〈原著論文〉

Shadowing Practice Through Xreading:
Does it Help Develop Speaking Abilities of Japanese EFL Learners?

Yuka MURAOKA, Robert J. S. ROWLAND

Abstract

This study examined how shadowing practice using an online extensive reading program called *Xreading* would help improve the speech performance of Japanese second language (L2) English learners. Nine participants engaged in shadowing practice twice a week for 15 to 30 minutes each, over four weeks. Their L2 speech performance was measured using three speaking tasks as follows: oral reading, picture description, and opinion expression tasks. A shadowing skill test was also administered to assess if their shadowing skill had improved. The speech data were analyzed in terms of complexity, accuracy, and fluency. A significant improvement was found in oral reading. Furthermore, fluency in the opinion expression task improved with large effect sizes. These results indicate that shadowing was at least effective for developing faster articulation.

Key words: Shadowing, Xreading, Complexity, Accuracy, and Fluency

Introduction

With the rapid progress of globalization and the spread of English as a lingua franca, acquiring the ability to communicate in English is essential. Needless to say, English has served as an international language to connect people from different parts of the world. Responding to this situation, the goal of English education in Japan has been to have students develop an ability to use English for real communication.

A variety of books have been published to introduce a wide range of communication activities easily conducted in classrooms. However, there is little clarity in the literature about how these activities affect the development of learners' L2 fluency. The present study focuses on *shadowing* as an oral practice method and explores how this method is useful for promoting

人文学部・欧米文化学科 論文受理日 2021 年 11 月 18 日

fluent English L2 speech performance. Shadowing is easy to practice and has been widely employed in English classrooms in Japan. Therefore, the investigation of shadowing as a research theme has not only empirical but also practical significance.

In what follows, shadowing will be explained in terms of its theoretical background and literature review. An explanation of speaking measures and materials follows. Then, the study method, analysis, results, and discussion will be presented.

Theoretical Background

Definition of Shadowing

Shadowing has been mainly recognized as a method to develop auditory processing in listening (Tamai, 2005; Hamada, 2017). Recently, Kadota (2019) has expanded the definition of shadowing to indicate "a technique for enhancing second language (L2) acquisition, in which learners repeat speech aloud as they hear it, as precisely as possible, while continuing to listen attentively to the incoming speech" (p. xiii). This definition suggests that shadowing is not only useful for facilitating listening process but also L2 acquisition processing as a whole.

Shadowing involves online processing since students must listen to a stream of sound and immediately repeat what is heard as accurately as possible, with a minimum time lag. Consequently, learners normally pay attention to phonological aspects rather than semantic aspects, especially at the beginning of shadowing training (Oki, 2011). However, processing incoming input and reproducing heard speech by focusing on meaning at the same time become possible as students develop automatized shadowing skill. Kadota (2019) argued that shadowing at the lexical level was no longer mere repetition of sound. It involves the processing of "linguistic information including grammatical, textual, semantic, and pragmatic cues" (p. 181).

Kadota (2015, 2018, 2019) has maintained that in real communication, three tasks have to be simultaneously implemented: understanding the message, conceptualizing the response, and responding to the message. These processes have to be automatized for impromptu interaction[1]. Since shadowing requires the incessant activation of both input speech perception (phonetic encoding) and output processing (reproducing heard speech), it can theoretically facilitate the automatization of the processes involved in real-world communication (Kadota, 2015, 2018, 2019).

Previous Studies on Shadowing

A great number of studies have examined the effects of shadowing on listening and speaking

skills in Japanese contexts. Tamai (1997, 2005), a pioneer researcher in the studies of shadowing, found that shadowing was effective for improving listening ability and a faster and more accurate sound reproduction rate in L2 English learners. Tamai argued that these effects implied that shadowing could enhance the function of phonological loop in working memory, which is responsible for maintaining and processing auditory input. In a similar vein, previous research has revealed that shadowing was effective for enhancing phoneme perception processing (Hamada, 2016), and a faster articulation rate (Miyake, 2009). Regarding English levels of L2 learners, a stronger favorable effect has been found for low-listening proficiency L2 learners than for those with high-listening proficiency (Hamada, 2016; Tamai, 2005). As for the level of materials used for shadowing training, Hamada (2011, 2012) found that both easy and difficult materials were useful for improving listening. Regarding effective methods of conducting shadowing practice in the classroom, Nakayama and Suzuki (2012) and Hamada (2015) found that providing learners with a chance to monitor and reflect on their own performance in the form of self-monitoring resulted in more improvement in phoneme perception and listening comprehension skill than pair-monitoring. In sum, previous studies have shown that shadowing was a useful method for improving listening skills in general, phoneme perception in particular, and was especially useful for learners with low-listening proficiency. Such effects can be enhanced with self-monitoring.

Compared to the positive effects shadowing has been shown to have on listening, the effect of shadowing on speaking skills has not been fully explored yet and results have been rather mixed across studies. Speaking can be measured in various ways since it entails multiple cognitive components and processing. Due to this multiplicity of speaking processes, the focuses of studies have been varied. Some studies examined the effect shadowing had on pronunciation (Hamada, 2017; Rowland, 2018), prosody (Mori, 2011), and speech performance in general (Iino, 2014; Iino & Yabuta, 2013, Kaneko, 2012; Muraoka, 2017, 2018, 2019; Shimizu & Saiki, 2011). Mori (2011) found that combined training methods of shadowing with oral reading improved prosodic features such as rhythm, intonation, and final lengthening of Japanese EFL learners. Rowland (2018) found a positive effect of shadowing on the improvement of L1 interference to enhance communicative comprehensibility. Hamada (2017) focused on pronunciation aspects and examined whether shadowing training could improve comprehensibility and accent for Japanese EFL learners. The study revealed little improvement in these aspects.

Regarding the effect of shadowing training on speech performance, Shimizu and Saiki (2011) explored whether synchronized reading and shadowing practice promoted speaking, listening,

and fluency and found gradual progress in both accuracy and fluency. With a clearer research design, Kaneko (2012) examined whether pseudo-speaking tasks consisting of a series of non-communicative speaking tasks such as shadowing, elicited imitation tasks, and read-and-look-up tasks, would lead to improvements in complexity, accuracy, and fluency (CAF). The results revealed some improvements in complexity and fluency measures, even though differences were not statistically significant. Accuracy was not improved. Kaneko argued that improvements in complexity and fluency could result from pseudo-speaking training. Iino (2014) explored whether shadowing-focused practice could improve the speaking skills of Japanese EFL college students. During 30 minutes of training, a treatment group engaged in oral reading and shadowing practice; another group engaged in listening and oral reading. In analyzing oral data elicited from a picture description task, the study found a significant increase in vocabulary level and grammatical accuracy in both groups, but the treatment group yielded a larger effect size. This study implied that shadowing practice together with oral reading activities would be useful for improving L2 oral performance.

Muraoka (2017, 2018, 2019) investigated the effect of shadowing practice on L2 fluency. In the 2017 study, eight Japanese EFL female college students engaged in shadowing training for 13 weeks, as a part of class activities. Oral tests (speeches) were run three times to examine changes in their oral fluency. Fluency was measured according to six variables: speech rate, mean length of runs (MLRs), total length of pauses, self-corrections, repetitions, and filled pauses. The results did not show significant improvements in any measures. On the other hand, in the 2018 study, the same participants were engaged in an additional 11-week shadowing training. The data collection and analysis methods were almost the same as those in the 2017 study. Unlike the previous study, a significant improvement in MLRs was found. In the 2019 study, 16 Japanese female EFL learners engaged in 11-week shadowing training in classroom. The speeches were analyzed in terms of MLRs, articulation rate, and speech rate. The study found a significant improvement in speech rate.

There are some differences between the previous studies and the current study. First, the current study was not conducted in a classroom context. The participants voluntarily joined shadowing training as an extracurricular activity. Therefore, time, ways, and places to practice shadowing were not tightly controlled. Second, it adapted *Xreading* as materials for practicing shadowing. The participants themselves could choose what material to practice. Finally, it examined L2 fluency by using more varied speaking measures. The use of multiple speaking tasks should allow more precious analysis of the effect shadowing training had on their L2

speech performance.

In sum, the findings of the previous studies exploring the effect of shadowing on L2 speech performance have not been consistent probably due to differences in the procedure of conducted shadowing training, materials used, and data elicitation techniques as well as measures for analyzing speech data. In this regard, more research is necessary to explore how shadowing practice promotes overall L2 speech performance.

Xreading for Speaking Practice

Xreading is an online library of extensive readers and learning management system (LMS) through which, at the time of this writing, students have online access to over 1,100 graded readers on smartphones, tablets, or personal computers anywhere they are connected to the Internet. Instructors have access to a variety of analytics in addition to the extensive library. Through the LMS, teachers can track their students' reading progress in terms of total number of books completed, total time read, total words read, and words per minute. The majority of books available on Xreading also include audio, which can be adjusted to various speeds without sacrificing audio quality thanks to the way they have been processed.

There are some strands of research which have focused on the use of Xreading with language learners. The first strand of research has investigated the effect of extensive reading (ER) on test scores using Xreading as a medium. Lyddon and Kramer (2019), for example, examined the effect of ER on TOEIC test score gains over the course of one semester, finding a statistically significant but small effect. The second strand of research has examined the use of Xreading as a medium for material expansion in language classes. For instance, action research by Harrold (2019) suggested that activities which require learners to expand upon Xreading content can be both stimulating for the students and allow teachers to meaningfully integrate digital reading into the classroom without reliance on sustained silent reading. The third strand of research has focused on the difference in student motivation to engage in ER with physical readers versus Xreading. Walker (2020) found that leaners were more likely to develop ER habits required to reach well-established ER goals for language acquisition with Xreading than with physical readers. Cote and Milliner (2015) found that students self-reported that Xreading provided a positive experience, was easy to access and use, and kept them motivated through real-time reporting of their progress.

There are no studies to date that have explored the effect of use of Xreading on speaking skills. However, given the foundation of some studies supporting the effects of disciplined

shadowing on gains of speaking ability, and the fact that Xreading seems to be an engaging and motivating method for interacting with ER texts, positive assumptions can be made about its use in the present study.

Speaking Measures and Tasks

There are various ways of assessing L2 speech performance. This study adopted the framework of the CAF dimensions, the most frequently used measure within task-based research (Housen, Kuiken, & Vedder, 2012; Skehan, 1998, 2009, 2014). The CAF have been used as dependent variables to describe general L2 performance and these components capture L2 proficiency which is multi-componential in nature (Bulté & Housen, 2012). Shadowing has been claimed to promote the whole speech production processes, not just the speed of delivery; therefore, the CAF is a suitable measure for the analysis.

Skehan (1998, 2009, 2014) suggested that advanced and challenging use of words and structures can lead to complexity. Gains were associated with risk-taking and development, but with increased possibility of errors. Avoiding and controlling errors in speaking resulted in greater accuracy. Higher accuracy was likely to be accompanied by halting and slow performance since L2 learners relied on the rule-based system, rather than the automatized system. When L2 learners acquired greater control over their language system, they would produce fluent speech. Fluency was correlated with proceduralization, automatization, less hesitation and pausing, and fewer errors.

The present study employed three types of speaking tasks: a read-aloud task, a picture description task, and an opinion expression task. They are called *Oral Reading*, *Picture*, *Opinion* respectively henceforward. Oral Reading is useful to examine how quickly and accurately learners can engage in phonological encoding. Comparing Picture with Opinion, the former is supposed to be cognitively less demanding than the latter since learners do not have to think about what to talk about and there are semantic contexts they can rely on. On the other hand, in Opinion, learners need to decide what ideas they are going to generate and choose appropriate words to express such ideas. The above suggests that the learner's speech be more complex, accurate, and fluent in Picture than in Opinion.

Research Questions

Based on the arguments put forth in the previous sections, this study has the following three research questions:

(1) Is shadowing training using Xreading useful for developing L2 speech performance?
(2) To what extent does shadowing training improve L2 speech performance in terms of the CAF?
(3) Are there any differences in the effects of shadowing on different speaking tasks?

Method

Participants

The participants were nine Japanese college students, five females and four males, whose majors were political science, child studies, and psychology[2]. Their ages were from 19 to 21. Their English levels were high-to-low intermediate, according to their English instructors' and their self-evaluation. Three students had studied abroad in an L1 English context; but the period of stay was no longer than one month. All of them volunteered to participate in the study and signed a consent form.

Research Schedule

The study began with an explanation session of the study. First, participants learned the research purpose and filled out a questionnaire. Then they were introduced to what the Xreading program was and how they used this system. Next, they watched a video which showed how to practice shadowing through Xreading and filled in a questionnaire. After the session, they were individually invited to take the first speaking test, called *Test 1* hereafter, in a quite room, using a computer with headphones. After *Test 1*, they began shadowing training at their own pace using Xreading. When participants practiced shadowing, they were instructed to use *the listening mode* on Xreading. They were asked to practice shadowing at least twice a week, 15 to 30 minutes each time for four weeks. Participants' performances were monitored and they were given regular feedback. Researchers gave struggling participants encouragement as necessary. After the training was completed, the participants individually took the second speaking tests, called *Test 2*, and filled out the second questionnaire.

Data Collection

As explicated in the previous section, this study employed three speaking tasks. The tests were conducted before and after the shadowing training (Test 1 and Test 2). The speaking tasks from each test are called *Oral Reading 1* and *2*, *Picture 1* and *2*, and *Opinion 1* and *2*

henceforth. In addition, in order to confirm whether their shadowing skill had actually improved through shadowing practice, shadowing skill tests were conducted (*Shadowing 1* and *2*, hereafter).

The test sessions were conducted by using Microsoft Power Point slides in which all the tasks ran automatically until the end. Participants' voices were recorded with recorders installed on the computer. The whole test took about 15 minutes.

As for Shadowing, the participants shadowed a conversation between a man and a woman. It consisted of 103 words and lasted 56 seconds, 110 words per minute. The conversation was taken from *Dialogue Basic 1200*, written by Akiba and Mori (2012). This book targets obtainment of 650 points on TOEIC. The same conversation was used for Test 1 and 2.

The texts for Oral Reading were taken from the GTEC past test (ALC, 2019). First, the participants read the instruction "You are going to introduce today's guest speaker through school public address system at school you are studying abroad. Read the following sentences clearly so that listeners can understand what you say." This instruction was helpful to establish a communicative situation for oral reading. Then, after 30 seconds of preparation, they started reading the text. Oral Reading 1 consisted of 68 words. Flesch Reading Ease of the text was 74.22 and Flesch-Kincaid Grade level was 5.36. As to Oral Reading 2, it consisted of 65 words. Flesch Reading Ease of the text was 83.91 and Flesch-Kincaid Grade level 4.25. These figures indicate that the two texts were similar in terms of readability.

Picture was also taken from the same GTEC past test. At first, the following instruction was provided: "You are going to talk with your foreign friend about what one boy experienced the other day. Speak clearly so that your friend can understand the story." Then, the next slide showed a cartoon scrip with four pictures. In Picture 1, the story is about a mother who lost her ring and her son who found the ring under the carpet while cleaning the floor. In Picture 2, a man and a woman were waiting at a bus stop. Then, an English speaking man came to the bus stop and asked him a question in English, but he could not make himself understood. The woman instead could answer the question in English. Compared to Picture 1, Picture 2 seemed to be more complicated in that three people with different genders and nationalities appeared in the story and that there were three interactional situations involving different speakers with different purposes each time. The participants were given 30 seconds of preparation before describing the pictures for one minute.

Opinion was created based on the past speaking tests of Eiken 2nd grade (Obunshya, 2019). To match the format of the test to the other sections, we included the instruction "You are going

to make a presentation about the following theme in an English class. First, express your opinion and then tell reasons. Speak clearly so that listeners can understand what you say." Then, in the next slide, the participants read a short passage which delineated the background of the problem, a question about the problem, and instruction on how to express their opinion. After one minute of preparation time, they made a short speech for one minute. The themes were about stress people feel in daily life in Opinion 1 and about putting personal information on the Internet in Opinion 2. While speaking, they could see the passage as a prompt appearing on the slides in both tests.

Data Analyses

In this study, the participants were instructed to shadow while listening to the story; therefore, the hours of listening were counted as the hours of shadowing. Reading time, shadowing time (listening time), words read, and books read were also counted.

For analyzing shadowing, the 30 seconds shadowing speech from the beginning were extracted, which resulted in 81 syllables. Then, one of the researchers checked whether each syllable was correctly produced by using Pratt version 6.1.39, a free software to analyze speech sound, developed by Boersma and Weenink. This analysis was conducted twice by the same researcher.

For Oral Reading, the time to read a passage was analyzed by using Pratt. Then, the reading time was reanalyzed as words per minute (wpm).

The speech data taken from Picture and Opinion were transcribed. Then one of the researchers examined the places and length of pauses within the speech to make baseline data by using Pratt. The study counted a pause exceeding 0.4 seconds, following Skehan (2014). Next, pruned data were created by excluding filled pauses, false starts, incomplete words, repetitions, and corrections. Then, the pruned data were segmented into the Analysis of Speech Unit (AS-unit) (Foster, Tonkyn, & Wigglesworth, 2000). AS-unit is a unit of analysis specially designed for spoken language data which often include incomplete sentences, hesitation, and repetitions. It is defined as "a single speaker's utterance consisting of *an independent clause, or sub-clausal unit,* together with any *subordinate clause*(s) associated with either" (p. 365). Finally, the data were analyzed in terms of CAF.

Following Norris and Ortega (2009), the current study measured syntactic complexity in terms of the mean length of AS-unit (the mean number of words produced per AS-unit). This index indicates overall complexity of speech performance. Accuracy was measured through

calculating the proportion of error-free clauses relative to the total number of clauses. This has been widely used in task-based research and proves high reliability (Kormos & Dénes, 2004). The accuracy was checked once by one researcher, and the second time together[3].

As for fluency, speech rate and MLRs were calculated. The speech rate was calculated by dividing the total number of syllables articulated in speech by the total speaking time, including pauses. This was further multiplied by sixty. The MLRs indicate the average number of syllables produced in utterances between pauses of 0.4 seconds and above. These measures have been found to be good predictors of fluency (Kormos & Dénes, 2004). All data were checked twice by the same author to ensure the reliability of the analyses.

Due to the small quantity of data and the violation of the normal distribution found in some variables, this study adapted non-parametric statistics. For comparing the means of paired samples, all the data set were submitted to the Wilcoxon signed ranks tests. The alpha level was set at .05. The effect size, Cohen's *d* was also calculated.

Results

Xreading Data

Table 1 provides overall descriptive statistics of Xreading data. For all measures, the ranges varied greatly. For instance, there was one participant who had just read one book while another had read 28 books for four weeks. These data generally indicated that there were great variations in how much time participants spent reading and practicing shadowing as well as how many words and books to read.

Table 1
Descriptive Statistics of Xreading Data

	M	*SD*	95% CI		Range
			Lower	Upper	
Reading time (min.)	205.11	269.35	−1.93	412.15	48–902
Shadowing time (min.)	97.56	60.98	50.68	144.43	32–205
Words to Read	8995.22	8225.25	2672.74	15317.71	489–26583
Books to Read	9.33	8.16	3.07	15.60	1–28

Shadowing Skill Test

Table 2 shows descriptive statistics of Shadowing. The mean scores increased from Test 1 to Test 2. The Wilcoxon signed ranks test indicated that there was a significant difference

between the two tests, $z = -2.668$, $p < .05$ with a large effect size ($d = 0.95$). This result indicates that the participants' shadowing skill had improved through shadowing training.

Table 2
Descriptive Statistics of Shadowing Skill Test

	M	*SD*	95% CI		Range
			Lower	Upper	
Shadowing 1	44.89	13.86	34.23	55.54	29–68
Shadowing 2	56.44	10.63	48.27	64.62	43–70

Speaking Measures

Oral Reading Task

The descriptive statistics of averages of wpm are displayed in Table 3. The wpm score improved from Oral Reading 1 to Oral Reading 2. The result of the statistical test indicated that a significant difference existed between the two tasks, $z = -2.310$, $p < .05$. However, the effect size was small ($d = 0.36$).

Table 3
Descriptive Statistics of Oral Reading Task

	M	*SD*	95% CI		Range
			Lower	Upper	
Oral Reading 1	106.58	26.10	86.52	126.65	68.67–154.88
Oral Reading 2	115.76	25.46	96.19	135.33	74.33–153.12

Picture Description Task and Opinion Expression Task

Table 4 displays the descriptive statistics of complexity of Picture and Opinion. The mean scores improved from Picture and Opinion 1 to Picture and Opinion 2, however an increase seems to be small. The Wilcoxon signed ranks test also showed that there were no statistical significances in both tasks, $z = -1.540$, $p > .05$, with a medium effect size ($d = 0.53$) for Picture and $z = -1.007$, $p > .05$, with a small effect size ($d = 0.31$) for Opinion. These results indicate even though the mean number of words per clause increased, such improved performance may not be due to the training the participants had received.

Table 4
Descriptive Statistics of Complexity

	M	*SD*	95% CI		Range
			Lower	Upper	
Picture 1	5.40	0.84	4.75	6.04	4.33–7.00
Picture 2	5.92	1.12	5.06	6.78	4.50–7.50
Opinion 1	4.15	2.05	2.57	5.72	0.00–6.20
Opinion 2	4.88	2.62	2.87	6.90	0.00–7.50

The descriptive statistics of accuracy are shown in Table 5. The mean scores largely decreased from Picture 1 to Picture 2, implying that the accuracy rate deteriorated, not improved. On the other hand, as for Opinion, the accuracy rate improved to a great extent. The Wilcoxon signed ranks test, however, showed no significant difference between the two tests in both tasks, $z = -1.483$, $p > .05$, with a medium effect size ($d = -0.6$) for Picture and $z = -1.352$, $p > .05$, with a medium effect size ($d = -0.52$) for Opinion. In sum, in terms of accuracy, different speaking tasks produced different results in that the error rate increased in Picture but decreased in Opinion, although not at a significant level, but with medium effect size.

Table 5
Descriptive Statistics of Accuracy

	M	*SD*	95% CI		Range
			Lower	Upper	
Picture 1	31.11	36.64	2.95	59.27	0.00–80.00
Picture 2	12.87	22.15	−4.16	29.90	0.00–62.50
Opinion 1	58.51	37.03	29.05	85.98	0.00–100.00
Opinion 2	73.33	22.61	55.96	90.71	50.00–100.00

Table 6 displays the descriptive statistics of speech rate. The mean score improved from Picture and Opinion 1 to Picture and Opinion 2. An improvement was relatively large in Opinion. The results of the Wilcoxon signed ranks test showed there was no significant difference between the two tests in Picture, $z = -1.007$, $p > .05$, with a medium effect size ($d = -0.41$). For Opinion, even though a statistically significant difference was not found, $z = -1.836$, $p > .05$, the p value was close to .05 level (.066) and a large effect size was found ($d = 1.01$). These results suggest that the speech rate had increased in Opinion 2, but not in Picture 2.

Table 6
Descriptive Statistics of Speech Rate (Fluency)

	M	*SD*	95% CI		Range
			Lower	Upper	
Picture 1	68.22	23.21	50.39	86.06	40.83–109.74
Picture 2	76.35	16.01	64.04	88.65	40.46–99.02
Opinion 1	67.73	42.99	34.69	100.77	0.00–142.86
Opinion 2	118.41	56.48	75.00	161.82	0.00–219.29

The averages of MLRs in Picture and Opinion tasks are presented in Table 7. The mean scores increased from Picture 1 to Picture 2 to a limited extent. The Wilcoxon test revealed no significant difference between the two tests. Conversely, the mean scores greatly improved from Opinion 1 to Opinion 2. Although a statistically significant difference was not found, $z = -1.718$, $p > .05$, the effect size was large ($d = 0.86$). This result indicates that the participants could produce more syllables between pauses in Opinion 2, compared to Opinion 1.

Table 7
Descriptive Statistics of Mean Length of Runs (Fluency)

	M	*SD*	95% CI		Range
			Lower	Upper	
Picture 1	3.57	0.82	2.50	5.00	2.50–5.00
Picture 2	3.82	0.97	3.07	4.57	2.42–5.00
Opinion 1	2.70	1.84	1.30	4.12	0.00–5.80
Opinion 2	4.90	3.31	2.35	7.44	0.00–12.00

Discussion and Conclusion

The first research question examined whether shadowing training using Xreading was useful for developing L2 speech performance in general. The answer for this question is positive, but in a narrow sense. The study found a significant improvement in Oral Reading. The analyses also revealed that speech rate and MLRs in Opinion improved from Test 1 to Test 2 with large effect sizes. These results imply that shadowing was useful to develop at least faster speech production. Similar results were also found in the studies by Miyake (2009) and Muraoka (2018, 2019). During the shadowing practice through Xreading, the participants were required to reproduce sound while listening to stories. Even though they could adjust the listening speed in Xreading, they still needed to catch up with the listening speed. This unique feature of speaking while listening without taking a break, mandatory in shadowing practice, may have enhanced the function of the speech articulation.

The second research question explored to what extent shadowing improved L2 speech performance in terms of CAF. As the results showed, complexity and accuracy were not significantly improved through shadowing training. However, comparing complexity with accuracy, there were some differences in the effect. The former generally revealed consistent speech production; that is, there were no prominent changes in the average number of words produced per AS-unit between Test 1 and Test 2, nor between Picture and Opinion. As for accuracy, it seems that there was some effect of a task type. In Picture, the accuracy score decreased, from Picture 1 to Picture 2. On the other hand, accuracy increased from Opinion 1 to Opinion 2, even though such differences were not statistically significant, there were medium effect sizes. These results may not be due to the shadowing training which the participants engaged in, but due to types of tasks they performed.

Several arguments can be put forward to interpret these findings. First, in Opinion, almost all participants used many chunks such as "I think." The expression "I think" is counted as one clause and this is a common expression that Japanese students have learned to use in expressing their opinion. Second, inspection of the data indicates that a priming effect might have occurred. That is, many of the participants took the phrases and expressions appearing in the instruction sentences and used them in their production. They did not recycle phrases from the instruction sentences in Opinion 1, but they did more in Opinion 2. In Picture, the instruction sentences were not provided, just a cartoon strip with four pictures. They had no language resources to rely on during their production. In sum, it can be argued that the use of a chunk phrase such as "I think" and expressions taken from the instruction sentences in Opinion may have contributed to the higher accuracy rate in this task.

Regarding fluency, speech rate and MLRs in Opinion improved from Test 1 to Test 2. Such differences were not statistically significant, but had large effect sizes. Improvement was not found in Picture. These results can be interpreted again that shadowing training was effective for reinforcing faster articulation.

The third research question examined if there were any differences in the effects of shadowing on different speaking tasks. The answer for this question is affirmative as discussed above. Shadowing training had different effects on different task types. Among the three speaking tasks, the effect of shadowing practice was best reflected in Oral Reading. The time to read given texts significantly improved in this study. However, the effect of shadowing practice on Picture and Opinion was not so explicit. Fluency was improved in Opinion with large effect sizes, but which was not statistically significant. One interpretation for this result is that both

the Picture and Opinion tasks were different from Oral Reading in that they involved automatized operation of the message generation and articulation by using learners' own linguistic resources. Kadota (2015, 2018, 2019) proposed that shadowing helps enhance L2 acquisition in general. This is through the online processing of linguistic information constantly implemented during shadowing practice. This repeated practice of shadowing may eventually result in the internalization and automatization of some linguistic information. Nonetheless, it is not clear at this point how much shadowing practice is necessary for promoting such processes. In this study, the participants were instructed to engage in shadowing for 15 to 30 minutes twice a week over four weeks. With no apparent improvements found in this study except the faster articulation, longer shadowing practice may have been necessary for improving other aspects of speaking such as message generation.

Finally, the study found that the shadowing skill improved from Test 1 to Test 2, and the difference was statistically significant. This suggests that shadowing training using the Xreading program was effective for improving shadowing skill itself to a certain extent. The improvement of shadowing skill has been reported in some studies (Tamai, 2005; Muraoka, 2019). Therefore, this study lent further empirical support to those studies. One caution is that the current study used the same shadowing text for both tests. Therefore, there is a possibility that the practice effect resulted in the positive improvement. To confirm the effect, different texts should be used in the future.

With regards to the use of Xreading as a medium for shadowing studies, the following can be considered. Although the way of shadowing practice was not controlled in this study, the participants engaged in meaningful, self-directed practice of shadowing with some positive results. This suggests that the convenience and ease-of-use of Xreading promotes motivation for task engagement, as suggested by prior research (Walker, 2020; Cote & Milliner, 2015)

There are some limitations for this study. First, without a control group to compare with, it is not certain that a positive result found in this study, such as the improvement in speech articulation, was caused solely by shadowing practice. Second, this study involves a small number of participates, only nine Japanese L2 English learners, so the results cannot be generalized to other contexts. Third, eight of the nine participants were taking regular English classes during the training; therefore, the fact that we failed to isolate the effect of this variable cannot denied. Regardless of these caveats, we hope this study made some contributions to shadowing studies specially focusing on its effect on L2 speech performance.

Notes

1. The key to the development of L2 fluency is the process called automatization. Schmidt (1992) argued that fluency is the automatized skill built from procedural knowledge. Procedural knowledge is knowledge that L2 learners can use without full attention or effort.
2. Originally, 16 students participated. However, four dropped out and one student was not included for the data analysis due to her long-term living experience in a foreign country. Two were not included because they were not students.
3. In analyzing accuracy, we found some cases which were grammatically accurate but were questionable in terms of the discourse. When such a case came up, we discussed and made a final judgement considering grammatical correctness.

References

Akiba, T., & Mori, H. (2012). *Dialogue Basic 1200* (3rd ed.). Obunsha

ALC (Ed.) (2019). *GTEC kako mondaishu* [GTEC past examination questions]. ALC.

Boersma, P., & Weenink, D. (2021). *Praat: doing phonetics by computer* (Version 6.1.56) [Computer program]. http://www.praat.org/

Bulté B., & Housen, A. (2012). Defining and operationalizing L2 complexity. In A. House, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 21–46). John Benjamins.

Cote, T., & Milliner, B. (2015). Implementing and managing online extensive reading: Student performance and perceptions. *International Association for Language Learning Technology*, *45*, 70–90. https://doi.org/10.17161/iallt.v45i1.8550

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, *21* (3), 354–375. https://doi:10.1093/applin/21.3.354

Kadota, S. (2015). *Shadowing ondoku to eigo communication no kagaku* [Science of shadowing, oral reading English communication]. Cosmopier

Kadota, S. (2018). *Gaikokugoo hanaseruyouninaru shikumi* [Mechanism of becoming a fluent L2 speaker: How shadowing enhances language acquisition]. SB Creative.

Kadota, S. (2019). *Shadowing as a practice in second language acquisition: Connecting inputs and outputs*. Routledge.

Kaneko, E. (2012). The impact of pseudo-speaking practices on L2 spontaneous speech. *IEICE Technical Report*, *7*, 73–78.

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

Iino, A. (2014). Effects of shadowing instruction on English speaking performance and perception of shadowing. *Hosei University Tama Bulletin*, *30*, 105–121. http://doi:10.15002/00009620

Iino, A., & Yabuta, Y. (2013). Relationship among reading aloud, shadowing, and speaking. *The Chubu English Language Education Society*, *42*, 139–146.

Hamada, Y. (2011). Improvement of listening comprehension skills through shadowing with difficult materials. *The Journal of Asia TEFL*, *8* (1), 139–162.

Hamada, Y. (2012). An effective way to improve listening skills through shadowing. *The Language Teacher*, *36* (1), 3–10.

Hamada, Y. (2015). Monitoring strategy in shadowing: Self-monitoring and pair-monitoring. *The Asian EFL Journal Professional Teaching Articles*, *81*, 4–25.

Hamada, Y. (2016). Shadowing: Who benefits and how? Uncovering a booming EFL teaching technique for listening comprehension. *Language Teaching Research, 20* (1), 35–52. https://doi.org/10.1177/1362168815597504

Hamada, Y. (2017). *Teaching EFL learners shadowing for listening*. Routledge.

Harrold, P. (2019). Integrating Xreading into class time using post-reading tasks. In F. Meunier, J. Van de Vyver, L. Bradley & S. Thouësny (Eds), *CALL and complexity-short papers from EUROCALL 2019* (pp. 174–179).

Housen, A., Kuiken, F., & Vedder, I. (2012). Complexity, accuracy, and fluency: Definitions, measurements and research. In A. House, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency-Complexity, accuracy and fluency in SLA* (pp. 1–20). John Benjamins.

Lyddon, P. A., & Kramer, B. (2019). Connecting extensive reading to TOEIC performance. In F. Meunier, J. Van de Vyver, L. Bradley & S. Thouësny (Eds), *CALL and complexity-short papers from EUROCALL 2019* (pp. 257–262).

Miyake, S. (2009). Cognitive processes in phrase shadowing: Focusing on articulation rate and shadowing latency. *JACET Bulletin, 48*, 15–28.

Mori, Y. (2011). Shadowing with oral reading: Effects of combined training on the improvement of Japanese EFL learners' prosody. *Language Education & Technology, 48*, 1–22. https://doi:10.24539/let.48.0_1

Muraoka, Y. (2017). Shadowing and fluency: How does shadowing practice promote fluency in the EFL classroom? *Keisen University Bulletin, 29*. 93–112.

Muraoka, Y. (2018). Shadowing to L2 ryuchosano koujoheno kouka: Chokiteki kansatsu karano kensho [Shadowing and its effect on the development of L2 fluency: The perspective of a long-term investigation]. *The Journal of Seigakuin University, 31* (1), 15–30. http://doi:10.15052/00003515

Muraoka, Y. (2019). The effect of shadowing training on the development of second language fluency. *Seigakuin University Bulletin, 66*, 46–69.

Nakayama, T., & Suzuki A. (2012). A study on learning strategies in shadowing training. *Journal of the Japan Association of Developmental Education, 7* (1), 131–140. https://doi.org/10.18950/jade.7.1_131

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics, 30* (4). 555–578. https://doi.org/10.1093/applin/amp044

Obunsha (Ed.) (2019). *Eiken 2kyu kako 6kai zenmondaishu* [Eiken 2nd grade past examination questions]. Obunsha.

Oki, T. (2011). shadowing kaishikiniokeru gakushushano fukushou sutoratejino bunrui [Classification of learners' repetition strategies in shadowing at the beginning stage]. *KATE Journal, 25*, 33–43. https://doi.org/10.20806/katejournal.25.0_33

Rowland, R. J. S. (2018). The Effect of Shadowing on Perception and Production of Function Words. *The Journal of Seigakuin University, 31* (1), 109–123.

Schmidt, R. (1992). Psychological Mechanisms Underlying Second Language Fluency. *Studies in Second Language Acquisition, 14* (4), 357–385.

Shimizu, K., & Saiki, Y. (2011). Shadowing wo mochiita chukyusha gakushushano seikakusato ryuchosa kyokatasukuno kenshou [Form focus and fluency tasks using shadowing for intermediate learners]. *Japanese Language Education Methods, 18* (1), 40–41. https://doi.

org/10.19022/jlem.18.1_40

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency and lexis. *Applied Linguistics*, *30* (4), 510–532. https://doi:10.1093/applin/amp047

Skehan, P. (2014). The context for researching a processing perspective on task performance. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 1–26). John Benjamins.

Tamai, K. (1997). Shadowing no koka to chokai process ni okeru ichizuke [The effectiveness of shadowing and its position in the listening process]. *Current English Studies*, *36*, 105–116. https://doi.org/10.11293/jaces1962.1997.36_105

Tamai, K. (2005). *Listening shidohou to shite no shadowing no kouka ni kansuru kenkyu* [Study on the effect of shadowing as an instructional method for listening]. Kazama

Walker, R. J. (2020). Successful users of Xreading: Toward a million words. *Extensive Reading World Congress Proceedings*, *5*, 23–31.

Xreadingを活用したシャドーイング
――日本人英語学習者のスピーキング能力をどのように向上させるか？――

村　岡　有　香，ローランド・ロバート・J. S.

抄　録

本研究では、オンライン多読プログラムであるXreadingを活用したシャドーイングの練習が、どのように日本人英語学習者のスピーキング力向上に役立つかについて検証した。9名の参加者が週2回､4週間にわたり毎回15～30分程度のシャドーイング練習に従事した。彼らの第二言語(L2)スピーチパフォーマンスは3つのスピーキング・タスク（音読、絵の描写、意見表明）によって評価された。また、参加者のシャドーイング･スキルが上達したか確認するために、シャドーイング･スキル・テストも実施した。スピーチ・データは、複雑さ、正確さ、流暢さの3つの観点から分析を行った。その結果、音読において有意な伸長が見られた。また、意見表明タスクにおける流暢さの向上で高い効果量が見つかった。これらの結果は、シャドーイングは少なくとも英語での発話速度を伸ばす上で効果的であることを示唆している。

キーワード：シャドーイング，Xreading，複雑さ，正確さ，流暢さ