# The Journal of Asia TEFL

http://journal.asiatefl.org/

# Using Auditory Word Repetition to Improve L2 Pronunciation of English Schwa by Japanese Learners: From the Perspective of Phonological Processing

**Kaori Sugiura**
*Ritsumeikan University, Japan*

This study investigates the effects of immediate repetition of auditory words on L2 pronunciation improvement of English schwa by Japanese EFL learners. More specifically, the study is concerned with how the amount of input (i.e., five or ten repetitions) and stimuli characteristics (i.e., the position of schwa in a word, word familiarity) influences the repetition effect. The study consists of a study and a test phase: In the study phase, participants listened to and repeated words. In the test phase, they repeated spoken stimuli including previously heard as well as new words, both ten minutes and one week after the study phase. The auditory materials consisted of 32 real words. The duration ratio of schwa to a stressed vowel in a word and the quality of the schwa were acoustically examined. A repetition effect was mainly found in the durational aspect. Repeating the stimuli more than five times did not differentiate the effect. One week after the study phase, the effect persisted only for initial schwa, which is weakly represented in the participants' mental representation. The pedagogical implications for L2 pronunciation teaching from the view of language processing are discussed.

**Keywords: L2 pronunciation improvement, English schwa, immediate repetition, auditory words, phonological processing**

## Introduction

A large amount of L2 pronunciation research has directly examined the effect of specific pronunciation instruction methods and has provided practical and useful insights for the ESL/EFL classroom; however, these studies are likely influenced by a variety of factors, including teachers, students, materials, and classroom environments. By controlling these influential factors, the present study sheds light on L2 pronunciation instruction from the perspective of language processing. That is, this study focused on L2 learners' sensitivity to phonological information in L2 input when learning L2 pronunciation by manipulating the amount (i.e., the number of repetitions) and characteristics (i.e., word stress patterns and word familiarity) of the input during auditory repetition tasks.

The effect of the immediate repetition of auditory information on L2 learning has been widely examined by

shadowing activities (i.e., learners listen attentively to incoming sentence-level information and vocalize it simultaneously) and other oral repetition tasks. Previous research has highlighted the potential effectiveness of the repetition of oral speech, even in adult L2 learners (e.g., Hori, 2008; Mori, 2011; Ofuka & Gilbert, 2013); however, during these activities, several factors of auditory input, such as discourse, semantics, structure, and word familiarity, may influence the effects of repetition on pronunciation. Thus, the issue that remains unclear is whether these effects occur due to the repetition of auditory input alone. To exclusively investigate the immediate auditory repetition effect on L2 pronunciation improvement, the present study used word-level stimuli in the framework of an auditory priming experiment, which is described in detail in this article.

The target sound investigated in the present study is the English sound schwa produced by EFL Japanese learners. A reduced vowel schwa appears most frequently in unstressed syllables in stress-timed languages, which are characterized by alternations of stressed and unstressed syllables (Bolinger, 1965). The phonetic characteristics of schwa are shorter duration and lower pitch and intensity compared to a vowel in a stressed syllable (Wallace, 1994). The quality of schwa is greatly influenced by the adjacent phonetic environments (e.g., Kondo, 2000). The present study focused on two phonetic aspects of schwa: the duration ratio of schwa to a stressed vowel in a word (henceforth, the duration ratio) and the quality (the first and second formant frequencies of schwa, henceforth F1 and F2), both of which influence the creation of English rhythm (e.g., for the duration ratio, Beckman, 1986; for the quality of the schwa, Grabe & Low 2002).

The English schwa sound has not been extensively explored in L2 production and perception research; however, the English schwa sound produced by L2 leaners should receive more attention in L2 pronunciation studies. Numerous studies have shown that the accurate production of prosody, including rhythm that involves schwa, by L2 leaners has a more serious impact on the perceived foreign accent and intelligibility to native listeners than segmental sounds (e.g., Anderson-Hsieh, Johnson, & Koehler, 1992; Moyer, 1999; Munro & Derwing, 1999; Pennington & Richards, 1986; Tien, 2013). Also, research that investigated Adult EFL learners with an L1 syllable-timed rhythm background has shown that the transfer of L1 rhythm to the L2 rhythm has a possible impact on reducing the learners' speech intelligibility (Tien, 2013).

On the other hand, some studies that focused on Japanese leaners of English have shown that segmental factors, such as consonant deletion (Suenobu, Kanzaki, & Yamane, 1992) and vowel errors (Kashiwagi & Snyder, 2010), significantly hinder intelligibility (for the former, intelligibility for native English listeners, and for the latter, intelligibility for both native English and Japanese listeners). Moreover, from the perspective of international communication, Jenkins' (2000) Lingua Franca Core proposes that the instruction of weak forms including schwa is unnecessary because those produced by L2 learners do not affect intelligibility among non-native speakers. However, as mentioned previously, a large number of studies have proposed the importance of prosody in intelligibility for native listeners. Most importantly, many L2 leaners desire to acquire native-like pronunciation (Derwing, 2003; Scales, Wennerstrom, Richard, & Wu, 2006). Thus, it is worthwhile for them to try to produce a native-like English rhythm, including pronouncing the English schwa sound accurately.

Despite the importance of learning English rhythms, in general, L2 learners have difficulty acquiring it due to the transfer from their L1 language. For Japanese English learners, there is a significant difference between the English and Japanese rhythm systems. English has a stress-timed rhythm in which schwa plays a crucial role, as previously mentioned, while Japanese has a mora-timed rhythm in which each syllable is of roughly equal length and the reduced vowel schwa is not required (Vance, 1987). In addition, Japanese does not possess a phoneme similar to schwa. Therefore, Japanese leaners tend to reflect their L1 phonology in their L2 English production (see Ueyama, 2000, for her acoustic study of prosodic transfer between English and Japanese).

In summary, the present study aimed to examine whether Japanese learners can improve their pronunciation of the English schwa sound through the immediate repetition of auditory words in a laboratory setting and to

determine the amount of input that should be used and whether or not input characteristics (i.e., word stress patterns and word familiarity) enhance pronunciation.

## Literature Review

### Auditory Word Repetition

Repetition of auditory words involves the auditory priming effect (Church & Fisher, 1998), a subconscious phenomenon in which previously encountered auditory input is processed faster and more accurately than previously unheard input (implicit learning). This implicit learning occurs because the phonetic/phonological information contained in spoken words is encoded and stored in listeners' implicit memory, facilitating subsequent speech processing. Previous studies have revealed that just one instance of hearing of a word can facilitate subsequent speech perception of the word (e.g., Hori & Sugiura, 2013; Sugiura & Hori, 2012; Trofimovich, 2005; Trofimovich, 2008; Trofimovich & Gatbonton, 2006). Since there is a strong relationship between speech perception and production (Galantucci, Fowler, & Turvey, 2006) in speech acquisition, increased repetition of presented auditory words can raise learners' sensitivity to phonetic/phonological information in L2 auditory input, and this greater sensitivity is expected to promote pronunciation accuracy as well. In addition, the repetition effect lasts from several minutes to days or even weeks (Church & Schacter, 1994; Goldinger 1996), and this characteristic seems to play a role in the constant development of speech processing (McDonough & Trofimovich 2009, p. 24).

To examine the immediate repetition effect on L2 pronunciation learning, the present study employed the psycholinguistic experimental paradigm of auditory priming. This type of experiment consists of study and test phases: In the study phase, participants encode novel phonetic and phonological information, and in the test phase, they are asked to retrieve learned information (Trofimovich & Gatbonton, 2006). In this experiment, participants are expected to pronounce words more accurately if they have been exposed to them before (here, through immediate repetition). To the author's knowledge, no previous studies have employed the framework of auditory priming experiment to investigate the role of repetitive practice in improving L2 pronunciation.

### English Schwa: Difficult to Pronounce for Japanese Learners of English

Japanese learners of English have difficulties in learning the English schwa sound (Akita, 2001; Kondo, 2000; Lee, Guion, & Harada, 2006; Sugiura, 2006; Tomita, Yamada, & Takatsuka, 2010) due to the difference between English and Japanese phonologies. In terms of vowels, American English includes nine phonemic vowels (Ladefoged, 1993), including full vowels /i/, /ɪ/, /ɛ/, /æ/, /e/, /ʌ/, /ʊ/, /ɔ/, and /a/ and a reduced vowel /ə/ in the vowel space. English has phonetically contrastive vowels, such as a lax vowel /ɪ/, as in *bit* /bɪt/, and a tense vowel /i/, as in *beet* /bit/. Conversely, Japanese contains only five full vowels (i.e., /i/, /e/, /a/, /o/, and /u/) (e.g., Imaishi, 1997). They are phonemically contrastive vowels in terms of length. The Japanese short vowels have long counterparts: /a/ vs. /aa/, /e/ vs. /ee/, /i/ vs. /ii/, /o/ vs. /oo/, and /u/ vs. /uu/. For instance, in *biru* (building), a short vowel /i/ is used, whereas in *biiru* (beer), a long vowel is used (Ueyama, 2000).

According to the phonetic feature hypothesis proposed by McAllister, Flege, and Piske (2002), L2 learners cannot utilize the phonetic features (i.e., pitch, intensity, duration, and quality) that they do not employ in L1 processing, and the established ways to process L1 speech thus tend to persist both in L2 processing and production. Therefore, Japanese speakers, who are sensitive to the length of vowels in their L1 languages, might find it relatively easy to pronounce the duration aspect of the English schwa but not the quality aspect. However, from the perspective of language rhythm, previous research has reported that Japanese learners tend to

pronounce schwa with inadequate shortening, reflecting their L1 with a mora-timed rhythm in which each syllable is roughly of equal length (Sugiura, 2006).

Regarding vowel quality, in addition to the fact that Japanese speakers are likely to be insensitive to the phonetic feature of quality, as mentioned earlier, Japanese does not possess a mid-central vowel that is equivalent to the English schwa; therefore, it is plausible that Japanese learners of English are inclined to mispronounce the English schwa sound in peripheral areas in the acoustic vowel space, reflecting their L1 vowels (e.g., Lee et al., 2006; Sugiura, 2006).

## Current Study

Of particular interest was the efficiency with which learners take in and integrate a target feature from language input during the repetition of auditory words and how they utilize the information to accurately pronounce the target sound. To facilitate learners' intake and integration of L2 input, two types of manipulation in auditory input were considered in the present study: the amount of input (i.e., five or 10 repetitions) and stimuli characteristics, or quality (i.e., the position of schwa in the word and word familiarity), that influence the repetition effect.

Regarding the amount of input, only minimal research on L2 shadowing has discussed the number of repetitions required for maximum improvement of target sound pronunciation, particularly related to pitch range (see Hori, 2008; Miyake, 2009b). To provide specific suggestions for target sound acquisition and learning, research should be expanded to include other phonetic features of segmental sounds and prosody. The number of repetitions was selected based on previous findings that Japanese learners of English attained maximum improvement in their pronunciation at the fifth (Hori, 2008) and tenth (Miyake, 2009b) instances out of 15 repetitions.

With respect to the position of schwa in a word (word stress pattern), a previous study showed that the acquisition of schwa is more problematic for intermediate Japanese learners of English in the initial syllable of a word (i.e., words with a weak–strong stress pattern, such as in *agenda* and *possess*), rather than in the final syllable of a word (i.e., words with a strong–weak stress pattern, such as in *medium* and *basket*) (Sugiura, 2006). It has remained unclear whether or not the weakness can be overcome by pronunciation training.

Word familiarity, which is an indicator of how often a person hears or sees a target word (Yokokawa, 2009), also potentially influences the improvement of L2 pronunciation. If learners can obtain a repetition effect with low-familiarity words, then it is hypothesized that they are able to genuinely learn from phonetic/phonological information (Murao & Sasaki, 2008).

Another topic of interest is the extent of the effect of auditory word repetition on pronunciation improvement. Research on L1 phonological processing indicates that repeated exposure to auditory input creates a long-lasting effect (days or weeks) (Church & Schacter, 1994; Goldinger, 1996); however, less attention has been focused on this issue in relation to L2 pronunciation improvement. If a long-term effect on auditory information indeed occurs, this phenomenon will contribute to the understanding of the mechanisms that govern L2 phonological acquisition and learning.

The present study focused on the production of the English schwa sound by Japanese learners of English and investigated whether Japanese learners can improve their pronunciation through the immediate repetition of auditory words considering the persistence of the effect in particular. Specifically, it aimed to determine how much input (i.e., how many repetitions) and whether and how input characteristics (i.e., word stress patterns, word familiarity) influence improvement. This study was guided by two primary research questions:

(1) Do Japanese learners of English improve their duration ratios of schwa to a stressed vowel (duration ratio) as well as the quality of schwa through auditory word repetition? If so, how do the differences in the number of repetitions (five vs. ten), the position of schwa within the word (initial vs. final syllable), and word familiarity (high vs. low) influence schwa pronunciation improvement in terms of duration ratio and vowel quality?

(2) Does the effect of repetitive practice with auditory words continue for one week?

# Method

## Participants

Twelve Japanese learners of English (nine females and three males), aged 19-21, participated in this study. All were native speakers of Japanese and were non-English-major students enrolled in their second English semester course at a university in Japan. Their English level was low intermediate, ranging from TOEIC 400 to 550 (i.e., CEFR level A2). They were paid to participate in this study. To provide a norm for the study, two male native English speakers, both instructors of English as a foreign language at Japanese universities, participated in the present study. One is from Boston, U.S.A., and the other is from Hawaii, U.S.A.

## Materials

Thirty-two words of either two or three syllables were selected from a list of 3,000 high-frequency words in *the British National Corpus* (2007) (Table 1), based on word familiarity for Japanese learners of English and schwa position within the word. To determine word familiarity, the present study used a word familiarity list that was created based on word familiarity judgments for 3,000 words by a large number of Japanese EFL students with various proficiency levels (auditory version; Yokokawa, 2009). The rating scale ranges from 1 (lowest familiarity) to 7 (highest). Regarding the position of schwa, previous studies have shown that the phonetic realization of schwa differs between initial and final syllables (Flemming & Johnson, 2007; Wallace, 1994). Therefore, the words were divided into two sets based on schwa position, with both low- and high-familiarity words in each set. The difference between high and low word familiarity was confirmed by Kruskal-Wallis tests for the initial- and final-schwa word sets (for initial schwa, high and low familiarity ratings were on average 6.0 and 2.8, respectively, $p < .005$; for final schwa, the ratings were on average 6.3 and 2.8, respectively, $p < .01$). In addition, words including schwa spellings <a>, <u>, <o>, and <ou> were used. Spellings with <e> and <i> were excluded because those schwas tend to exhibit different phonetic qualities (Flemming & Johnson, 2007).

All the words were read using the text-to-speech software *Natural Reader 10.0* (2011) and recorded with the audio-processing software *Audacity 1.3.12* (Mazzoni & Dannenberg, 2010) with a sampling rate of 44kHz. The same U.S.-accented female voice was used throughout. The words were then used to construct four study–test list pairs, with each pair containing a 16-word study list and a 32-word test list (i.e., all the words). Table 2 shows the phonetic characteristics of the auditory word stimuli produced by the text-to-speech software.

TABLE 1
*Linguistic Materials*

| Position of Schwa | Spelling Familiarity | <a> Word | Rating | <u>, <ou> Word | Rating | <o> Word | Rating |
|---|---|---|---|---|---|---|---|
| Initial | High | abroad | 6.50 | supply | 6.20 | correct | 6.32 |
| | | attend | 6.41 | support | 5.87 | control | 6.14 |
| | | again | 6.10 | suggest | 5.71 | | |
| | | alarm | 6.06 | suspend | 5.52 | | |
| | Low | facility | 2.88 | sustain | 3.58 | convey | 2.85 |
| | | award | 2.78 | supplier | 2.90 | condemn | 2.85 |
| | | absorb | 2.47 | | | | |
| | | capacity | 2.36 | | | | |
| Final | High | special | 6.71 | famous | 6.51 | station | 6.31 |
| | | camera | 6.34 | August | 6.22 | lesson | 6.14 |
| | | media | 6.27 | dangerous | 5.82 | | |
| | Low | agenda | 3.18 | enormous | 2.36 | reckon | 2.96 |
| | | opera | 2.89 | religious | 2.71 | carbon | 2.84 |

*Note.* The bolded letter indicates schwa's position.

TABLE 2
*Phonetic Information of Auditory Word Stimuli (Female Voice): Duration Ratio (ms) and the F1 and F2 (Hz) Values of Schwa*

| Phonetic features | Initial schwa | Final schwa |
|---|---|---|
| Duration ratio | 0.32 | 0.72 |
| F1 (Hz) | 626 | 671 |
| F2 (Hz) | 1711 | 1759 |

## Experimental Paradigm and Procedures

The experiment consisted of study and test phases, as described above. During the study phase, participants listened to and repeated auditory words (eight words five times, and another eight words 10 times; 120 repetitions in total) individually presented at five-second intervals using *SuperLab 4.0* (Cedrus Corporation, 2006). The words were randomized in both phases for each participant in order to minimize unwanted order effects. To ensure that all the participants concentrated on all stimuli in the study phase, they were asked to indicate on an accompanying form whether the auditorily presented words included a /t/ sound. After the study phase, the participants performed simple calculations for about ten minutes to clear their short-term memory and to create an interval between the study and test phases of the experiment. The test phase consisted of Test 1 (ten minutes after the study phase) and Test 2 (one week after the study phase). For the test phase, the participants repeated spoken stimuli as quickly as possible (Onishi, Chambers, & Fisher, 2002), including those already presented in the study phase as well as new words. Productions from the test phase were recorded using *Audacity 1.3.12* (Mazzoni & Dannenberg, 2010) with a sampling rate of 44 kHz. Note here that both the study and test phases were preceded by three practice trials. In addition, to avoid strategic learning, participants were not told that the test phase would include words from the study phase. All sessions were conducted in a quiet room. The study and Test 1 sessions lasted approximately thirty minutes. Test 2 continued for ten minutes. To obtain the norm, two native English speakers were asked to repeat 36 auditory word stimuli. The native speakers' pronunciation was recorded using *Audacity 1.3.12* (Mazzoni & Dannenbeg, 2010) with a sampling rate of 44 kHz.

## Analysis

Only words that were repeated correctly by the participants were used for the auditory measurements. For example, the production of words that differed from the target words (e.g., *alarm → alone, convey → today, special → station*), incorrect words (e.g., *enormous → endromas\*, condemn → condent\*, capacity → compacity\**), and the words with incorrect stress placement in Tests 1 and 2, 13% and 11%, respectively, were excluded from the analysis. The repetition effect was acoustically examined, using spectrograms created with the speech analysis software *WaveSurfer 1.8.5* (2011). As a guide to identify schwas in the productions by the participants, the analysis of schwa was conducted in reference to the example acoustic data of the English schwa from previous studies: A duration ratio of schwa to a stressed vowel is 0.47 in a non-final syllable, while it is 0.63 in a final syllable (Wallace, 1994), and the quality of schwas occurring in the first or middle positions of a word has an F1 of 449 Hz and an F2 of 1922 Hz, while word-final schwas are 539 Hz for F1 and 1797 Hz for F2 (Flemming & Johnson, 2007). These F1 and F2 values of schwa vowels were obtained from female speakers. It should be noted that, to avoid gender effects on the acoustic properties due to the vocal tract lengths of the participants, all formant values in the productions of the participants, including both Japanese and English speakers, were normalized to one female participant. To do so, first, the average third formant frequency (henceforth, F3) in the female speaker of /ɛ/ in *lesson*, *special*, *correct*, *agenda* and *attend* was calculated. Second, the mean F3 of this speaker was used as the norm and was divided by the mean F3 for each participant. Finally, the F1 and F2 for each participant were multiplied by the factor obtained from their own F3 (see Lee et al., 2006).

Regarding the judgment of whether the participants can benefit from the repetition effect, if the duration ratio in repeated words becomes significantly smaller than that in unrepeated words and also approaches native ratios, it can be assumed that a repetitive practice effect is obtained because Japanese learners of English tend to inaccurately produce high duration ratios (closer to 1.0) due to the transfer from Japanese mora-timed rhythm, in which each mora is of approximately equal length (e.g., Sugiura, 2006). As for vowel quality, if the F1 becomes significantly lower and F2 becomes higher in repeated words than those in unrepeated words, moving towards native norms, it can be said that a repetitive practice effect was obtained because relatively higher F1 and lower F2 values of schwa are expected from non-proficient Japanese participants, corresponding to articulatory targets closer to Japanese full vowels (e.g., Akita, 2001; Kondo, 2000; Sugiura, 2006; Tomita et al., 2010).

Statistical analyses for the present study were separately conducted for each position of schwa (initial schwa and final schwa in a word) for both duration ratio and the quality of schwa because their acoustic characteristics are intrinsically different.

## Results and Discussion

## Duration Ratio

### Initial schwa (weak-strong syllables)

Table 3 and Figure 1 illustrate descriptive statistics for the duration ratio in Tests 1 and 2 (ten minutes and one week after the study phase, respectively). In order to examine whether the initial schwa produced by participants improved through repetition, and whether the repetition effect lasted, the duration ratio data was submitted to a three-way analysis of variance (ANOVA), performed using SPSS, with Repetition (unrepeated, repeated five times, repeated ten times), Familiarity (high, low) and Interval (ten minutes, one week) as factors. The analysis yielded a significant main effect only for Repetition, $F(2, 354) = 5.029$, $p < .05$, $\eta^2_p$

= .028. A further analysis revealed that the duration ratios derived during both the five and ten repetition trials were significantly smaller than the duration ratios obtained in the unrepeated condition, $p < .05$, $d = .038$ (a small-sized effect) and $p < .05$, $d = .036$ (a small-sized effect), but no significant difference was found between the duration ratios obtained in the five and ten repetition conditions. No significant main effects were found for Familiarity, $F (1, 355) = 0.392$, $p = .532$, $\eta^2_p = .001$ or Interval, $F (1, 355) = 0.209$, $p = .647$, $\eta^2_p = .001$, and there were no significant interactions for Repetition × Familiarity, $F (1, 355) = 0.705$, $p = .495$, $\eta^2_p = .004$, for Repetition × Interval, $F (2, 354) = 1.022$, $p = .361$, $\eta^2_p = .006$, or for Repetition × Familiarity × Interval, $F (2, 354) = 0.458$, $p = .633$, $\eta^2_p = .003$.

These results suggest that the repetitions were effective, but increasing the number of repetitions to ten does not differentiate the effect on pronunciation of initial schwa. In addition, word familiarity did not seem to influence the repetition effect on pronunciation. Moreover, the significant difference in duration ratios between the repeated and unrepeated conditions, and the absence of a significant difference in duration values between Test 1 (ten minutes) and Test 2 (one week) indicate that the practice effect obtained ten minutes after the repetition-based study phase remained unchanged for one week. That is, the repetitive effect lasted one week in this particular phonological aspect.

TABLE 3
*Mean Duration Ratios of Initial Schwa (ms) for Unrepeated and Repeated Words (Five, Ten Times) in Test 1 (Ten Minutes) and Test 2 (One Week)*

|  | Test 1 (ten minutes) | | Test 2 (one week) | |
| --- | --- | --- | --- | --- |
|  | Word Familiarity | | | |
|  | High | Low | High | Low |
| Unrepeated | 0.46 (0.17) | 0.45 (0.20) | 0.45 (0.17) | 0.39 (0.29) |
| 5 times | 0.36 (0.12) | 0.39 (0.15) | 0.39 (0.17) | 0.37 (0.45) |
| 10 times | 0.38 (0.11) | 0.37 (0.18) | 0.38 (0.11) | 0.39 (0.41) |
| Native speakers | 0.37 (0.11) | 0.40 (0.13) | 0.37 (0.11) | 0.40 (0.13) |

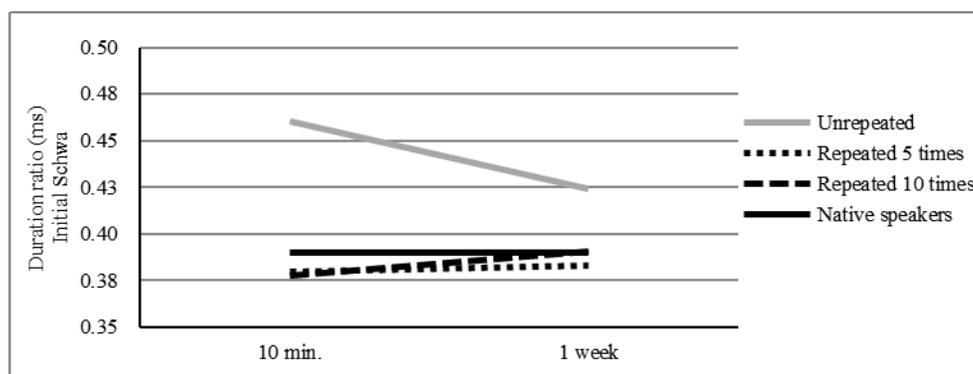*Note.* The value in parentheses is the standard deviation (SD).



*Figure 1.* Mean duration ratios for initial schwas in Test 1 (ten minutes) and Test 2 (one week). As the factor of word familiarity did not affect the repetition effect, it is not indicated in this graph.

## Final schwa (strong-weak syllables)

Table 4 and Figure 2 show descriptive statistics for the duration ratio of final schwa in Tests 1 and 2 (ten minutes and one week after the study phase, respectively). To investigate whether the final schwa produced by the participants improved as a result of the repetition task and whether the effect continued for one week, the

data on the duration ratio of final schwa were treated in the same manner as in the statistical analysis for initial schwa. The analysis yielded a significant effect of Repetition, $F(2, 251) = 10.615$, $p < .001$, $\eta^2_p = .081$. A post-hoc test revealed that the duration ratios in the five- and ten-repetition conditions in Test 1 were significantly smaller than in the unrepeated condition, $p < .05$, $d = 0.96$ (a large-sized effect) and $p < .001$, $d = 0.97$ (a large-sized effect). Therefore, there was an improvement in pronunciation immediately after practice. There were no significant effects of Familiarity, $F(1, 252) = 0.240$, $p = .625$, $\eta^2_p = .001$ or Interval, $F(1, 252) = 1.792$, $p = .182$, $\eta^2_p = .007$. There was a significant interaction for Repetition × Interval, $F(2, 251) = 5.026$, $p < .01$, $\eta^2_p = .040$, but not for Repetition × Familiarity, $F(2, 251) = 1.889$, $p = .153$, $\eta^2_p = .015$ or for Repetition × Familiarity × Interval, $F(2, 251) = 2.210$, $p = .112$, $\eta^2_p = .018$.

These results show that, again, word familiarity did not seem to influence the repetition effect on the improvement of final schwa. As for the persistent effect of repetition, for both five and ten repetitions in the study phase, the duration ratios in Test 2 (one week) were significantly larger than those in Test 1 (ten minutes), $p < .05$, $d = 0.55$ (a medium-sized effect) for five repetitions, and $p < .05$, $d = 0.80$ (a large-sized effect) for ten repetitions. That is, the repetition effect did not last one week for the final schwa.

TABLE 4
*Mean Duration Ratios of Final Schwa (ms) for Unrepeated and Repeated Words (Five, Ten Times) in Test 1 (Ten Minutes) and Test 2 (One Week)*

|  | Test 1 (ten minutes) | | Test 2 (one week) | |
|---|---|---|---|---|
|  | Word Familiarity | | | |
|  | High | Low | High | Low |
| Unrepeated | 0.72 (0.18) | 0.78 (0.27) | 0.67 (0.16) | 0.71 (0.19) |
| 5 times | 0.53 (0.10) | 0.62 (0.14) | 0.71 (0.19) | 0.55 (0.18) |
| 10 times | 0.59 (0.13) | 0.53 (0.12) | 0.69 (0.17) | 0.65 (0.18) |
| Native speakers | 0.61 (0.32) | 0.54 (0.13) | 0.61 (0.32) | 0.54 (0.13) |

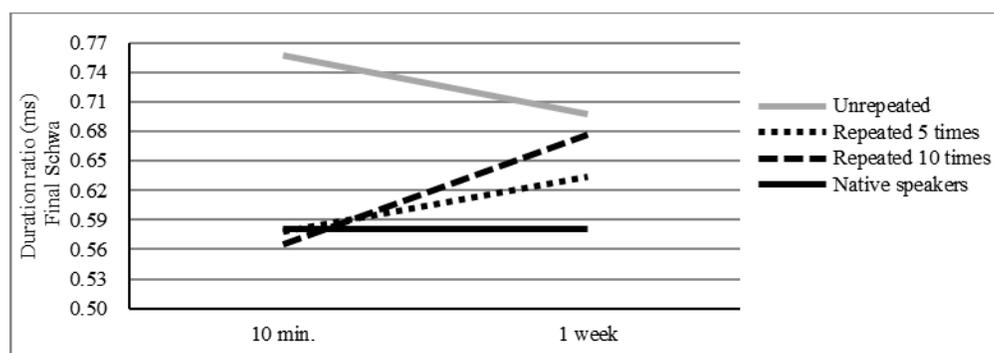*Note.* The value in parentheses is the standard deviation (SD).



*Figure 2.* Mean duration ratios for final schwas in Test 1 (ten minutes) and Test 2 (one week).

## Discussion on the Duration Ratio

### Number of repetitions

The results of this study indicate that no difference in the repetition effect was found between the five and ten repetition conditions either in the initial or final schwa sound; the participants required at least five spoken repetitions of words to improve their pronunciation of schwa in conditions in which the repetition effect was observed (all conditions, except for the final schwa in Test 2). This finding implies that more than five

repetitions would not have positively influenced the acquisition of the target sound or the persistence of the repetition effect. As mentioned, previous studies also demonstrated that the performance of Japanese learners of English reaches a maximum improvement at around the fifth (Hori, 2008) or tenth (Miyake, 2009b) instances out of 15 repetitions. Thus, it seems unlikely that the number of repetitions correlates with the degree of prosodic improvement in L2 pronunciation (pitch range or duration ratio of schwa to a stressed vowel). This non-correlation may be attributed to extra-linguistic factors, such as attentional or motivational constraints.

## Position of schwa and the persistence of the repetition effect

The results of the present study showed that the repetition effect continued for one week for the initial schwa, while a long-lasting effect was not obtained for the final schwa. This contrasting effect could have been due to differences in the frequency of the stress patterns of the target words. The final schwa (a strong–weak stress pattern) should be much more common than the initial schwa (weak–strong) among L2 speakers considering the input frequency: 90% of the content words in a corpus of over 20,000 English words followed a strong–weak stress pattern (Cutler & Carter, 1987). Thus, it can be assumed that the initial schwa (weak–strong stress) is more weakly represented in participants' mental representations. According to formal theories of learning (e.g., Rescorla & Wagner, 1972), increased learning practice has a significant impact on less familiar linguistic material. Several earlier studies on structural priming have supported this theory (e.g., Hartsuiker & Westenberg, 2000; Luka & Barsalou, 2005; Luka & Choi, 2012; Reitter, Keller, & Moore, 2011; Scheepers, 2003). In particular, Luka and Choi's (2012) study on the persistent effect of structural priming indicated that for a novel structure (i.e., less familiar linguistic material), the effect continued for one week. Although these structural priming studies did not focus on learners' sensitivity to auditory information in general or on word stress patterns in particular, their findings serve as useful scaffolds in understanding the present results, given that both syntactic and stress patterns are characterized by structural systems (i.e., regular and exceptional patterns).

## Generalization to novel words

Interestingly, the numerical data showed that a repetition effect did not occur in Test 1 (ten minutes), but it did occur in Test 2 (one week) for novel words (i.e., unrepeated words) with low familiarity and an initial schwa, which should be less common for Japanese learners of English in terms of duration. More specifically, the duration ratio in Test 1 was 0.45, but it improved to 0.39 in Test 2 when the duration ratio was 0.37 for a five-time repetition and 0.39 for a ten-time repetition of words. The native norm was 0.4. This might indicate that the effect of learning in the study phase was gradually generalized to unrepeated words over one week.

This finding can be explained by "memory consolidation," a theory that has been extensively discussed in the field of motor skill learning. It traditionally refers to a process in which a memory becomes increasingly stabilized as time passes (McGaugh, 2000). This suggests that immediately after training, skills are still only weakly consolidated in memory, but after a certain period, the repetitive training effect becomes stable (e.g., Karni, Tanne, Rubenstein, Askenasy, & Sagi, 1994). Note that this phenomenon is specifically related to procedural memory (i.e., knowledge about how to do things), which requires repetition training (Walker & Stickgold, 2006).

Consolidation has recently been investigated in the context of learning new spoken words (e.g., Davis & Gaskell, 2009; Dumay & Gaskell, 2012; Henderson, Powell, Gaskell, & Norbury, 2014). Davis and Gaskell's (2009) PET and fMRI study showed that for new words that are learned by listening and repeating (i.e., repetition), it takes some time to achieve a similar lexical status as familiar words because the novel mapping must be integrated with items of existing mapping and with several representations (e.g., semantic, phonological, and morphosyntactic). In the present study, a similar phenomenon was observed in a

phonological aspect (i.e., word stress) of words: an accurate phonological representation was generated for untrained words (i.e., new words). That is, the representation of the target sound in trained words became stronger and more solid in itself due to repetition training (Truscott, 2014). This reinforced representation was gradually combined with a new representation (the target sound in untrained words) (Truscott, 2014). Because consolidation is an essential part of L2 learning, this finding is worth consideration.

## Quality: F1 Values

### Initial schwa (weak-strong syllables)

Table 5 and Figure 3 illustrate descriptive statistics for the F1 of the initial schwa in Tests 1 (ten minutes after the study phase) and Test 2 (one week after the study phase). The data on the F1 of the initial schwa in Tests 1 and 2 were subjected to a three-way ANOVA, with Repetition (unrepeated, repeated five, repeated ten times), Familiarity (high, low), and Interval (ten minutes, one week) as factors. The analysis yielded no significant main effects of Repetition, $F (2, 317) = 0.588, p = .556, \eta^2_p = .004$, Familiarity, $F (1, 318) = 0.149$, $p = .700, \eta^2_p = .000$, or Interval, $F (1, 318) = 0.371, p = .543, \eta^2_p = .001$. No significant interactions were found for Repetition × Interval, $F (2, 317) = 0.009, p = .991, \eta^2_p = .000$, for Repetition × Familiarity, $F (2, 317)= 0.675, p = .510, \eta^2_p = .004$, or for Repetition × Familiarity × Interval, $F (2, 317)= 0.210, p = .811, \eta^2_p = .001$.

These findings indicate that no improvement in the F1 value of the initial schwa occurred in Tests 1 and 2, demonstrating that the F1 values of the initial schwa remained higher than the native norm. Thus, immediate repetition (five or ten times) with auditory words did not contribute to pronunciation improvement in terms of vowel quality for initial schwa.

TABLE 5
*Mean F1 of Initial Schwa (Hz) for Unrepeated and Repeated Words (Five, Ten Times) in Test 1 (Ten Minutes) and Test 2 (One Week)*

|  | Test 1 (ten minutes) | | Test 2 (one week) | |
|---|---|---|---|---|
|  | Word Familiarity | | | |
|  | High | Low | High | Low |
| Unrepeated | 605 (179) | 627 (186) | 609 (141) | 593 (166) |
| 5 times | 629 (181) | 603 (122) | 614 (141) | 599 (160) |
| 10 times | 605 (175) | 674 (171) | 620 (176) | 634 (192) |
| Native speakers | 490  (98) | 507 (137) | 490  (98) | 507 (137) |

*Note.* The value in parentheses is the standard deviation (SD).
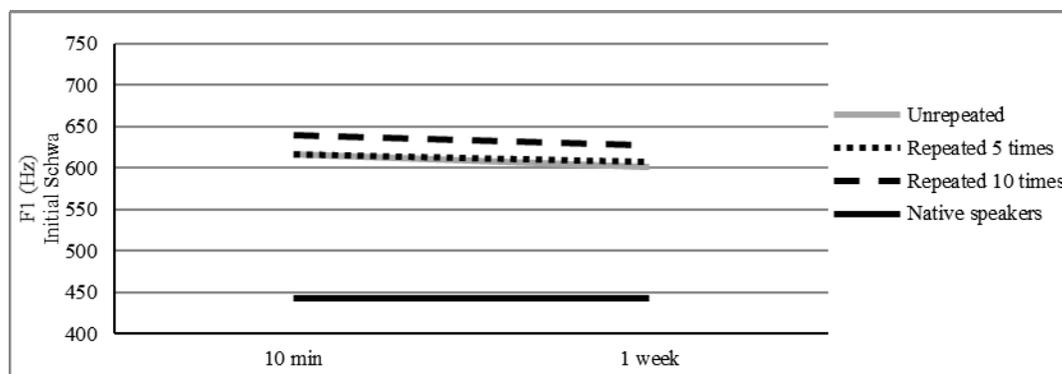


*Figure 3.* Mean F1 values of initial schwas: ten minutes vs. one week.

**Final schwa (strong-weak syllables)**

   Table 6 displays descriptive statistics for the F1 values of final schwa for Test 1 (ten minutes) and Test 2 (one week later). Figure 4 illustrates the values in graph form. The data on the F1 of the final schwa in Tests 1 and 2 were treated in the same manner as in the analysis of the initial schwa. This analysis yielded no significant main effects of Repetition, $F$ (2, 271) = 0.528, $p$ = .591, $\eta^2_p$ = .004, Familiarity, $F$ (1, 272) = 0.129, $p$ = .720, $\eta^2_p$ = .000, or Interval, $F$ (1, 272) = 0.540, $p$ = .816, $\eta^2_p$ = .001. No significant interactions were found for Repetition × Interval, $F$ (2, 271) = 0.151, $p$ = .860, $\eta^2_p$ = .001, for Repetition × Familiarity, $F$ (2, 271) = 0.656, $p$ = .520, $\eta^2_p$ = .005, or for Repetition × Familiarity × Interval $F$ (2, 271) = 0.128, $p$ = .880, $\eta^2_p$ = .001.

   These findings reveal that no improvement in the F1 value of the final schwa occurred in Tests 1 and 2, indicating that immediate repetition (five or ten times) with auditory words did not facilitate improvement in final schwa pronunciation. As with the initial schwa, the F1 values of the final schwa produced by the Japanese participants tended to be higher than the native norm.

TABLE 6
*Mean F1 of Final Schwa (Hz) for Unrepeated and Repeated Words (Five, Ten times) in Test 1 (Ten Minutes) and Test 2 (One Week)*

|  | Test 1 (ten minutes) | | Test 2 (one week) | |
| --- | --- | --- | --- | --- |
|  | Word Familiarity | | | |
|  | High | Low | High | Low |
| Unrepeated | 654 (254) | 640 (173) | 672 (154) | 633 (142) |
| 5 times | 674 (171) | 701 (167) | 694 (162) | 677 (137) |
| 10 times | 666 (170) | 675 (170) | 653 (124) | 692 (155) |
| Native speakers | 577 (144) | 586 (128) | 577 (144) | 586 (128) |

*Note.* The value in parentheses is the standard deviation (SD).
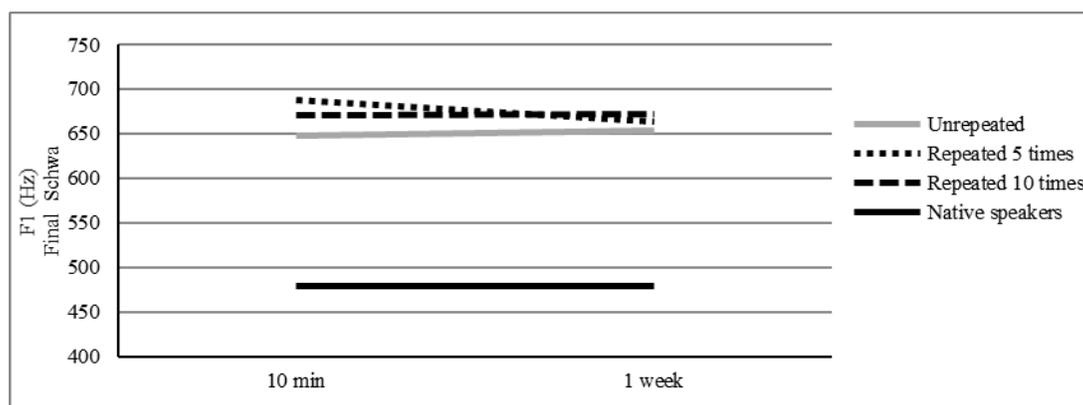


*Figure 4.* Mean F1 values of final schwas: ten minutes vs. one week.

## Quality: F2 Values

### Initial schwa (weak-strong syllables)

   Table 7 displays descriptive statistics for the F2 values of initial schwa for Test 1 (ten minutes) and Test 2 (one week). Figure 5 illustrates the values in graph form. The analysis of F2 values of the initial schwa in Tests 1 and 2 was conducted in the same manner as for F1 values. The result yielded a significant main effect

only for Interval, $F(1, 318) = 3.959$, $p < .05$, $\eta^2_p = .013$. No significant main effects were found for Repetition, $F(2, 317) = 0.157$, $p = .855$, $\eta^2_p = .001$ or Familiarity, $F(1, 318) = 3.312$, $p = .070$, $\eta^2_p = .011$. No significant interactions were obtained for Repetition × Interval, $F(2, 317) = 1.570$, $p = .210$, $\eta^2_p = .010$, for Repetition × Familiarity, $F(2, 317) = 0.583$, $p = .559$, $\eta^2_p = .004$, or for Repetition × Familiarity × Interval, $F(2, 317) = 0.300$, $p = .971$, $\eta^2_p = .000$.

The main effect of Interval suggests that the mean F2 value of the initial schwa was significantly lower in Test 2 than in Test 1, showing that performance worsened one week after the study phase. Additionally, the absence of a significant difference between the repeated and unrepeated conditions suggests that a repetition effect did not occur. The F2 value of the initial schwa produced by the participants was lower than the native norm. Taken together, these findings indicate that the F2 value of the initial schwa did not improve through repetition.

TABLE 7
*Mean F2 of Initial Schwa (Hz) for Unrepeated and Repeated Words (Five, Ten Times) in Test 1 (Ten Minutes) and Test 2 (One Week)*

|  | Test 1 (ten minutes) | | Test 2 (one week) | |
|---|---|---|---|---|
|  | Word Familiarity | | | |
|  | High | Low | High | Low |
| Unrepeated | 1680 (444) | 1610 (408) | 1520 (278) | 1384 (425) |
| 5 times | 1590 (307) | 1600 (307) | 1570 (282) | 1539 (481) |
| 10 times | 1590 (307) | 1500 (348) | 1597 (437) | 1455 (374) |
| Native speakers | 1828 (475) | 1705 (313) | 1828 (475) | 1705 (313) |

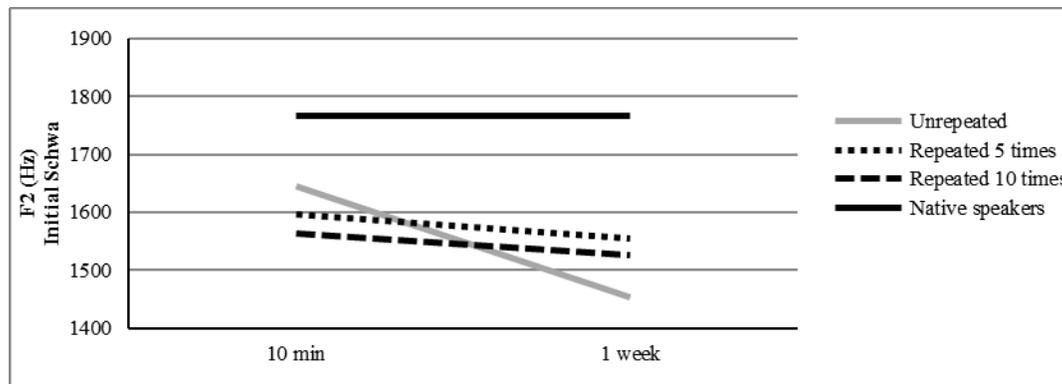*Note.* The value in parentheses is the standard deviation (SD).



*Figure 5*. Mean F2 values of initial schwas: ten minutes vs. one week.

### Final schwa (strong-weak syllables)

Table 8 shows descriptive statistics for the F2 values of the final schwa in Tests 1 and 2. Figure 6 plots the values. The analysis was identical to that performed for the F2 initial schwa. The analysis yielded significant main effects of Repetition, $F(2, 271) = 4.466$, $p < .05$, $\eta^2_p = 0.033$, and Familiarity, $F(1, 272) = 4.128$, $F(1, 272) = 4.128$, $p < .05$, $\eta^2_p = .016$, but no significant main effect of Interval, $F(1, 272) = 0.498$, $p = .481$, $\eta^2_p = .002$. There were no significant interactions for Repetition × Interval, $F(2, 271) = 1.030$, $p = .359$, $\eta^2_p = .008$, for Repetition × Familiarity, $F(2, 271) = 0.879$, $p = .416$, $\eta^2_p = .007$, or for Repetition × Familiarity × Interval, $F(2, 271) = 0.652$, $p = .522$, $\eta^2_p = .005$.

A post-hoc test showed significant differences in the F2 value of the final schwa only between the no-

repetition and five-repetition conditions, $p < .05$, $d = 0.40$ (a small-sized effect). The findings on the main effect of Repetition suggest that the overall F2 values of the final schwa under five repetitions were significantly higher than those under no repetition, and close to the native norm. The absence of a significant effect of Interval suggests that no difference occurred between the ten-minute and one-week intervals. These results indicate that the repetition effect that occurred in Test 1 (ten minutes) lasted for one week under the five-repetition condition in terms of the F2 values of the final schwa.

TABLE 8
*Mean F2 of Final Schwa (Hz) for Unrepeated and Repeated Words (Five, Ten Times) in Test 1 (Ten Minutes) and Test 2 (One Week)*

|  | Test 1 (ten minutes) | | Test 2 (one week) | |
| --- | --- | --- | --- | --- |
|  | Word Familiarity | | | |
|  | High | Low | High | Low |
| Unrepeated | 1731 (315) | 1570 (302) | 1765 (303) | 1611 (304) |
| 5 times | 1894 (342) | 1759 (257) | 1796 (367) | 1803 (368) |
| 10 times | 1800 (233) | 1822 (437) | 1759 (341) | 1660 (280) |
| Native speakers | 1807 (201) | 1909 (286) | 1807 (201) | 1909 (286) |

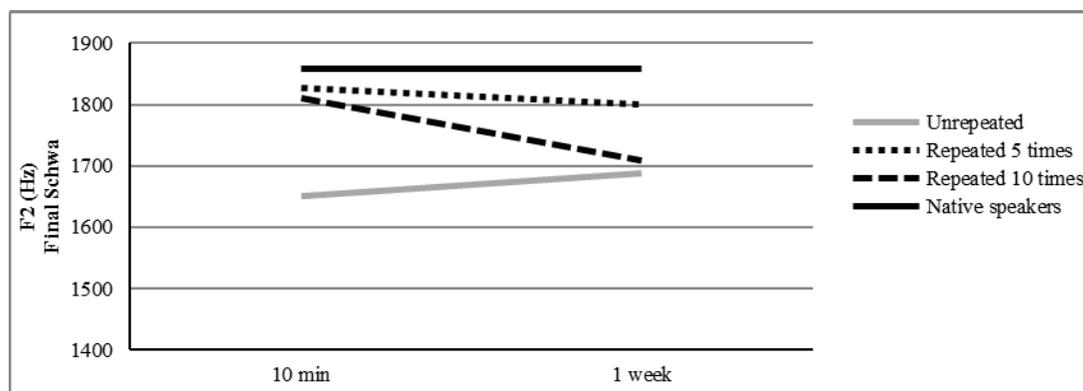*Note.* The value in parentheses is the standard deviation (SD).



*Figure 6.* Mean F2 values of final schwas: ten minutes vs. one week.

## Discussion on Schwa Quality

### Overall findings

In terms of the quality aspect of schwa, overall repetition effects were not observed (except for F2 of final schwa). In general, the participants' F1 was relatively higher than the native norm, and their F2 was lower than the norm. These were characteristic of articulatory targets closer to the Japanese low vowel /a/ or /o/, as previously stated. These findings suggest that because schwa is spelled as <a>, <o>, and <ou> in the word stimuli, the participants may have pronounced it by reflecting the one-to-one grapheme–phoneme correspondence in their L1 phonological representations, which is consistent with previous studies (e.g., Lee et al., 2006; Sugiura, 2006). However, it might not be immediately clear why orthography would affect pronunciation in a test where the spelling is not provided. Young-Scholten and Archibald (2000) argued that adult learners in a foreign language learning context are likely to have had contact with written L2 words from the beginning of learning and are thus expected to possess orthographic representations for words that they have already stored in their lexicon. Given this situation, such learners tend to access written input and

rely on these orthographic representations to pronounce words or to speak in a target language, even when written input is unavailable. This study indicates that the mere repetition of auditory words is only minimally effective in improving schwa quality, suggesting that participants were unable to sufficiently encode the phonetic information available in the auditory input and to restructure the preexisting phonological representation.

### The persistence of the repetition effect for final schwa F2 values

An interesting finding regarding the repetition effect on the quality of schwa is that contrary to duration ratio, a persistent repetition effect was found for final schwa (strong–weak stress pattern) under the five-repetition condition in F2 values. Explaining the finding is somewhat challenging because this result contradicts the persistent effect found only with the initial schwa for duration ratio. However, as argued by Flemming and Jonson (2007), the final schwa is intrinsically more similar in quality to a full vowel than the initial schwa. Therefore, it can be assumed that even the Japanese participants, whose language does not have a central vowel, were able to easily articulate it. An issue that remains unclear; however, is why the repetition effect was observed only in F2 and not in F1. One of the possible accounts lies in individual variations in the accurate productions of schwa by the participants. Tomita et al. (2010) who investigated the pronunciation of schwa by Japanese learners with sufficient English ability has reported that one learner produced a target-like schwa in terms of F1 but not in terms of F2, while another participant exhibited the opposite pattern. Further research is necessary to fully explain this finding.

## Conclusion

## Summary of Findings and Pedagogical Implications

By employing the concept of the auditory priming effect, which is widely studied in L1 and L2 speech processing, the current study suggests the importance of reconsidering the value of the repetition of auditory words for L2 pronunciation improvement. The present study particularly focused on how the amount of input (five and ten repetitions) and the quality of input (initial/final schwa, high/low familiarity) influences a learner's sensitivity to the phonetic/phonological information of presented auditory words as well as how these factors facilitate pronunciation learning. The experiment also evaluated the repetition effect, which was observed immediately after practice and was preserved for one week. Thus, this study provided several noteworthy findings, which are discussed in detail.

Immediate auditory word repetition (five and ten times) improved the pronunciation of schwa by Japanese learners of English, particularly in terms of relative duration; however, repeating the target word more than five times did not differentiate the effect. It is unlikely that there is a linear relationship between the number of repetitions and pronunciation. This finding supports previous research on the optimum number of repetitions for improving prosody in terms of pitch range (Hori, 2008; Miyake, 2009b). From a pedagogical viewpoint, determining the optimal number of auditory word repetitions is meaningful in the pursuit of appropriate pronunciation instruction and learning the schwa sound; however, five repetitions might induce boredom. Therefore, researchers and instructors should develop ways to maintain learner motivation to continue practicing in communicative settings.

Interestingly, the schwa position (word stress patterns) appeared to be related to the persistence of the repetition effect. For the duration ratio, the effect was observed in both initial and final schwas after a 10-minute interval; however, one week later, repetition effectiveness was observed only for the initial schwa, which indicates an exceptional stress pattern for Japanese speakers and shows that it is more difficult for them

to acquire than the schwa sound in the final syllable of a word. This finding suggests that a stress pattern that is uncommon to L2 learners may be more strongly embedded and remembered than a common stress pattern learned by auditory word repetition, which is meaningful from a teaching perspective.

The most intriguing finding of this study was that in the durational aspect the generalization of repetition effects for novel words (i.e., unrepeated words) with low familiarity and the initial schwa sound, which are less common features for the participants, occurred one week after the study phase. This finding suggests that the information obtained from auditory input and repetition is gradually included in the language memory and processing system. ESL and EFL teachers should note that the repetitive effect might not appear for new items immediately after practice. Although this finding is limited to English stress rhythms in the present study, if the repetitive effect can be generalized to novel (unrepeated) words in other linguistic aspects, then this is a crucial finding for L2 language learning.

Regarding vowel quality (i.e., F1 and F2), it appeared that mere repetition did not significantly contribute to improving the vowel quality of the schwa sound. Only the F2 value of the final schwa, which was a common stress pattern for Japanese participants, exhibited a persistent effect (one week); however, for the duration ratio, a persistent effect was observed in the initial schwa sound, which was a less common stress pattern for the participants. These asymmetric results may be attributed to the process in which information about phonetic features is stored in the mental lexicon and/or encoded in speech production; the duration ratio is involved in metrical structures accompanying regular and irregular patterns (rules), whereas vowel quality is related to phoneme information and/or movements of articulators, which do not have these patterns.

Notably, the quality of the schwa sound produced by the participants tended to reflect their L1 one-to-one grapheme–phoneme correspondence in the present study, even though orthography was not presented in the practice or elicitation tasks. That is, the participants were less sensitive to formants and thus failed to improve their pronunciation in this respect. Based on the phonetic feature hypothesis proposed by McAllister et al. (2002), which is mentioned earlier in this paper, since Japanese contains phonemic vowel distinctions regarding length (i.e., long and short), Japanese speakers could have been able to utilize the phonetic feature of duration to produce the reduced English vowels in the manner of native English speakers but may not have been able to use quality information (Lee et al., 2006).

The method of pronunciation training used in the present study was simple repetition, or implicit learning. Therefore, learners were not explicitly taught in regards to strong and weak vowels in words. To facilitate improvement of the quality of the schwa sound, it may be important to include explicit educational intervention in L2 pronunciation training. One method would be to direct the learners' attention to the coarticulation strategy and to emphasize it during their exposure to spoken English input. The term coarticulation was proposed by Menzerath and Lacerda (1933) to refer to successive sounds articulated simultaneously, which is language-specific. Native speakers of English produce full vowels at the targeted position and leave the unstressed vowels, including schwa (only in the F2 value, Kondo, 2000), to contextual assimilation. That is, the English schwa sound is coarticulated more strongly with the surrounding segmental sounds than full vowels (Magen, 1998). Thus, the vowel contrast between full and unstressed vowels plays a crucial role in the coarticulatory pattern of English; however, a study has shown that the F2 values of non-proficient Japanese speakers are quite similar to Japanese vowels (Kondo, 2000) without allowing schwa to be assimilated to the neighboring segments. Japanese speakers of English should therefore learn an English coarticulation pattern by contrasting targeted full vowels and targetless (assimilation to surrounding segments) schwa in their pronunciation. A teaching method based on coarticulation could draw the participants' attention to the visual information necessary for articulating schwa by showing internal articulatory movements, for example, high-low (F1) and front-back (F2) movements of the tongue on the computer screen, to improve the quality of schwa. Future studies are necessary to investigate the effectiveness of this approach.

## Limitations and Future Studies

There are several limitations that call for further studies. Firstly, while the present study examined acoustic properties to investigate the improvement of pronunciation, it would also be informative to take speech intelligibility into account, considering the optimal goal of speech communication (Saito & Lyster, 2012). Secondly, the present study was limited to students at the low intermediate level (CEFR A2 level). Studies with learners at different levels will provide a better understanding of the repetition effect on improving L2 pronunciation of schwa. Moreover, further research is necessary to consider what specific linguistic features (e.g., prosodic information, rhythmic patterns) and under what conditions (e.g., speech rate, quality of voice) the phonetic information of auditory input are incorporated into phonological representations during auditory word repetition tasks. Such studies of pronunciation from the perspective of speech processing will provide further understanding of L2 phonological development, as well as more effective methods of teaching L2 pronunciation.

## The Author

*Kaori SUGIURA* has an M.A. in TESOL from New York University in the United States and a Ph.D. in Language, Communication, and Culture from Kwansei Gakuin University in Japan. She is currently an associate professor in the Department of Science and Engineering of Ritsumeikan University. Her research interests include L2 speech perception and production.

College of Science and Engineering
Ritsumeikan University
1-1-1 Noji-higashi, Kusatsu City, Shiga,
525-8577, Japan
Tel: +81 77-599-4352
Email: k-sugi51@fc.ritsumei.ac.jp

## References

Akita, M. (2001). *The phonological development of adult Japanese learners of English: A longitudinal study of perception and production.* (Doctoral dissertation). University of Durham.

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language*

*Learning, 42*(4), 529-555.

Beckman, M. E. (1986). *Stress and non-stress accent.* Holland, Dorrecht, The Netherlands: Foris Publicati.

Church, B. A., & Fisher, C. (1998). Long-term auditory word priming 270 in preschoolers: Implicit memory support for language acquisition. *Journal of Memory and Language, 39*, 523-542.

Church, B. A., & Schacter, D. L. (1994). Perceptual specificity of auditory priming: implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 521-533.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech & Language*, *2*(3), 133-142.

Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: neural and behavioural evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1536), 3773-3800.

Derwing, T. (2003). What do ESL students say about their accents? *Canadian Modern Language Review*, *59*(4), 547-567.

Dumay, N., & Gaskell, M. G. (2012). Overnight lexical consolidation revealed by speech segmentation. *Cognition, 123*(1), 119-132.

Flemming, E., & Johnson, S. (2007). Rosa's roses: Reduced vowels in American English. *Journal of the International Phonetic Association, 37*, 83-96.

Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, *13*(3), 361-377.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1166-1183.

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in Laboratory Phonology, 7*, 515-546.

Hartsuiker, R. J., & Westenberg, C. (2000). Word order priming in written and spoken sentence production. *Cognition, 75*, B27-B39.

Henderson, L., Powell, A., Gaskell, M.G., & Norbury, C. (2014). Learning and consolidation of new spoken words in autism spectrum disorder. *Developmental Science,17,* 858-871.

Hori, T. (2008). *Exploring shadowing as a method of English pronunciation training.* (Unpublished doctoral dissertation). Kwansei Gakuin University.

Hori, T., & Sugiura, K. (2013). The auditory priming effect in Japanese learners of English: effects of voice specificity and word stress patterns. *The Bialystok Linguistic Archives [Bialostockie Archiwum Jezykowe], 13*, 65-79.

Imaishi, M. (1997). *Nihongo onsei no jikken tekikenkyu*. Osaka: Izumishoin.

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford University Press.

Karni, A., Tanne, D., Rubenstein, B. S., Askenasy, J. J., & Sagi, D. (1994). Dependence on REM sleep of overnight improvement of a perceptual skill. *Science, 265*(5172), 679-682.

Kashiwagi, A., & Snyder, M. (2014). Intelligibility of Japanese college freshmen as listened to by native and nonnative listeners. *JACET Journal, 58*, 39-56.

Kondo, Y. (2000). Production of schwa by Japanese speakers of English. In M. B. Broe & J. B. Pierrehumbert (Eds.), *Laboratory phonology V: Acquisition and the lexicon* (pp. 29-39). Cambridge: Cambridge University Press.

Ladefoged, P. (1993). *A course in phonetics* (3rd ed.). New York: Harcourt Brace Jovanovich.

Lee, B., Guion, S. G., & Harada, T. (2006). Acoustic analysis of the production of unstressed English vowels by early and late Korean and Japanese bilinguals. *Studies in Second Language Acquisition, 28*, 487-513.

Luka, B. J., & Barsalou, L. W. (2005). Structural facilitation: Mere exposure effects for grammatical

acceptability as evidence for syntactic priming in comprehension. *Journal of Memory & Language, 52*(3), 436-459.

Luka, B. J., & Choi, H. (2012). Dynamic grammar in adults: Incidental learning of natural syntactic structures extends over 48 hours. *Journal of Memory and Language, 66*, 345-360.

Magen, H. S. (1998). The perception of foreign-accented speech. *Journal of Phonetics,* 381-400.

Mazzoni, D., & Dannenberg, R. (2010). *Audacity (1.3.12)* [Computer software]. Available from http://audacity.sourceforge.net/

McAllister, R., Flege, J., & Piske, T. (2002). The influence of the L1 on the acquisition of Swedish vowel quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics, 30*, 229-258.

McDonough, K., & Trofimovich, P. (2009). *Using priming methods in second language research*. New York: Routledge.

McGaugh, J. L. (2000). Memory--a century of consolidation. *Science, 287*(5451), 248-251.

Menzerath, P., & Lacerda, A. (1933). *Koartikulation, steuerung und lautabgrenzung: Eine experiemntelle untersuchung.* Berlin and Bonn: Gerd. Diimmler.

Miyake, S. (2009b). Cognitive processes in phrase shadowing: Focusing on articulation rate and shadowing latency, *JACET Journal, 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.

Moyer, A. (1999). Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition*, *21*(1), 81-108.

Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *49*(s1), 285-310.

Murao M., & Sasaki, Y. (2008). Effects of auditory repetition on identification of spoken words, *The Science and Engineering Review of Doshisha University, 49*(1), 67-75.

*Natural Reader 10.0* [Computer software] (2011). Available from http://naturalreader.software.informer.com/10.0/

Ofuka, E., & Gilbert, J. E. (2013). Effects of tasks and awareness on oral imitation accuracy by Japanese EFL learners. *The Journal of Asia TEFL*, *10*(3), 1-29.

Onishi, K. H., Chambers, K. E., & Fisher, C. (2002). Learning phonotactic constraints from brief auditory experience. *Cognition*, *83*(1), B13-B23.

Pennington, M. C., & Richards, J. C. (1986). Pronunciation revisited. *TESOL Quarterly*, *20*(2), 207-225.

Reitter, D., Keller, F., & Moore, J. D. (2011). A computational cognitive model of syntactic priming. *Cognitive Science*, *35*(4), 587-637.

Rescorla, R.A., & Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A.H. Black & W.F. Prokasy (Eds.), *Classical conditioning II: Current theory and research* (pp. 64-99). New York: Appleton-Century-Crofts.

Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /r/ by Japanese learners of English. *Language Learning, 62*(2), 595-633.

Scales, J., Wennerstrom, A., Richard, D., & Wu, S. H. (2006). Language learners' perceptions of accent. *TESOL Quarterly, 40*, 715-738.

Scheepers, C. (2003). Syntactic priming of relative clause attachments: persistence of structural configuration in sentence production. *Cognition, 89*, 179-205.

Suenobu, M., Kanzaki, K., & Yamane, S. (1992). An experimental study of intelligibility of Japanese English. *IRAL: International Review of Applied Linguistics in Language Teaching*, *30*(2), 146-153.

Sugiura, K. (2006). On the difference in schwa produced by native speakers of English and Japanese speakers of English. *JASEC Bulletin, 15*, 1-12.

Sugiura, K. (2015). *Production of English schwa by Japanese speakers.*(Unpublished doctoral dissertation). Kwansei Gakuin University.

Sugiura, K., & Hori, T. (2012). Auditory priming effect in Japanese learners of English: An investigation using a repetition task of spoken words*, Kanto-Koshinetsu Association of Teachers of English Bulletin*, *26,* 39-51.

*SuperLab Pro (4.0)* [Computer software] (2006). San Pedro, CA: Cedrus Corporation.

*The British National Corpus, version 3* (BNC XML Edition) (2007). Available from http://www.natcorp.ox.ac.uk/

Tien, D. A. N. G. (2013). Impact of rhythm on Vietnamese adult EFL learners intelligibility in term of mid-level tone. *International Journal of Applied Linguistics and English Literature*, *2*(4), 98-109.

Tomita, K., Yamada, J., & Takatsuka, S. (2010). English vowel spaces produced by Japanese speakers: The smaller point vowels and the greater schwas. *Journal of Psycholinguistic Research, 39*, 375-391.

Trofimovich, P. (2005). Spoken-word processing in a native and a second language: An investigation of auditory word priming. *Applied Psycholinguistics*, *26*, 479-504.

Trofimovich, P. (2008). What do second language listeners know about spoken words? Effects of experience and attention in spoken word processing. *Journal of Psycholinguistic Research*, *37*, 309-329.

Trofimovich, P., & Gatbonton, E. (2006). Repetition and focus on form in processing L2 Spanish words: Implications for pronunciation instruction. *The Modern Language Journal, 90*(4), 519-535.

Truscott, J. (2014). *Consciousness and second language learning* (Vol. 83). Bristol: Multilingual Matters.

Ueyama, M. (2000). *Prosodic transfer: An acoustic study of L2 English vs. L2 Japanese*. (Unpublished doctoral dissertation). University of California, Los Angeles.

Vance, T. J. (1987). *An introduction to Japanese phonology*. Albany, NY: State University of New York Press.

Walker, M. P., & Stickgold, R. (2006) Sleep, memory, and plasticity. *The Annual Review of Psychology, 57*, 139-166.

Wallace, K. (1994). *An acoustic study of American English Schwa in multiple speaking modes*. (Unpublished doctoral dissertation). New York University.

*WaveSurfer 1.8.5*. [Computer software] (2011). Available from http://wavesurfer.en.uptodown.com/windows

Yokokawa, H. (Ed.). (2009). *Database for second language pedagogy and research: English vocabulary familiarity of Japanese EFL learners: auditory version*. Tokyo: Kuroshio Publishers.

Young-Scholten, M., & Archibald, J. (2000). Second language syllable structure. In J. Archibald (Ed.), *Language linguistics and linguistic theory* (pp.64-101). Oxford: Blackwell.