulted in the distinction between “soft” and “hard” CLIL. Hard CLIL is close to full immersion. Subjects are taught by subject teachers in L2 for all contact hours for a year or more and possibly lead to public examinations in the subject. In hard CLIL, language teachers take on a consultative or supportive role where necessary. In contrast, soft CLIL is the form commonly seen in ELT settings such as Japan, where language teachers, who are not necessarily subject experts, import content into language lessons. The balance between language and content is based on nebulous factors including lesson stage, learner needs, and institutional requirements. These courses are typically less than a year, do not involve all contact hours, and have a lot of language teacher involvement (Ball et al., 2015).

Having briefly reviewed common CLIL definitions, the conclusion reached is that CLIL cannot be called a single approach or even a “uniform set of practices” (Ball et al., 2015, p.



2). Most recently, Coyle and Meyer (2021) has helpfully redefined CLIL as “a constantly evolving approach to learning and teaching in our multilingual and multicultural classrooms” (p. 3). Perhaps, the flexibility implicit in being a broad umbrella is one of CLIL’s greatest strengths. In other words, various versions can be readapted and redefined to suit local, cultural, and institutional conditions. However, as Scrivener (2011) warns “The hope is that achievement in both subject and language will be higher. The fear is that both will end up worse” (p. 327). Therefore, it is now appropriate to take a closer look at particular benefits and drawbacks of CLIL.

### **Benefits of CLIL**

Coyle et al. (2010) group the goals of CLIL under the heading of 4Cs: Content, Cognition, Communication, Culture. This 4Cs framework also offers a succinct way of examining and understanding the benefits of CLIL.

Firstly and perhaps most importantly, content is the subject matter that is being taught through CLIL. Content is most important because students are supposedly learning about an actual subject in a semi-authentic manner. This could be particularly beneficial for motivating more advanced students who already have a good grasp of communicative English. It could also benefit language teachers who have grown demotivated with communicative language teaching or other functional syllabus types (Breen, 1987). Furthermore, language teachers could focus on subjects that they have studied and are personally interested in, and institutions could promote various syllabus options.

Perhaps the most complicated but unique of the 4Cs is cognition. This relates to the development of thinking skills based on the content studied. Coyle et al. (2010) base cognition on Anderson and Krathwohl’s adaptation of Bloom’s taxonomy of thinking processes. This is a general approach to teaching that encourages students to be independent thinkers and develop processing skills such as analyzing, evaluating, and creating alongside the development of factual, conceptual, procedural, and metacognitive knowledge.

Although closely connected to content, the third C (communication) relates to language learning and implementation (Coyle et al., 2010). Communication is a reminder that CLIL lessons are not the same as subject lessons delivered to L1 students. For example, to encourage communication skills, CLIL teachers should modify content so that it is at a slightly higher level than students’ current ability. This gives students a scaffold in which to develop their language skills while they learn about the subject without the possible intimidation of total immersion. Teachers then provide any necessary feedback and support for students to learn the content. Teachers are also expected to encourage peer interaction and productive practice that helps develop communicative competencies such as breakdown and repair strategies (de Graaff et al., 2007). However, subjects taught as CLIL should not be an oversimplified or translation-assisted replication of previous disciplines studied in L1 (Coyle & Meyer, 2021).

Finally, culture relates to developing intercultural understanding and global citizenship, which is particularly important in the context of CLIL as an increasingly global phenomenon. As Ball et al. (2015) note, in many ways, the idea of simply learning English or any language, unless for a hobby, as a goal in itself is rapidly vanishing as the 20th century recedes. This

is particularly true of English in a global setting, where English is no longer simply one of many possible languages to learn but a “core skill” (Ball et al., 2015, p. 26). Although this “post-millennial, utilitarian view of the English language” (Ball et al., 2015, p. 26) as a passport to global citizenship may sound attractive, it begs questions about how difficult it is to implement CLIL.

### **Drawbacks of CLIL**

Though Coyle et al.’s (2010) version of CLIL has been praised as a framework or toolkit, criticism of CLIL revolves around the implementation process (Ball et al., 2015; Banegas, 2011). Furthermore, Scrivener’s (2011) concern about the balance of learning in CLIL relates to how teachers integrate content and language. Another related problem in CLIL is assessment.

If language and content are meant to be integrated as the CLIL acronym declares, this integration should also occur at the level of teacher roles. CLIL cannot be implemented with the assumption that teaching in L2 to non-native speakers is the same as teaching in L1 to native speakers. It is not CLIL if there is no or little language development. There must be some language support and some form of integration (Ball et al., 2015). However, when a CLIL course is taught solely by a teacher with an ELT background, if not a subject or content expert, then it is simply ELT. Having said that, perhaps the dichotomy between subject and language teachers is false. One solution might be to view all teaching as related to content and language. Ball et al. (2015) point toward the emergence of a hybrid teacher: “If all language teachers begin to consider themselves content teachers, then instead of maintaining the divide between the language and subject worlds, we begin to build bridges between them” (p. 272).

The duality implied by separating content and language as if content is somehow separate from the language of the content is also pertinent when it comes to assessment. Courses that foreground linguistic objectives could be soft CLIL, whereas courses that foreground procedural and conceptual content objectives could be hard CLIL (Ball et al., 2015). Therefore, the main difficulty in implementing CLIL assessment is deciding where the course sits on this spectrum. Ball et al. (2015) argue that “the priority should always be the content, but as Coyle et al. (2010) suggest, the assessment can also consider language features if the teacher deems it appropriate” (p. 215). Having reviewed the benefits and difficulties of CLIL, it is now appropriate to discuss how CLIL could work with CW in Japan.

### **Discussion: English for Creative Purposes**

English might be the language of globalization and internationalization; however, Japan is still relatively monocultural and monolingual despite several decades of ambivalent government attempts to alter the situation (Tada, 2016). Learning English in Japan has a long and sometimes controversial history, too complex to cover here; however, leading up to the postponed Tokyo 2020 Olympic and Paralympic Games, the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT) implemented guidelines for active learning

in Japanese schools (McMurray, 2018). MEXT also planned to increase English as a medium of instruction courses at Japanese universities. The ultimate goal is to raise the international standing of Japanese universities, transforming higher education in Japan into a global hub, where local students can study alongside international students. For example, “Hiroshima University is seeking to increase ‘foreign language taught courses’ to 50% across the university” (Rose & McKinley, 2018, p. 123), and although the university has not mentioned a date to reach this target, Waseda University is aiming for the same plateau for both undergraduate and graduate courses by 2032, which would be “up from 9% and 6% respectively in 2012” (Elwood, 2018, p. 37).

A result of MEXT’s plans is the proliferation of courses based on so-called soft CLIL (i.e., language teacher facilitated subject lessons) at Japanese universities (Ikeda et al., 2021). This combination of increased MEXT plans to globalize the Japanese education sector and a rich literary culture could thus lead to innovative approaches such as CLIL and CW. This section discusses the creation of an undergraduate CLIL/CW course. The section is divided into four parts that follow the 4Cs framework for a CLIL syllabus.

### **Content (Course Design and Content)**

Breen (1987) classically asks of any course, “what knowledge does it focus on?” He also asks what capabilities does it focus on and prioritize. The answer to the first question for a CLIL course would usually be that the knowledge is the content being imparted to the students, what they are expected to demonstrate knowledge of by the end of the course. For example, in a science or history course, students would learn about particular aspects of science or periods in history. However, in CW, as in other fine arts subjects, it could be argued that there is no specific content. This belief is perhaps related to the antiquated argument that CW cannot or should not be taught, particularly at the university level, because it is an exclusive art form, rather than a craft, that relies most on inborn talent (Kamata & Guenther, 2014).

However, teaching CW has a long history in academia, where it was once seen as an anomaly in literature departments but eventually gained its own space and is now seen as an essential part of the tertiary academic landscape (Cantrell, 2012). Furthermore, it is not only essential because of popularity or because students necessarily want to become published novelists, poets, or screenwriters. As Leahy (2012) says, although “most undergraduates will not go on to publish creative work, we should recognize that these language skills are transferable” (p. 75). This is an important point to consider from a soft CLIL/ELT perspective. A main difference with CW and other academic subjects is that the students bring much of the content. They may study different genres to learn how to appreciate or apply different styles and techniques, particular beats and themes, but what is most important is their own work. Furthermore, language learning is automatically integrated into the creation of that content. CW classes continuously integrate content and language into their curricula. In CW, it is assumed that everyone can improve their language use. There is always language development. Moreover, offering CW courses to L2 speakers could make a significant contribution to the field because new students could bring new perspectives to the use of global English.

The American Association of Writers and Writing Programs recommends that introductory undergraduate CW courses include “the reading and critical analysis of canonical and contemporary works of literature” as well as practice in “specific craft techniques” in the four most common genres: poetry, non-fiction, fiction, and drama/screenwriting and peer review with instructor moderation (Association of Writers and Writing Programs, n.d.). Therefore, CLIL/CW courses at Japanese universities could follow a similar format. As there are typically 14 weeks in a Japanese tertiary semester, the first lesson could be an introduction (and needs analysis), whereas the last lesson could be a round up or an additional workshop if required. The other 12 lessons could be divided into four modules: poetry, non-fiction, fiction, and screenwriting.

Each module would last three weeks (two seminars and a workshop). In the seminar lessons, students would describe, compare, examine, discuss, and/or rate different examples of particular genres, forms, styles, and techniques. Students would have homework reading from published literature that exemplifies that week’s particular craft skill or genre. This would be followed by a workshop lesson, where students would share their own work and give and receive critical feedback. For example, two seminars on poetry could be divided into canonical and contemporary works. Shakespearean sonnets could be contrasted with the Western haiku, and students could choose to write their own versions following or breaking each genre’s particular constraints. Similarly, alternative poetic forms, such as blank verse, narrative, and concrete poetry could be analyzed, evaluated, and composed.

By following a standard CW trajectory and eventually submitting a portfolio of selected work, the students would demonstrate that “tasks (and activities) in CLIL should not be seen in isolation, but as part of a larger sequence” (Ball et al., 2015, p. 207). Finally, students would also be expected to participate in workshops and communicate with each other about the process, and it is in the interlinked areas of content and communication that potential language issues would be addressed.

### **Communication (Needs Analysis)**

Institutional requirements would ideally cap the number of students on such a course at a maximum of 20, each with an IELTS score of at least 6.5 (B2/C1 of the Common European Framework of Reference [CEFR]), which is an average minimum level for entering an English-speaking university program or studying abroad, and it is the recommended minimum for doing CLIL (Ball et al., 2015). As such a course would be voluntary, it could also be expected that applicants were interested in English literature and CW. Additionally, enthusiasm is as important as experience. However, ideally, in a CW course, students would submit a brief written introduction and portfolio of previous work as part of the application process to serve as a needs analysis. Because student lists are perhaps only available shortly before the semester begins, the first lesson, as mentioned, could itself function as a form of needs analysis.

In regular lessons, although communication needs would be addressed as required, the sequencing of the four modules also points to particular areas of language use that may reveal the need for support strategies. For instance, the poetry module would focus on vocabulary, lexical range, and accuracy. Students would be encouraged to explore etymology and word

choice and areas such as figurative language. The second module, non-fiction, would focus on sentence-level grammatical structures and building paragraphs. This would also relate to the use of past and perfect tenses for reporting on memorable experiences. The fiction module would extend the focus on paragraph structures and include punctuation. This extension would include developing cohesion and coherence through the use of discourse markers and the conventions of formatting dialog. The focus on dialog in the fiction module would be further explored in the final module, screenwriting, which also requires particular formatting. Screenwriting would focus on the use of present tenses because the conventions of a screenplay require it to be read like the report of a film unfolding in real time.

Finally, an essential aspect of the workshop component would be student interaction. This aspect of the course would help students practice and develop communicative competencies such as breakdown and repair strategies (De Graaff et al., 2007). Furthermore, when sharing written work and engaging in peer feedback, Kamata and Guenther (2014) agree, in the context of L2 CW tertiary programs in Japan, students should feel relaxed and supported. Having a positive atmosphere would encourage students to experiment and take risks with their writing, some of which may be very personal. Therefore, students could be encouraged to remain neutral. Rather than identifying what they liked or did not like about a piece, from a traditionally critical perspective, a more equitable, rhizomatic approach could see students noting what resonated with them (Vanderslice, 2012). Students could also develop metacognitive skills to discuss and explain their own work as well as responding to examples of literature in different genres and to each other's work. This communicative process circles back to the second C: cognition.

### **Cognition (Materials Development)**

As mentioned previously, though a complicated area, cognition has been labeled the most important of the 4Cs in CLIL for Japanese students because it can encourage their development as critical and independent thinkers (MacGregor, 2016).

Creativity itself is an important aspect of cognition based on Anderson and Krathwohl's adaptation of Bloom's taxonomy of thinking processes (Coyle et al., 2010). Although Bloom's taxonomy is often mentioned in CLIL papers, little attention is given to creativity. However, CW courses, by the nature of their process and student-centered orientation, encourage creativity in the reading and writing process. They also encourage critical and independent thinking in the workshop process, where students are initially exposed to published materials before developing and sharing their own work.

Consequently, the most important materials developed in a CW course are the ones the students develop themselves. As mentioned above, materials that exemplify canonical and contemporary literature could be shared with students, but as the syllabus would be partly negotiated, these materials might change depending on initial needs analysis. However, these materials would not be held up as untouchable, superior to potential student output. Initial exposure to published work would not only give students different models to try on for size and discover their textual preferences but also allow students to practice responding to creative work before responding to each other's. This response would be "not in the traditional literary sense but with an eye towards interrogating technique" (Vanderslice, 2012,

p. 117). Students would not always be expected to judge the work but to comment on what resonates with them or which aspects of craft are noticeable and how those effects are achieved. Materials would also allow new CW students the chance to practice the metalanguage required to actively participate in the workshop setting without the risk of having to expose their own work first.

In producing their own materials, students would potentially engage with English in a more creative manner than they have ever experienced in an ELT setting, “putting together pieces to construct something new or recognizing components of a new structure, e.g. generating, planning, producing” (Coyle et al., 2010, p. 66). In addition to the ability to improve creativity, students would develop ways to analyze and evaluate their own work. They would explain their work, reflect on it, share and supportively critique each other’s. This is important because the “development of students as critics and reflective writers usually goes hand in hand with their development as creative writers” (Vanderslice, 2012, p. 116). As has been shown, CW has the potential to make a lasting impact on the cognitive development of students.

Finally, CW courses contribute to self-knowledge and self-expression because students think and write about themselves and their experiences. “Students can learn about the craft of writing, as well as how to think critically, skills that will prove useful to students even outside of English classes” (Kamata & Guenther, 2014, p. 527). It is outside of the classroom that students may eventually contribute to a wider society. With the students’ permission, their creative materials could be gathered into a journal for publication. Such a publication could serve to encourage students and add to the broader development of L2 CW culture.

### **Culture (Assessment and Evaluation)**

An aspirational objective of CW courses could include helping students with the “difficult task of writing works that might profoundly affect readers and enrich the culture” (O’Connor, 2012, p. 164). However, this goal is, with rare exceptions, virtually impossible to achieve during a full postgraduate degree, let alone a single undergraduate semester. It is also very difficult to gauge the impact on individual readers and the wider culture in years to come.

One recurring criticism of CW in general has been that it is impossible to grade or rank creativity. For as long as it has been part of academia, CW lecturers and teachers have grappled with the seemingly contradictory impulses of encouraging artistic freedom and submitting marks (Leahy, 2012). Cantrell (2012) points out that assessment in CW is often seen as an arbitrary judgment based on tastes and individual notions of objectivity. O’Connor (2012) agrees that assessing use of basic grammar, structure, and plot are relatively straightforward, but other aspects of literature such as vision or originality are too connected to taste.

Therefore, particularly at the undergraduate level, it is recommended that assessment should focus on elements of craft, attendance, participation, resilience, and effort and not necessarily on IELTS/TOEFL/CEFR equivalencies. The later process areas are considered more important than the final product. For example, in marking poetry, areas of craft to consider include demonstrating knowledge of different poetic forms used in the course and

the ability to “self-edit and build a poem through successive drafts” (Symmons Roberts, 2012, p. 152). Regarding fiction, in their previous Japan L2 CW trial, Kamata and Guenther (2014) also recommend focusing on craft aspects such as consistency of viewpoint, clarity of setting, three-dimensional characters, and purposeful dialog as well as some grammar and spelling.

Literature is the perfect place to share, explore, and exchange different cultural viewpoints. CW encourages students to share their experiences in written English as a form of cultural exchange. As English continues to be the global language, or *lingua franca*, it can also be a means through which people from different cultures can share the stories of their culture and heritage. Therefore, exposing students to literature, reading, and writing and discussing their perspectives is an objective that fits firmly within the CLIL framework and MEXT plans. If students are inspired to continue writing creatively in English and choose to pursue publishing or further studies in the future, that would be magnificent. However, on a practical level, those considerations fall outside the assessment remit. Having said that, institutional requirements stipulate that students receive a grade, and the university culture expects grades.

The assessment for elective lecture courses, regardless of content or subject, might be broken down into participation (30%), coursework (40%), and final paper (30%). For a CW course, participation could be subdivided into attendance and active participation during individual feedback sessions, whereas in lectures and workshops, the coursework could be divided into four short pieces (10% each) submitted at the end of each module: poetry, non-fiction, fiction, and screenwriting. The final paper (30%) would include a portfolio of an extended piece or combination of pieces alongside a short reflection that discusses considerations of the different techniques explored and responses to feedback given. The reflection would also include a self-assessment component that could be used for grading in consultation and negotiation with the teacher. For example, students could justify their own grades based on how far they believe they have achieved what they set out to accomplish. Additionally, a course evaluation from the students could be used for research purposes and to improve the course.

In offering this course to Japanese and other international L1 or L2 undergraduate students, it is hoped that a contribution can be made to the field of CLIL and CW.

## **Conclusion**

This paper reported on initial research into implementing a CLIL framework for undergraduate CW courses at Japanese universities. Future research could compare L1 and L2 CW courses and survey student opinions. The current paper explored CLIL definitions and offered a critical evaluation of the benefits and difficulties of CLIL. There was then a detailed discussion of the process of designing an undergraduate CLIL course in CW. In this post-millennium world, universities implementing plans to globalize Japan’s tertiary culture could surely benefit from offering students courses that encourage English for creative purposes.

## References

- Association of Writers and Writing Programs. (n.d.). *Guide to writing programs*. Association of Writers and Writing Programs.  
[https://www.awpwriter.org/guide/guide\\_writing\\_programs](https://www.awpwriter.org/guide/guide_writing_programs)
- Ball, P., Kelly, K. & Clegg, J., (2015). *Putting CLIL into practice*. Oxford University Press.
- Banegas, D. L. (2011). A review of “CLIL: Content and language integrated learning.” *Language and Education*, 25(2), 182–185.  
<https://doi.org/10.1080/09500782.2010.539045>
- Breen, M. P. (1987). Contemporary paradigms in syllabus design part II. *Language Teaching*, 20(3), 157–174. <https://doi.org/10.1017/S026144480000450X>
- Cantrell, M. (2012). Assessment as empowerment: Grading entry-level creative writing students. In H. Beck (Ed.), *Teaching creative writing* (pp. 156–159). Palgrave Macmillan.  
[https://doi.org/10.1057/9781137284464\\_22](https://doi.org/10.1057/9781137284464_22)
- Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and language integrated learning*. Cambridge University Press.
- Coyle, D., & Meyer, O. (2021). *Beyond CLIL: Pluriliteracies teaching for deeper learning*. Cambridge University Press.
- De Graaff, R., Koopman, G. J., Anikina, Y., & Westhoff, G. (2007). An observation tool for effective L2 pedagogy in content and language integrated learning (CLIL). *International Journal of Bilingual Education and Bilingualism*, 10(5), 603–624.  
<https://doi.org/10.2167/beb462.0>
- Elwood, K. (2018). CLIL in Italian universities: Lessons for Japan. *The Cultural review*, 51–52, 37–59.
- Ikeda, M., Izumi, S., Watanabe, Y., Pinner, R., & Davis, M. (2021). *Soft CLIL and English language teaching: Understanding Japanese policy, practice and implications*. Routledge.
- Kamata, S., & Guenther, D. (2014). Fundamentals of creative writing for Japanese university students. In N. Sonda & A. Krause (Eds.), *JALT2013 Conference Proceedings, Tokyo: JALT*. (pp. 521–528).
- Leahy, A. (2012). Undergraduate creative writing in the United States: Buying in isn’t selling out. In H. Beck (Ed.), *Teaching creative writing* (pp. 73–79). Palgrave Macmillan.  
[https://doi.org/10.1057/9781137284464\\_9](https://doi.org/10.1057/9781137284464_9)
- MacGregor, L. (2016). CLIL focus: An interview with Professor Makoto Ikeda. *The Language Teacher*, 40(1), 18–21. <https://doi.org/10.37546/JALTTLT40.1-4>
- McMurray, D. (2018). MEXT’s new course of study guidelines to rely on active learning. *The Language Teacher*, 42(3), 27–29.
- O’Connor, S. (2012). Ranking student writing as bad pedagogy and a bogus pretense of objectivity. In H. Beck (Ed.), *Teaching creative writing* (pp. 160–165). Palgrave Macmillan. [https://doi.org/10.1057/9781137284464\\_23](https://doi.org/10.1057/9781137284464_23)
- Ohmori, A. (2014). Exploring the potential of CLIL in English language teaching in Japanese universities: An innovation for the development of effective teaching and global awareness. *The Journal of Rikkyo University Language Center*, 32, 39–51.
- Rose, H., & McKinley, J. (2018). Japan’s English-medium instruction initiatives and the



globalization of higher education. *Higher Education*, 75(1), 111–129.

<https://doi.org/10.1007/s10734-017-0125-1>

Scrivener, J. (2011). *Learning teaching* (3rd ed.). Macmillan Education.

Symmons Roberts, M. (2012). Assessment of poetry in higher education courses: What are the limits? In H. Beck (Ed.), *Teaching creative writing* (pp. 151–155). Palgrave Macmillan. [https://doi.org/10.1057/9781137284464\\_21](https://doi.org/10.1057/9781137284464_21)

Tada, M. (2016). Recent reform to the English education system in Japan. *21st Century Education Forum*, 11, 21–29.

Vanderslice, S. (2012). The lynchpin in the workshop: Student critique and reflection. In H. Beck (Ed.), *Teaching creative writing* (pp. 116–120). Palgrave Macmillan.

[https://doi.org/10.1057/9781137284464\\_16](https://doi.org/10.1057/9781137284464_16)

# 外国語教育研究センター 外国語教育研究ジャーナル投稿規程

施行 2020年4月14日

改正 2022年9月27日

2023年3月13日

## (投稿資格)

- 第1条 執筆者は（共著の場合はファーストオーサーのみ）、原則として現職の外国語教育研究センター（以下、センター）所属教員とする。特別号、特集号などの場合はこの限りではない。
- 2 投稿論文はそのカテゴリーを問わず1教員につき1本までとする（共著による論文も1本と数える）。ただし、投稿論文数によっては、ジャーナル&リサーチ委員会（以下、委員会）による決定を経て掲載号・論文枚数等の調整を受ける場合がある。
  - 3 投稿論文は過去に出版されておらず、他のジャーナルに現在投稿されているものではないこと。

## (使用言語)

- 第2条 原稿執筆にあたっては、センターに所属する教員が広く互いの研究及び知識の共有を活性化するにあたり、原則として日本語あるいは英語での執筆を基本とする。ドイツ語、フランス語、スペイン語、中国語、朝鮮語及びロシア語による原稿執筆も可。

## (執筆要項)

- 第3条 原稿の種類は、次のいずれかの区分に属するものとする。区分は執筆者が判断するが、委員会が確認し、必要に応じて、変更を促すこともある。

### (1) 研究論文

当該分野で一定の結論を出し、その分野に貢献すべき内容が分析、議論されているもの。従って、執筆者は少なくとも、セクションとして導入と結論を明記し、先行研究を踏まえた当該論文の位置づけ、分析や議論（またはそれに代わるもの）にあたる記述を含めること。

### (2) 研究報告

当該分野における研究の進捗状況を公表し、将来その分野において貢献すべき可能性のある研究経過をまとめたもの。

### (3) 書評

書籍や文献を紹介しまとめたもの。

- 2 書式は以下の項目全てを厳守すること。英語で執筆する場合は、英語版の投稿規程及び別紙テンプレートに従って作成すること。英語以外の言語で執筆する場合は、別紙テンプレートを参考にして作成すること。なお、原稿の書式が以下の項目に則していない場合は、執筆者に原稿を返却し、修

正を求める場合がある。

(1) 原稿サイズ

A4版を使用し、上下各25mm、左右各25mm ずつあける。

横書きでフォントは日本語 (MS明朝)、英語は (Times New Roman) 12ポイントを使用。その他の言語のフォントは同一論文内で齟齬のないよう当該言語における学術論文で一般的に使用されるフォントを選択し使用すること。行間は1行とし、本文は左揃えとすること。

(2) 字数

字数 (語数) は以下のとおり言語毎に定める。全言語において、図表、参考資料、参考文献、注、付録など全てを字数 (語数) に含める。なお、図表については、明瞭なものを当該個所へ貼りつける。貼り付けられない場合は、別ファイルを用意し、挿入個所を明示する。

	研究論文	研究報告	書評
日本語	20000 字程度	6000 字程度	2500 字程度
英語	8000 語程度	3000 語程度	1000 語程度
ドイツ語	60000 字程度 (30ページ程度)	18000 字程度 (9~10ページ程度)	7500 字程度 (3~5 ページ程度)
フランス語	8000 語程度	3000 語程度	1000 語程度
スペイン語	8000 語程度	3000 語程度	1000 語程度
中国語	13000 字程度	4000 字程度	1600 字程度
朝鮮語	20000 字程度	6000 字程度	2500 字程度
ロシア語	8000 語程度	3000 語程度	1000 語程度

(3) 原稿タイトル

左寄せ、18ポイント、太字とする。フォントは上記書式に従う。執筆言語が英語の場合はタイトルケースとすること。

(4) 氏名

右寄せ、ゴシック体で12ポイントとする。タイトルとの間は1行あける。

(5) 要旨

全ての原稿に、要旨と3~5項目のキーワードをつけること。要旨は日本語500字程度または英語150~250語程度で執筆すること。書式は、左右15mm ずつ全行インデントし、フォントはMS明朝、11ポイントを使用する。なお、要旨の言語は執筆言語に関わらず英語あるいは日本語とする。

(6) 註釈

本文への註釈は、対応する註記を各ページの下に9ポイントで表記する。

**(原稿の提出)**

第4条 原稿の提出は、指定のGoogleフォームに必要事項を入力の上、投稿内容を収めた電子媒体をアップロードし、送信すること。

**(脚注及び参考文献の形式)**

第5条 原則としてAPAスタイル(第7版)を用いること。英語以外の言語で執筆する場合は、APA(第7版)の形式に可能な限り沿うよう配慮し、執筆者の責任において同一論文の中で齟齬のないよう確認すること。剽窃を行わないよう十分に注意すること。

**(投稿開始及び締切日)**

第6条 投稿の受付は毎年4月春学期の開始から行う。初稿の提出締め切り日は6月初週最終日とする。

**(査読に関して)**

第7条 研究論文カテゴリへの投稿は、匿名の査読員1名による査読が行われる。査読は規定のループブックを用いて行い、査読員は任意のコメント、修正案等を追記することができる。査読終了後委員会より、春学期終了前までに執筆者へ査読結果を報告し校正等の打診を行う。研究報告及び書評への投稿は原則として査読は行わない。ただし基本的な内容の確認を委員会が行い、掲載可否を判断する。

**(校正及び再提出)**

第8条 査読を受けた執筆者は、論文の校正を行い、秋学期第1週日末日までに再提出を行う。校正後の原稿は委員会による最終確認を経て掲載可否の判断を行うものとし、執筆者に結果を通知する。なお出版社より体裁等の修正依頼があった場合は、再度の校正を執筆者に依頼する場合がある。

**(出版)**

第9条 当該ジャーナルは毎年12月に出版される。

**(CiNii及び立教リポジトリへの登録)**

第10条 掲載された論文は、外国語教育研究センターのウェブサイトに掲載され、CiNii(国立情報学研究所論文情報ナビゲーター[サイニィ])及び立教大学学術リポジトリに登録される。

**(その他の要件)**

第11条 その他の要件は以下に定める。

- (1) 原稿料は支払わない。
- (2) 掲載された論文の著作権は、原則として立教大学外国語教育研究センターに帰属する。ただし、著者が著者自身の研究・教育活動に使用する際は、許可なく使用することができるものとする。
- (3) 万が一出版後、剽窃等の不正が発覚した場合は当該論文をジャーナルから削除する。

**(規程の改廃)**

第12条 この規程の改廃は、外国語教育研究センター教授会の議を経て、外国語教

育研究センター長が行う。

附 則

この規程は、2020年4月14日から施行する。

附 則

この規程は、2022年9月27日から適用する。

附 則

この規程は、2023年3月13日から適用する。

# Center for Foreign Language Education and Research

## Journal of Foreign Language Education and Research

### Manuscript Submission Guidelines

Established on April 14, 2020  
Revised on September 27, 2022  
Revised on March 13, 2023

#### **(Eligibility)**

- Article 1 Contributions to the journal are primarily limited to individuals affiliated with the Center for Foreign Language Education and Research (hereafter “FLER”). In the case of co-authored papers, this requirement applies only to the first author. Exceptions may be made for special editions.
- 2 A maximum of one contribution per issue is accepted (co-authored papers are also counted as one contribution). Due to space limitations, your submission may be considered by the Journal and Research Committee (hereafter “Committee”) for publication in a later issue, or you may be asked to reduce the length of the submitted article.
- 3 Work submitted to the journal should not have been previously published and should not be under consideration for potential publication by other journals.

#### **(Language)**

- Article 2 In order to effectively share knowledge and research activity amongst FLER-affiliated instructors, we accept, in principal, manuscripts written in either English or Japanese. Manuscripts written in German, French, Spanish, Chinese, Korean, or Russian may be accepted.

#### **(Content and Formatting Guidelines)**

- Article 3 Upon submission, authors are required to select one of the categories listed below for their contribution. Appropriateness of the selected category will be reviewed by the Committee. You may be requested to change the submission category, if necessary.

##### **(1) Research Articles**

An article which can be considered as making a substantial contribution to the relevant field through explicit analysis of data or theoretical discussion which leads to a clear and sound conclusion. The author is expected to include an introduction section in order to situate the research in reference to previous empirical research and/or theoretical discussion and a clear conclusion section well-connected to the research findings or discussion presented.

##### **(2) Research Briefs**

An article which reports an aspect of a research project underway at the time of

writing and demonstrates the potential of the topic for generating further questions of interest or prospective contributions in the relevant field.

(3) Book/Article Reviews

A critical review of a published book or article relevant to the field.

- 2 Please ensure to follow all formatting guidelines listed below. Authors who intend to write in Japanese should refer to the Japanese version of this document. When submitting an article written in English, be sure to format the article using the supplied template. For languages other than English, format the article whilst referring to the template. Submissions that do not follow these formatting guidelines may be returned to the author(s) for revision.

(1) Size

Use A4-sized paper, leaving margins of 25mm on all sides. The font used should be Times New Roman 12 point and single-spaced. For Japanese, use (MS *Mincho*), for any other languages, use the font type considered standard in the selected language. Use the same font consistently throughout the manuscript unless otherwise noted. The line spacing should be set at single-line spacing, and the text should be left-justified.

(2) Length

The number of words (characters) shall be determined for each language as follows. In all languages, include all graphs, charts, reference list, and any appendices in the word (character) count. Graphs and charts should be embedded in the text. However, if it is difficult to do so, please submit as a separate file, but leave space and indicate where they should be placed in the text.

	Research Articles	Research Briefs	Book/Article Reviews
Japanese	Approximately 20000 characters	Approximately 6000 characters	Approximately 2500 characters
English	Approximately 8000 words	Approximately 3000 words	Approximately 1000 words
German	Approximately 60000 characters (30 pages)	Approximately 18000 characters (9-10 pages)	Approximately 7500 characters (3-5 pages)
French	Approximately 8000 words	Approximately 3000 words	Approximately 1000 words
Spanish	Approximately 8000 words	Approximately 3000 words	Approximately 1000 words
Chinese	Approximately 13000 characters	Approximately 4000 character	Approximately 1600 characters
Korean	Approximately 20000 characters	Approximately 6000 characters	Approximately 2500 characters
Russian	Approximately 8000 words	Approximately 3000 words	Approximately 1000 words

(3) Title

The title should be in 18 point bold, left-justified, with Title-Case. Font as above.

(4) Author's name

The name of the author/s should be indented to the right side and written in Gothic 12 point. Leave one line between the title and the name of the author/s.

(5) Abstract

Abstracts should be approximately 500 characters in Japanese or 150-250 words in English, with 3 to 5 keywords for the article at the bottom. For the abstract, the entire text should be indented 15mm from the left and right and written in Times New Roman 11 point.

(6) Footnotes

Footnotes should be placed at the bottom of each page, in 9 point.

**(Manuscript Submission)**

Article 4 To submit a manuscript, fill in the necessary details in the designated Google form, upload the electronic file containing the article content, and submit.

**(Footnotes and Referencing)**

Article 5 The author is responsible for consistently adhering to APA (7<sup>th</sup> edition). If the manuscript is written in any language other than English, adopt APA style format as much as possible, and make sure that the style used is consistent throughout the manuscript. Make sure to avoid committing plagiarism.

**(Call and Deadline for Submission)**

Article 6 Submissions begin at the beginning of every spring semester. The deadline for submission is the last day of the first week in June.

**(Peer Review)**

Article 7 All submissions to the research article category will be reviewed by one anonymous reviewer. The review is conducted with a designated rubric and the reviewer may provide you with comments and suggestions for further improvement of your work. After the review is completed, review results will be sent to the author from the Committee before the end of Spring semester. No peer review will be conducted for research briefs or book/article reviews. However, the Committee will check if the submitted work meets the journal submission guidelines and determine if the work is publishable.

**(Revision and Resubmission)**

Article 8 After receiving the results of peer review, the author is expected to revise the manuscript for resubmission. The revised manuscript must be submitted by the end of the first week in Fall semester. A final review will be conducted by the Committee to determine if the work is publishable. The author will be notified of the decision once the final review is completed. The author may be asked to further revise the manuscript if there is any stylistic/format issue.

**(Journal Publication)**

Article 9 The journal is published annually in December.



**(Registration on CiNii and Rikkyo Repository)**

Article 10 Contributions to the Journal of Foreign Language Education and Research will be published on the FLER webpage and registered on the national CiNii database and the Rikkyo University Academic Repository.

**(Other Conditions)**

Article 11 Other terms and conditions are set out below.

- (1) No remuneration is offered to the author(s).
- (2) The copyright of articles published in the Journal of Foreign Language Education and Research resides with the Center for Foreign Language Education and Research, Rikkyo University. However, the author(s) retains the right to use his/her work for future research and/or educational purposes without permission.
- (3) If any plagiarism or misconduct is discovered after the work is published, the published work will be removed from the journal.

**(Amendment or Abolishment of Guidelines)**

Article 12 Amendment or abolishment of these guidelines will be made by the Dean of the Center for Foreign Language Education and Research after deliberation at the Faculty Meeting.

These Guidelines shall come into effect as of April 14, 2020.

These Guidelines shall come into effect as of September 27, 2022.

These Guidelines shall come into effect as of March 13, 2023.

## 【執筆者・Authors】

### **Heather Woodward**

Heather Woodward is a Lecturer in English Education at Rikkyo University. She has been teaching higher education since 2019 and is currently conducting research on L2 spoken fluency and mobile application development.

### **Joshua Rappenecker**

Joshua Rappenecker is an Adjunct Lecturer at Rikkyo University. He has developed and taught higher education courses in Japan for more than 10 years and is currently researching L2 spoken fluency development. He designs and builds software to assist with classroom management and formative feedback.

### **Kevin Thomas**

Kevin Thomas is an Adjunct Lecturer at Rikkyo University. He enjoys all his classes at the university, especially his self-created Lecture and Discussion course on British music and society. The development of this course led to his interest in content and language integrated learning (CLIL) and how students can be given language support to engage with challenging authentic material. He wishes to continue his research on CLIL and would appreciate the opportunity to collaborate with teachers with similar interests.

### **Michael Green**

Michael Green is an Adjunct Lecturer at Rikkyo University. He has a Ph.D. from Cardiff University in Applied Linguistics, and his research interests include L2 vocabulary acquisition.

### **李 菲 (リ フェイ)**

立教大学外国語教育研究センター教育講師（中国語）。研究分野は、中国語学と認知文法。現代中国語（北京語）の様々な文法現象について、認知文法の理論と観点からとらえ、記述することを中心に研究を行う。これまで、中国語の結果補語と離合詞に関する論文を公表している。近年、言語習得の観点から見た文法記述のあり方、構文から見た文法体系の再構築といった研究課題にも関心をもつ。

### **Anna Belobrovy**

Anna Belobrovy is currently employed as an Adjunct Lecturer at Rikkyo University. She has vast teaching experience in tertiary English education and has taught in multiple academic institutions in Japan. She is a graduate of a Master's program in TESOL from Teachers' College, Columbia University. Her research interests are learner development, learner autonomy, and CLIL policy and evaluation. She is presently pursuing her Ph.D. in Higher Education at Lancaster University.

### **Michael Peragine**

Michael Peragine is an Adjunct Lecturer in the Center for Foreign Language Education and

Research at Rikkyo University. He holds an M.A. in English and Creative Writing, a B.A. in Education focusing on Literature, and teaching certifications in both California and New York (Pre K–6 and English 7–12). His current research fields are peer feedback, student motivation, and extensive reading.

### **Christopher Mattson**

Christopher Mattson is an Adjunct Lecturer in the Center for Foreign Language Education and Research at Rikkyo University. He holds an M.S. in Teaching, Learning, and Curriculum from Drexel University, and a B.A. in English and German from the University of Arkansas. His research interests are literature, film, and extensive reading.

### **Paul Curran**

Paul Curran is an Adjunct Lecturer at Rikkyo University. He has also taught at Waseda University and Tokyo University of Foreign Studies. He has a B.A. (Hons 1st) in English Literature (James Cook University), an M.A. in Creative Writing (University of Queensland), and an M.A. in Applied Linguistics and TESOL (University of Leicester). His writing has appeared online and in print in various international publications. His novels include *Left Hand* (2020) and *Generation Bloodbath* (2022).

---

外国語教育研究ジャーナル 第4巻  
(JOURNAL OF FOREIGN LANGUAGE EDUCATION AND RESEARCH, Vol. 4)

---

発行日 2023年12月11日

発行者 立教大学 外国語教育研究センター  
(Center for Foreign Language Education and Research, Rikkyo University)  
〒171-8501 東京都豊島区西池袋3-34-1

製作 株式会社 遊文舎  
〒101-0061 東京都千代田区神田三崎町2-20-7 水道橋西口会館8F

# JOURNAL OF FOREIGN LANGUAGE EDUCATION AND RESEARCH

Vol. 4

December 2023



**RIKKYO UNIVERSITY**  
Center for Foreign Language  
Education and Research