Shadowing and Fluency: How does Shadowing Practice Promote Fluency in the EFL classroom?

Yuka Muraoka

Abstract

This study explores how shadowing training can promote more fluent speech production by second language (L2) learners. The key cognitive process of L2 fluency is proceduralization (automatization). To make fluent delivery possible in L2, explicit linguistic knowledge needs to be converted into implicit and automatized knowledge. It is theoretically argued that the process of proceduralization can be promoted through shadowing practice. The present study tests this theoretical claim. The participants were eight Japanese female college students who are studying English as L2. They engaged in shadowing training for eleven weeks. Oral tests were provided three times to examine changes in their oral fluency. Fluency was measured according to six variables: speech rate (unpruned and pruned), mean length of runs, total length of pauses, self-corrections/T-unit, repetitions/T-unit, and filled pauses/T-unit. The first three are temporal variables and the other three are related to disfluency markers. In addition, as secondary data, the study qualitatively analyzed participants' responses in a questionnaire about the effects of shadowing training. The results of the analysis did not show any significant improvements. Discussion is presented regarding how shadowing can be introduced in the classroom to promote L2 fluency.

Keyword : shadowing, fluency, proceduralization
キーワード：シャドーイング、流暢さ、手続き化

One of the important current goals of English education in Japan is for students to acquire oral fluency for global communication. This is because communicating with people from different cultural backgrounds has recently been seen as essential in order to improve education and economy in a modern society. For this purpose, Japanese students need to acquire automatized skills of putting thoughts into words

to carry out a spontaneous conversation within a limited time. Such practical commutative skill is especially desired in English education at the college level. In order to meet the needs of a globalized society where communicating in L2 is unavoidable, teachers who engage in English education in Japan have to think seriously about how students' L2 fluency can be promoted even in English as Foreign Language (EFL) settings. In this regard, this study explores possible effects of shadowing practice on the development of fluency.

In what follows, the theoretical backgrounds of fluency and shadowing are presented. Then the exploration of how fluency can be promoted through shadowing is explicated. Next, the methodology, results, discussion, and conclusion will follow.

Fluency in SLA

There are fluent speakers and less fluent speakers of L2; and we usually judge if someone is fluent in speaking L2 or not through listening to their speech. However, what we exactly mean by L2 fluency is difficult to define in technical terms. In the field of second language acquisition (SLA) studies, Lennon (1990) conducted the first study investigating L2 fluency. Lennon argues that there are two senses of fluency: a broad and a narrow sense. In a broad sense, fluency indicates general oral proficiency. L2 learners with high fluency often get a higher score; those with low fluency get a lower score. In a narrower sense, fluency can be regarded as one component of oral proficiency which consists of other variables such as correctness, appropriateness, pronunciation, and lexical range. Therefore, it is possible to conclude that someone can speak a given language fluently, but his or her grammar is not accurate. He claims that fluency is a performance phenomenon and is closely related to "an impression on the listener's part that the psycholinguistic processes of speech planning and speech production are functioning easily and efficiently" (p. 391). In a later study, Lennon (2000) summarized earlier definitions and study findings of fluency and suggests that "a working definition of fluency might be the rapid, smooth, accurate, lucid, and efficient translation of thought or communicative intention into language under the temporal constraints of on-line processing" (p. 26). He maintains that this definition can be applicable not only to native speakers but also to L2 speakers.

Measurement of Fluency

Fluency has been generally measured by examining two variables: temporal features such as the number of syllables per minute, mean number of syllables between pauses, and the length of pause and disfluency markers such as self-corrections or repetitions. Even though the findings are not always consistent, the measures that best predict fluency are speech rate, the mean length of runs, and phonation-time ratio (Derwing, Rossiter, Munro, & Thomson, 2004; Freed, Segalowitz, & Dewey, 2004; Fujio, 2011; Kormos, 2006; Kormos & Dénes, 2004; Taguchi, 2008; Towell, Hawkins, & Bazergui, 1996; Wood, 2001). Furthermore, it is argued that filled and unfilled pauses as well as repetitions, restarts, and repairs are not clear indicators of fluency (Kormos, 2006). For instance, Lennon (1990) examined the spoken performance of four advanced learners and how it changed after long-term residence in Britain. The study found that there was an increase in self-corrections and suggests that self-correction may not be a good indicator of fluency (Derwing et al., 2004).

In brief, fluent speech entails such features as few pauses (pause length), lack of hesitation, production of relatively long speech between pauses (mean length of runs), and high speed of delivery (speech rate).

Psycholinguistic Processes Underlying Fluency

Defining fluency in cognitive terms, it is closely related to the mechanism of speech production and language learning theories. It is the aspects of speech production that are responsible for delivery. The development of fluency involves the learning of new grammar and vocabulary and automatized use of that knowledge in authentic communication.

The most well-known speech production model was proposed by Levelt (1989, 1999) and adapted by de Bot (1992) to describe its specific relevance to L2 use. Levelt (1989) proposes five components involved in the understanding and generation of human speech: conceptualizer, formulator, articulator, audition, and speech-comprehension system[1]. Among these five components, the first three elements are mainly responsible for production. The generation of messages takes

place in the conceptualizer by accessing two kinds of knowledge: procedural knowledge and declarative knowledge (e.g., encyclopedic knowledge, situational knowledge, and discourse knowledge). The output of the conceptualizer is called a *preverbal message*, which becomes the input to the formulator. Within the formulator, the preverbal message, which is a conceptual unit, is converted into a linguistic structure in two stages. First, grammatical encoding, consisting of procedures for accessing lemmas and syntactic building procedures in the mental lexicon, operates on the preverbal message. The product of the grammatical encoding is called the *surface structure*, which is "an ordered string of lemmas grouped in phrases and subphrases of various kinds" (Levelt, 1989, p. 11). Second, morpho-phonological encoding operates on this surface structure to build a phonetic or articulatory plan (internal speech) by accessing lexical form, which contains information about a lexicon item's morphology and phonology. The phonetic plan generated by the phonological encoding is then translated into overt speech by the articulator. Both overt speech and internal speech eventually enter the speech-comprehension system to enable recognition of words and retrieve their meanings by accessing the form information and the lemma information in the lexicon.

As to learning theories involved in L2 fluency, Schmidt (1992) attempts to explain how fluency develops by describing a psychological learning mechanism. He argues that fluency is an automatic skill depending on procedural knowledge (processing) rather than declarative knowledge. Procedural knowledge indicates that L2 learners can use such knowledge without full attention or efforts, which results in automatization; declarative knowledge is conscious knowledge requiring a great deal of attention in application. L2 speech is a quite complex cognitive task involving message creation, transformation of a message into appropriate expressions, and actual articulation. Each process requires some level of attention. It is only when these processes are automatized that fluent L2 speech production is possible. Therefore, the degree of automatization of speech processes is the key to the development of L2 fluency. Fluent speech is based on automatic procedural knowledge and processing.

Still, an important empirical question is how L2 fluency can be developed-that is to say, how declarative knowledge or processing can be converted into procedural

knowledge or processing. Schmidt (1992) and Towell et al. (1996) single out Anderson's model of Adaptive Control of Thought (ACT) and argue that this model could account for how declarative knowledge can be transformed into procedural knowledge. Precise descriptions of Anderson's model are beyond the scope of this study; but it is generally argued that the power law of practice is the most prominent characteristic of skill development (Schmidt, 1992). That is to say, rehearsal or repetition in meaningful contexts is the key for automatization. However, it has not been clear what kinds of practice are effective to develop L2 fluency and how. In this line, one suggestion is proposed by Kadota (2007, 2011, 2012, 2014, 2015). He claims that shadowing can promote oral fluency in meaningful ways, as introduced in the next section.

Fluency and Shadowing Practice

Shadowing is one of the methods adapted to train simultaneous interpreters in Japan for fluency development and this method has recently been utilized in Japanese English classroom. It involves the verbal reproduction of sound as accurately as possible soon after it is heard. When shadowing is practiced by listening to sound and looking at a written text simultaneously, this is called *parallel reading* (Kadota, 2007). In this study, the participants engaged in shadowing with or without a text. Moreover, Kadota (2012) explains that there are two different kinds of shadowing: bottom-up and top-down shadowing. In bottom-up shadowing, L2 learners engage in shadowing before they study the structures and words of a passage. On the other hand, in top-down shadowing, they practice shadowing after studying the target text. Therefore, the former is phonology-focused rehearsal in which the process of sound perception is promoted. The latter is knowledge-focused rehearsal in which memorization of vocabulary can be facilitated. In this study, since the participants engaged in shadowing training after they studied the vocabulary and structures of given materials, it can be presumed that the type of shadowing was top-down.

Kadota (2007, 2011, 2012, 2014, 2015) argues that shadowing plays an important role in improving the listening ability, speech production ability, learning of new linguistic items, and process of proceduralization, which is the basis of fluent

performance as argued by Schmidt (1992) and Towell et al. (1996). According to Kadota, to understand incoming aural input, human beings first need to perceive sound. In perceiving sound, it is necessary to transform it into phonetic or phonological representation in order to process it further. The next stage is understanding. Understanding involves lexical, syntactic, semantic, contextual, and schema processing. Shadowing is especially helpful in automatizing the perception stage because it requires the immediate repetition of incoming sound without taking time to activate the process of understanding. When a shadowing task is automatized, this implies that speech processing- the conversion of sound into phonetic representation- is automatized. This further leads to automatic search of speech knowledge in the mental lexicon in the long-term memory and the reconstruction of that database including native-like phonetic and phonological representation. If automatization in speech processing is accomplished, understanding of aural input can be accelerated, resulting in the improvement of listening ability.

In addition, shadowing can also enhance the ability to articulate sound which is close to that of the target language. Once the reconstruction of speech knowledge and automatization of accessing such knowledge in the mental lexicon has been attained, the results will be the articulation of L2 sound with native-like prosody and with faster articulation. In short, shadowing practice could develop fluency in terms of automatic sound perception and native-like and faster speech articulation (Miyake, 2009; Mori, 2011).

Kadota (2007, 2011, 2012, 2014, 2015) also maintains that shadowing can help internalize new information such as grammatical structures or words in the long-term memory. It is generally known that the human memory system consists of the working memory and the long-term memory. The working memory has the function of temporarily retaining perceived information and processing it by accelerating the long-term memory. The phonological loop, one composite of the working memory, plays a major role in retaining and processing verbalized input. In order to maintain input in the working memory and further store it in the long-term memory, the activation of the subvocal rehearsal in the phonological loop within a limitation of two seconds is necessary. In other words, if perceived information is not rehearsed within two seconds through subvocal rehearsal in the phonological loop, it will

disappear from the working memory. This in turn implies that the internalization of new information such as words or numbers in the long-term memory is not possible without subvocal rehearsal. As delineated above, shadowing practice can promote the learning of new words by automatizing the processes of phonological representation and phonological coding. The automatizing of these processes further accelerates the speed and efficient use of subvocal rehearsal. When the speed of subvocal rehearsal is quickened, a larger amount of input can be rehearsed within two seconds and maintained in the working memory. This can lead to greater possibilities of new words being internalized in the long-term memory.

Moreover, Kadota (2015) argues that shadowing practice could not only promote the learning of new words, but also the proceduralization of explicit memory. As is stated, shadowing practice has an effect on the acceleration of articulation speed. Faster articulation speech means that more rehearsal of input information in the phonological loop is possible within a limitation of two seconds. If more rehearsal of input can be done within two seconds, more learning of new words and the proceduralization of already-learned-yet-not-fully-automatized linguistic knowledge are possible. This is because subvocal rehearsal in the phonological loop is an essential cognitive process which is critically involved in the conversion of explicit knowledge into implicit knowledge.

All these arguments indicate that shadowing practice is one plausible method to promote fluent performance in L2 together with improvements in the ability to listen, articulate sound, and learn new words as well as the proceduralization of explicit knowledge or memory.

A large number of studies have examined the effects of shadowing on listening skills (Kojima & Ota, 2012; Nakayama & Mori, 2012; Tamai, 2005), speed of articulation (Miyake, 2009), English prosody (Mori, 2011), motivation for English learning (Shiraki, Yasukawa, Yoshida, & Sasaki, 2008), and L2 learners' reaction to shadowing training (Horiyama, 2012). One clear benefit found from these studies is that shadowing is an effective learning method to improve listening skill. However, its effect on the development of the production ability has not been fully explored yet. Even though several studies examined how shadowing could help improve L2 learners' speaking fluency (Kaneko, 2012; Shimizu & Saiki, 2011), more research is

necessary.

Research Questions

The present study focuses on the following two research questions:

(1) Can shadowing practice facilitate the development of L2 fluency?

(2) What were the participants' reactions to shadowing training?

Method

Participants

The participants of the study were eight female Japanese college students studying at women's university in Tokyo, Japan. All of them were 19 years old at the time of the study. Their major was English and their proficiency levels of English ranged from low- to high-intermediate, based on their self-evaluation and the researcher's observation. All of them, except one, had study-abroad experiences of one to four months in English-speaking countries such as Canada, America, and England. Originally, 44 students were enrolled in the shadowing training sessions; however, only eight had less than three absences from the training sessions and took all three tests. Their voices were also clear enough to be analyzed.

Material

The material used in the current study was *Dialogue Basic 1200*(3rd edition). This textbook contains 171 dialogues using 1204 basic words that L2 learners should know to obtain a score of 650 on TOEIC. Most of the dialogues are up to one minute long. CDs are attached to the book, so students can listen to the dialogues while reading them. The book was selected because it contains basic words which can be used for communication and 171 interesting topics which can help maintain L2 learners' motivation to study.

Procedure

The data was taken from an intact class where the focus was on the development of oral fluency in English. Training sessions lasted for 13 weeks, with an additional two test periods. Each training session was 90 minutes long, divided

into five parts: (1) a review task (15 minutes), (2) vocabulary and grammar explanation (15 minutes), (3) top-down shadowing practice (30 minutes), (4) pair reading (15 minutes), and (5) a dictation task (15 minutes). The time for each part was controlled; however, it was often the case that the explanations of vocabulary and grammar took more than 15 minutes. If this happened, the time for the pair reading was reduced to assure 30 minutes for shadowing practice.

In the review task, the participants checked how well they remembered words from the previous training session. Then, they received explicit instruction on key vocabulary and grammar contained in the materials they were going to shadow next. The participants worked individually on desktop computers with headphones, all connected to a master computer. First, they inserted their CDs and listened to a dialogue while trying to shadow. Then, they listened again and shadowed, but this time paying closer attention to prosody features such as pronunciation and intonation. The first two rounds of shadowing were synchronized shadowing since they could look at the text. Lastly, they shadowed again without looking at the text. The shadowing of each dialogue with these steps took about three to five minutes. In a pair-reading task, they paired up with their partners and took turns reading the dialogues. Finally, a dictation task was provided using one dialogue of the ten. In the dictation task, there were ten blanks the participants had to fill out within the context of sentences in a dialogue. Before the task, the students were given five minutes to study for the dictation task.

For collecting data, they were instructed to record their voices individually using an application called Sound Recorder in Windows. After they finished recording, they sent this oral data to the master computer. The data was collected three times in total; a Pretest, Posttest 1, and Posttest 2. The Pretest was administered on the first day of class, Posttest 1 in the 8thweek, and Posttest 2 in the 15thweek. In each test, the participants were instructed to make a speech on a given topic for one minute. First, they practiced how to record their voices using Sound Recorder and how to save the data on their computers. Then they were given a topic. After spending one minute planning what they were going to say, they were told to start recording at the same time. The topics were “My spring break” (Pretest), “My last weekend” (Posttest 1), and “My summer vacation” (Posttests 2). Time was controlled

using a time watch.

Fluency Measures

The collected oral data was transcribed and analyzed by examining two components: temporal variables and vocal disfluency markers. The temporal variables analyzed in the present study were (1) speech rate (unpruned and pruned), (2) mean length of runs (MLRs), and (3) total pause length. These measures were selected because they have been suggested as good indicators of fluency (Kormos, 2006; Lennon, 1990; Taguchi, 2008; Towell *et al*. 1996; Freed *et al*. 2004; Derwing *et al,* 2004; Fujii & Tomoda, 2005; Segalowitz, 2010). As to disfluency markers, (4) self-corrections/T-unit, (5) repetitions /T-unit, and (6) filled pauses/T-unit were examined, following Lennon (1990).

Each variable was defined in the following terms: for calculating the speech rate, the total number of syllables was divided by the total amount of time spent speaking including pauses, then multiplied by sixty. For pruned data, false starts, repeated syllables, and the self-corrected syllables were excluded from the total number of syllables. To measure MLRS, the average number of syllables produced in utterances between pauses of more than 0.4 second was calculated[2]. Finally, pause length was operationalized as the percentage of unfilled pauses of more than 0.4 seconds out of the total time of speech delivery.

With respect to disfluency markers, self-corrections/T-unit, repetitions/T-unit, and filled pauses/T-unit were defined as the total occurrences of these variables, divided by the total number of T-units[3]. Repetitions were operationalized as words immediately repeated in production. An example of the repetitions can be found in the utterance such as "and speak *speak* and writing in English." In this utterance *speak* was repeated, so this word was excluded from the data. Self-corrections were operationalized as a word immediately corrected by the speaker. An example of self-corrections is seen in the utterance: "I will practice to listening *listen*." In this case, the word *listening* was corrected to *listen*. For self-corrections, the word *listen* (corrected word) was excluded to make pruned data. Filled pauses were defined as non-silent and non-lexical pauses such as *um, er*, and *ah.*

Analyses

The recorded speech samples were first transcribed by the researcher. Then, it was segmented into syllables and t-units. For analyzing temporal measures, the study used the speech analysis software Sound Creation (Onsei Koubou) developed by NTT AT. Sound Creation displays waveforms of sound segments and allows the detailed analyses of the length of each segment. For examining disfluency markers, the number of repetitions, self-corrections, and false starts was counted and categorized accordingly.

To compare the means of the temporal and disfluency measures, the Friedman Test was used. A nonparametric test was used because the sample size was small (only eight cases) and normality was not assumed in several sets of data. Still, the data shows similar variability across distribution. All the statistical analyses were conducted using SPSS software. The alpha level was set at 0.05.

Results

Analysis of Fluency (Temporal Measures)

Table 1 shows the descriptive statistics for the four temporal measures: unpruned and pruned speech rate, MLRs, and total pause length. As seen in Figures 1 and 2, the means of the speech rate and MLRs declined from the Pretest to Posttest 2, implying that improvements in fluency were not observed in these measures. As to speech rate, differences in the means between unpruned and pruned were not so great in any of the tests; subsequently, these two measures did not differ greatly in terms of analyzing how many syllables were produced per minute. Comparing the means of the tests, declines were clearer from the Pretest to Posttest 1 than from Posttest 1 to Posttest 2. This tendency can be seen especially in the MLRs, with a mean of 5.01 in the Pretest, 3.22 in Posttest 1, and 3.12 in Posttest 2. Contrary to the speech rate and MLRs, as Figure 3 shows, the means of the total pause length increased from the Pretest (M = 54.58) to Posttest 1 (M = 58.02), but decreased a bit in Posttest 2 (M = 55.68). This result indicates that the participants produced more pauses in Posttest 1 than in the other two tests.

The results of the Friedman Test did not show significant differences in any of the temporal measures: speech rate (unpruned), $X^2(2, N= 8) = 3.25$, $p> .05$; speech

rate (pruned), $X^2(2, N= 8) = 1.00$, $p> .05$; MLRs, $X^2(2, N= 8) = 5.25$, $p> .05$; or total pause length, $X^2(2, N= 8) = 0.25$, $p> .05$. In sum, there was no significant improvement in any of the measures.

Table 1. *Descriptive Statistics of Temporal Measures*

	Pretest (Week 1)		Posttest 1 (Week 8)		Posttest 2 (Week 15)	
	M (SD)	Range	*M (SD)*	Range	*M (SD)*	Range
Speech Rate (unpruned)	63.61 (5.63)	48.75-95.20	58.36 (5.94)	40.67-81.01	57.36 (7.62)	28.62-84.09
Speech Rate (pruned)	59.26 (5.91)	47.02-92.07	56.01 (5.95)	35.74-81.01	51.81 (7.28)	27.70-81.58
Mean Length of Run (MLRs)	5.01 (1.06)	2.77-11.38	3.22 (0.52)	2.23-6.67	3.12 (0.19)	2.21-3.74
Total Pause Length	54.58 (4.58)	24.61-65.06	58.02 (4.44)	36.17-74.92	55.68 (4.35)	36.27-68.86

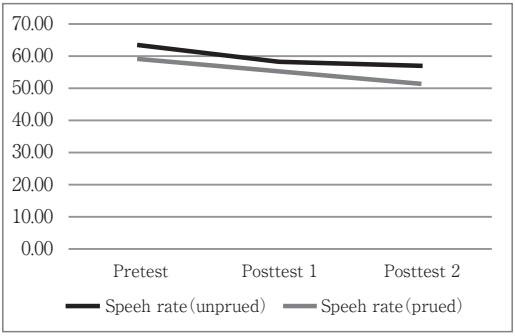

Figure 1. Means of Speech Rate (unpruned & pruned)

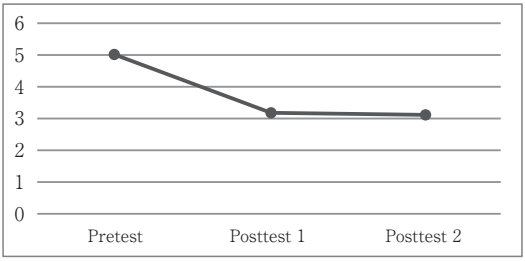

Figure 2. Means of Mean Length of Runs (MLRs)

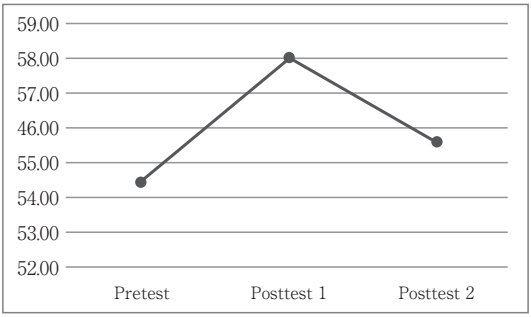

Figure 3. Means of Total Pause Length

Analysis of the Disfluency Measures

Table 2 shows the descriptive statistics for the three disfluency markers: self-corrections/T-unit, repetitions/T-unit, and filled pauses/T-unit. As Figure 4 displays, the means went up from the Pretest to Posttest 2 in all the disfluency measures, indicating that the participants made more self-corrections, repetitions, and filled pauses in the Posttest 2 compared to Pretest. Regarding the repetitions and filled pauses, the means declined a bit from the Pretest (0.26 for repetitions and 0.36 for filled pauses) to Posttest 1 (0.21 for repetitions and 0.34 for filled pauses); however, they increased again in Posttest 2.

Being similar to the temporal measures, the results of the Friedman Test showed no significant differences in any of the disfluency measures: self-corrections/T-unit, $X^2(2, N= 8) = 0.5$, $p> .05$; repetitions/T-unit, $X^2(2, N= 8) = 0.24$, $p> .05$; filled pauses, $X^2(2, N= 8) = 0.33$, $p> .05$ for MLRS; or pause length, $X^2(2, N= 8) = 0.25$, $p> .05$. In brief, this study found no improvement in disfluency markers.

Table 2. *Descriptive Statistics of Disfluency Markers*

	Pretest (Week 1)		Posttest 1 (Week 8)		Posttest 2 (Week 15)	
	M (SD)	Range	*M (SD)*	Range	*M (SD)*	Range
Self-corrections/T-unit	0.08 (0.14)	0.0-0.40	0.10 (0.16)	0.0-0.43	0.16 (0.24)	0.0-0.60
Repetitions/T-unit	0.26 (0.35)	0.0-0.83	0.21 (0.34)	0.0-0.29	0.38 (0.32)	0.0-0.80
Filled pauses/T-unit	0.36 (0.31)	0.0-0.75	0.34 (0.34)	0.0-0.75	0.40 (0.5)	0.0-1.25

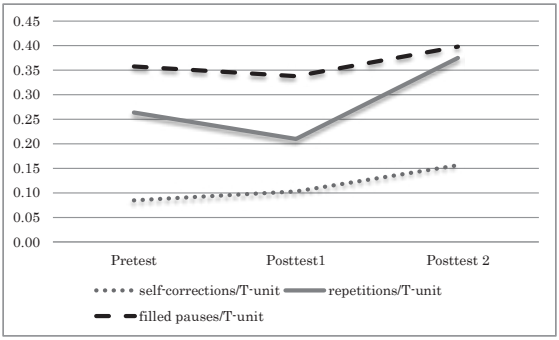

Figure 4. Means of Disfluency Markers

Questionnaire Results

Responses to a questionnaire were obtained from 35 students, including eight whose oral data was analyzed for fluency variables. All of them participated in shadowing training for 13 weeks. The questionnaire contained the following five questions: (1) Did you work hard in the class?; (2) Did you enjoy conducting shadowing practice in class?; (3) Do you think shadowing training is useful to improve your English skills?; (4) Do you think shadowing training improves your fluency in English?; (5) Do you think shadowing training helped increase your vocabulary knowledge? The participants chose their responses from the five-point Likert scale: *strongly agree, agree, neither, disagree, strongly disagree*. The responses were first coded into the numbers 5 to 1 and then submitted them for further analysis.

Table 3 shows the averages and detailed percentages of the five questions contained in the questionnaire. Question 3 had the highest average score (3.94) and the precise percentages of each scale show that 74 percent of the participants strongly agreed or agreed with the statement that shadowing training was useful to improve their English skills in general. The second highest average score was found in Question 1 (3.74), indicating that 74 percent of them claimed that they worked hard on shadowing practice in class. The rest of the questions displayed similar averages: 3.60 for Question 2, 3.64 for Question 3, and 3.57 for Question 5. When looking at the precise percentage of each response, different response patterns were

observed. In Question 2, 66 percent of the respondents strongly agreed or agreed that they liked shadowing training; however, only around half of them (49 percent for Question 4 and 57 percent for Question 5) strongly agreed or agreed that their fluency in L2 improved or that their vocabulary knowledge increased through shadowing practice. Around 30 percent of them were not sure about the effects of shadowing on these measures.

Table 3. *Responses in Questionnaire*

	Question 1	Question 2	Question 3	Question 4	Question 5
Averages	3.74	3.60	3.94	3.64	3.57
Strongly agree	2 (6 %)	3 (9 %)	8 (23 %)	4 (12 %)	3 (9 %)
Agree	24 (68 %)	20 (57 %)	18 (51 %)	13 (37 %)	17 (48 %)
Neither	7 (20 %)	8 (23 %)	8 (23 %)	13 (37 %)	12 (34 %)
Disagree	2 (6 %)	3 (8 %)	1 (3 %)	5 (14 %)	3 (9 %)
Strongly disagree	0	1 (3 %)	0	0	0
Total	35 (100 %)	35 (100 %)	35 (100 %)	35 (100 %)	35 (100 %)

In sum, the analyses of the responses to the questionnaire indicated that around 70 percent of the participants thought that shadowing training was useful to improve their English level and that they liked shadowing practice, so they worked hard. However, only about half of them agreed that their fluency and vocabulary knowledge improved.

Discussion

If fluency is developed through shadowing practice, the following changes could be observed in learners' performance: there would be an increase in speech rate and MLRs. There would be a decrease in pause length, repetitions, and filled pauses. Regarding self-corrections, as the study by Lennon (1990) argues, there would be an increase in self-correction.

The results of the present study did not match the above assumptions in most cases. With respect to temporal measures, there was a slight decrease in speech rate and a moderate decrease in MLRs from the Pretest to Posttests 1 and 2. This indicates that the participants' speed of articulating English was not especially

accelerated and that the number of syllables produced between pauses reduced slightly. Averages of pause length increased slightly from the Pretest to Posttest 1, but decreased in Posttest 2. This suggests that the participants produced a similar amount of pauses in their production after three months. Regarding disfluency markers, there was a slight decrease from the Pretest to Posttest 1, but a small increase from Posttest 1 to Posttest 2. If the process of proceduralization had occurred, the participants should have produced fewer repetitions and filled pauses. However, the data showed contrary results: the participants repeated more syllables and produced more filled pauses in Posttest 2. In terms of self-correction, the amount increased gradually from the Pretest to Posttests 1 and 2. This is the only measure that was congruent with the results of previous studies maintaining that self-correction is not related to fluency.

Still, due to a number of limitations in this study, the results should be interpreted with caution and it cannot be concluded that shadowing training method has a minimal effect on the development of L2 fluency. First, even though the measures of fluency, temporal markers and disfluency markers were appropriate to examine changes in oral performance, the elicitation task was not sufficient. The present study used only speech samples; multiple tasks including a picture description task or a communication task involving interactions between participants should have been employed to obtain more comprehensive data. Also, the selection procedure of the participants might have affected the results. The data collection was conducted for about three months, but many students could not be included in the analysis since some of them did not take one of the tests or did not participate fully in the training sessions. Only eight participants met the condition of taking all three tests and having less than three absences. Seven of the eight happened to have studied English in Canada for one month prior to this experiment. Even though seven weeks have passed since they came back to Japan, the fluency they developed through study-abroad experiences in Canada still remained at the time of Pretest. This, however, implies that their frequent performance had deteriorated over time.

Furthermore, it is possible that the speech topics might have affected their performance. In the Pretest, the topic was *My spring break*, and many of them had many things to talk about since they had just spent a month in Canada. Although the

topics of the Posttests were similar such as *My weekend* and *My summer vacation,* their tone of speaking indicated that they did not know what to talk about because they had not done anything worthy of talking about or did not have a concrete plan. In this regard, how speech topics influence the quality and quantity of fluency is worth investigating in the future.

Besides these limitations, there seem to be flaws in the research design itself. It can be argued that mere engagement in shadowing practice might not be so effective to promote L2 fluency. More focused shadowing practice is necessary. For example, Kadota (2015) proposed three types of speech production components on which L2 fluent performance is based: rule-governed sentence production, primed sentence production, and formulaic sentence production. Among the three, formulaic sentence production is argued to contribute most to fluent speech (Kadota, 2007). Formulaic chunks can reduce the consumption of cognitive resources in the formulator. Consequently, L2 learners can use more cognitive resources to plan, formulate, and articulate their messages. Future research can examine how shadowing practice plus chunk learning incorporated into communicative activities can promote fluency.

Conclusion

The present study explored whether shadowing training implemented once a week for 13 weeks could promote L2 fluency. The study showed some changes in means in three tests; however statistically significant differences were not found in any of the temporal and disfluency measures. Based on these results, it can be concluded that the shadowing practice the participants had engaged for three months did not promote either fluency or the process of proceduralization in the production model, at least as measured by the variables used in the current study and the elicitation task. It is possible that measurement using different tasks and using different measures, including a larger number of participants, would have produced different results.

Notes

1. Levelt (1999) later reframed and revised the processing systems of production. This study uses his 1989 model since this model is still widely used in L2 studies.

2. Riggenback (1991) suggested pauses of shorter than 400 ms are within the range of normal speech.
3. T-unit indicates a unit consisting of one independent clause together with whatever dependent clauses are attached to it (Richards & Schmidt, 2010).

References

de Bot, K. (1992). A bilingual production model: Levelt's' speaking model adapted. *Applied Linguistics, 13*(1), 1 -24.

Derwing, M. T., Rossiter, J. M., Munro, J. M., & Thomson, I. R. (2004). Second language fluency: Judgments on different tasks. *Language Learning, 54*(4), 655-679.

Freed, F. B., Segalowitz, N., & Dewey, P. D. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition, 26,* 275-301.

Fujio, M. (2011). A longitudinal study on the fluency of Japanese learners studying in the UK. *JACET-KANTO Journal, 7,* 15-26.

Horiyama, A. (2012). The development of English language skills through shadowing exercises. *Bulletin of Bunkyo Gakuin University and Junior College, 12,* 113-123.

Kadota, S. (2007). *Shadowing to ondoku no kagaku* [Science of shadowing, oral reading, and English Acquisition]. Tokyo: Cosmopier.

Kadota, S. (2011). Inputo to autoputo o tsunagu shadowing: Riron to Jissen no Renkei [Shadowing connecting input and output: Association of theory and practice]. *JACET Chubu Bulletin, 9,* 41-55.

Kadota, S. (2012) *Shadowing to ondoku to eigoshutoku no kagaku* [Science of shadowing, oral reading, and English acquisition]. Tokyo: Cosmopier

Kadota. S. (2014), *Eigo jotatsu 12 no pointo* [12 points for improving English]. Tokyo: Cosmopier

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

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

Kojima, S. & Ota, S. (2012). Shadowing, dictation and reading aloud: Which is effective? *JACET-TOHOKU Journal, 4,* 29-40.

Kormos, J. (2006). *Speech production and second language learning*. Mahwah, NJ: Lawrence Erlbaum.

Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System, 32*(2), 145-164.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning, 40*(3), 387-417.

Lennon, P. (2000). The lexical element on spoken second language fluency. In H. Riggenback (Ed.), *Perspectives on fluency* (pp. 25-42). Ann Arbor: University of Michigan Press.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* Cambridge, MA : MIT Press.

Levelt, W. J. (1999). Producing spoken language: A blueprint of the speaker. In C. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 83-122). Oxford University Press.

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.

Nakayama, T. & Mori, T. (2012). Efficacy of visual-auditory shadowing. *The Society of English Studies, 42,* pp. 55-68.

Richards, J. & Schmidt, R. (2010). *Longman dictionary of language teaching and applied linguistics*. New York: Pearson

Riggenbach, H. (1991). Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse processes*, *14*(4), 423-441.

Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition, 14,* 357-357.

Shimizu, K. & Saiki, Y. (2011). Form focus and fluency tasks using shadowing for intermediate learners. *Japanese Language Education Methods, 18*(1), 40-41.

Shiraki, S., Yasukawa, K., Yoshida, K., & Sasaki, M. (2008). The effects of shadowing

training on arising motivation for English learning, *Studies on Communication, 6*, 27-36, Research Institute for Communication, Kansai University of International Studies.

Taguchi, N. (2008). Building language blocks in L 2 Japanese: Chunk learning and the development of complexity and fluency in spoken production. *Foreign Language Annuals, 41*(1), 132-156.

Tamai, K. (2005). *Listening shidoho to shite no shadowing no koukani kansuru kenkyu* [Study on the effect of shadowing as a instructional method for listening]. Tokyo: Kazama-shoo.

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics, 17*(1), 84-119.

Wood, D. (2001). In search of fluency: What is it and how can we teach it? *Canadian Modern Language Review, 57,* -589.