Exploring Differences Between Shadowing and Repeating Practices: An Analysis of Reproduction Rate and Types of Reproduced Words

Osato SHIKI
Kwansei Gakuin University
Yoko MORI
Doshisha University
Shuhei KADOTA
Kwansei Gakuin University
Shinsuke YOSHIDA
Kansai University

Abstract

This study compares the effects of shadowing and those of repeating during six repetition trials in terms of two aspects: reproduction rate and type of reproduced words. 48 university students performed both shadowing and repeating six times each, and their recorded voices were analyzed to see if there were any differences between the two kinds of training methods in terms of the reproduction rate and the number of function and content words reproduced correctly on each repetition trial. The results indicated that four or five repetition trials were enough to improve the rate of reproduction to a level near the ceiling point. It was also found that the participants could reproduce more content words with shadowing than with repeating, whereas some reduced function words were reproduced from the first trail of repeating, all of which may reflect differences in the language processing between shadowing and repeating performances.

1. Introduction

In the discussion of effectiveness of shadowing and oral reading practices, Kadota (2007) hypothesizes two possible effects of shadowing training in L2, particularly in EFL learning:
(1) It facilitates automatic perception of input speech, the lower-level process which precedes comprehension of the message, and the improvement of listening skill.
(2) It enhances vocal rehearsal rate of a speech input and promotes internalization (or
memorization) of words, formulaic chunks, etc. (see Kadota, 2007 for a discussion also of the effects of oral reading practice in ESL / EFL)

Kadota (2007) also suggests that shadowing (SH) and repeating (RE), both of which are basically the practice of repeating sounds, are cognitively different from the psycholinguistic viewpoint. SH is assumed to be the on-line immediate process of repeating speech, whereas RE training provides students with silent pauses in which they are supposed to reproduce the sounds off-line. This difference of articulation latency seems to be very important. We can assume that the former task forces students to focus more on the reproduction of the input speech, while the latter is supposed to give the learner enough time to analyze the input message syntactically and semantically, such as parsing and semantic proposition constructions. The difference may lead us to the hypothesis that the two methods should differ in their reproduction rates and types of words correctly reproduced.

Two studies hint at how many repetitions of SH are needed to significantly change its performance in terms of the number of correctly reproduced words and speech rates. Hori (2008) reported that reproduction rates sharply increased in the first five repetitions of SH (p. 96). Miyake (2009) found that speech rates significantly improved for some phrases through six repetitions of SH. However, it has not been examined in detail how much the reproduction rate increases by each additional repetition from the 1st through 6th trials. Moreover, it can be assumed that RE may differ in the improvement of reproduction rate because of the cognitive aspect specific to RE.

The present paper, thus, attempts to investigate the cognitive processes involved in shadowing and repeating English auditory inputs by EFL learners in Japan based on the analyses of their reproduction rates and types of words reproduced. The present study also aims to provide practical insight about the efficiency of the two methods in English classrooms in Japan. To be more specific, it examines how many trials are needed for
Japanese learners to make sufficient progress in SH vs. RE practice.

To explore the processing difference between SH and RE, this study, thus, addresses the following research questions:
1) Are there any differences in the improvement of reproduction rate between SH and RE over six repetition trials?
2) Are there any differences in the type of words correctly reproduced by the participants between SH and RE?

2. Method

2.1 Participants

The participants in the present study were 48 Japanese university students learning English as a foreign language, 24 being given SH first and then RE, and the other 24 being given a reverse treatment of RE followed by SH. Their English proficiency fell between low intermediate and intermediate levels.

2.2 Materials

Model speech for SH and RE trainings was a recorded reading of a short story taken from Shukan ST published by the Japan Times, lasting about 1 minute and 42 seconds in length and containing a total of 220 words, with a speaking rate of about 135 words per minute. The story consisted of basic English words with a mean level of 1.02 in JACET 8000, and was assumed to be easy enough for the participants to repeat or shadow without viewing the text. For the two types of speaking trainings, the story was divided into two parts, which are referred to as Part 1 and Part 2 in this paper. For SH practice, the original materials were used six times in succession. For RE practice, pauses were inserted into 18 boundaries in Part 1, and 19 in Part 2. Most of the boundaries were consistent with major syntactic boundaries marked by pausing in the original material. Pause length was 1.5 times the utterances they preceded and lasted 945 to 3640 ms with a mean duration of 2131.7 ms for Part 1, and 2017.8 ms for Part 2.

2.3 Experiment Procedure

The experiment was conducted in a CALL Room with 48 Windows notebook PCs and with software specifically designed for recording SH. The procedure was as follows:
1) Participants listened to Part 1 of the material once through a headphone.
2) They then were required to either shadow Part 1 of the material or repeat it a total of six times by using a microphone attached to the headphone.
3) After the recordings were over, the experimenter collected the recorded data in the PC from the participants.
4) Then the participants proceeded to Part 2 of the material and recorded the reproductions in the same way as in Part 1.

2.4 Reproduction Analyses

To what extent the participants were able to repeat the model stimuli was analyzed by the four authors. Each author analyzed four different trainings: SH of Part 1, SH of Part 2, RE of Part 1, and RE of Part 2. Each of us listened to the recordings several times, checked syllables that were not correctly reproduced, and calculated a reproduction rate for each trial by each participant. It was found that in some of the recordings, the first and last sentences in Part 1 and Part 2 were totally missing presumably due to a delayed start or some recording problems. Thus, the first and last sentences in Part 1 and Part 2 were excluded from reproduction analyses.

For the consistency of rating criteria, we listened to several samples and discussed the criteria for incorrect reproduction (errors) before we actually started rating. The criteria used are listed in (1) through (4) below:

(1) Insertion of additional words or phrases should be neglected.
(2) Order reversal between two words should be counted as one error.
(3) In principal, errors are checked syllable by syllable.
    e.g. “Roommate” for “teammate” is counted as one error.
(4) Substitution by totally different words should be regarded as errors.
    e.g. “Wonderful” for “one of the” is counted as 3 errors.

In order to confirm the consistency of our use of the rating criteria, each of us checked the reproduction errors of each participant twice. In order to further exclude inter-rater variation, one of the authors checked one of the six trials by all the participants (96 trials in total) and confirmed the consistency of the rating criteria among the four raters.

3. Results & Discussion

3.1 Reproduction Rate

In order to carry out a statistical analysis to examine the difference between the effects of SH and those of RE on reproduction rate, the conditions (two groups of participants, and two different parts of the story) were counterbalanced to reduce the effects of those differences; thus, the obtained data were divided by the method factor (SH and RE) to have each of them include the data of both groups of participants (one group who recorded SH first and the other group who recorded RE first) and both parts of the story.
(Part 1 and Part 2). For the analysis of the reproduction rate, a three-way ANOVA with Method (SH vs. RE) and Session (Part 1, Part 2) as between-subject factors, and Trials (1st to 6th) as a within-subject factor was conducted on the reproduction rate as a dependent variable. Table 1 shows the descriptive statistics for the rates of reproduction over six repetition trials. The ANOVA revealed that there was a significant main effect for Trial \( F(5, 460) = 37.941, p < .001 \). There were significant interactions between Method and Trial \( F(5, 460) = 4.323, p < .001 \), and between Session and Trial \( F(5, 460) = 5.575, p < .001 \).

Table 1. Descriptive statistics for reproduction rates (%)

<table>
<thead>
<tr>
<th></th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>5th</th>
<th>6th</th>
<th>7th</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>SH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Part 1 (24)</td>
<td>74.7</td>
<td>13.5</td>
<td>76.4</td>
<td>13.4</td>
<td>76.8</td>
<td>14.1</td>
<td>79.5</td>
</tr>
<tr>
<td>Part 2 (24)</td>
<td>72.7</td>
<td>11.4</td>
<td>75.3</td>
<td>13.1</td>
<td>76.9</td>
<td>12.0</td>
<td>77.3</td>
</tr>
<tr>
<td>RE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Part 1 (24)</td>
<td>62.9</td>
<td>15.9</td>
<td>66.2</td>
<td>15.4</td>
<td>69.3</td>
<td>13.8</td>
<td>74.3</td>
</tr>
<tr>
<td>Part 2 (24)</td>
<td>72.8</td>
<td>12.0</td>
<td>79.9</td>
<td>9.4</td>
<td>80.3</td>
<td>10.1</td>
<td>81.1</td>
</tr>
</tbody>
</table>

SH: shadowing RE: repeating

As the post hoc analysis for the effects of Method, \( t \)-tests were run to examine the differences in the reproduction rate between SH and RE on each repetition trial. The reproduction rate of SH was higher than that of RE only on the 1st trial \( (p < .05) \). Table 2 shows the results of LSD (least significant difference) multiple comparisons conducted to examine the differences in the rate of reproduction among the six trials for each type of method (SH and RE). In Figure 2, the vertical lines indicate standard errors.

Table 2. LSD multiple comparisons for SH & RE

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1&lt;2,3,4,5,6</td>
<td>( p &lt; .001 )</td>
<td>2&lt;4,5,6</td>
<td>( p &lt; .05 )</td>
<td>3&lt;5, 3&lt;6</td>
<td>( p = .09, p &lt; .05 )</td>
<td>1&lt;2,3,4,5,6</td>
</tr>
<tr>
<td>ST</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2&lt;3,2&lt;4,5,6</td>
<td>( p = .07, p &lt; .001 )</td>
<td>4&lt;6</td>
<td>( p &lt; .001 )</td>
<td>5&lt;6</td>
<td>( p = .08 )</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. Reproduction rates of SH & RE over six repetition trials
As Table 2 indicates, putting equal importance on the marginal significances (indicated by underlines), it can be said that significant differences were constantly obtained up to 4th trial, but the effects are sharply reduced after 4th trial. This is further indicated with the result that there was no significant difference between 4th and 5th trials. In Figure 2, it can be easily seen that the rate gains effectively disappear around 4th and 5th trials both for SH and RE.

As the post hoc analysis for the effects of Session, t-tests were also run on each trial. The reproduction rate of Part 2 was higher than that of Part 1 on the 2nd and 3rd trials (p < .05). This means that there was a sudden increase in the reproduction rate from the 1st to the 2nd trial in Part 2. Table 3 shows the results of LSD multiple comparisons conducted to examine the differences in the rate of reproduction among the six trials for each session (Part 1 and Part 2). The underlined part indicates a marginal significance. As Table 3 indicates, significant differences were not detected again for the comparison between the 4th and 5th trials both for Part 1 and Part 2. As mentioned above for the comparison of SH and RE, this tendency again gives an indication of decline in the improvement of the reproduction rate around the 4th or 5th trials. This can be seen easily in Figure 3, especially with the line of Part 1 (the first half session).

<table>
<thead>
<tr>
<th>Table 3: LSD multiple comparisons</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>1 &lt; 2,3,4,5,6</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td>2 &lt; 3,4,5,6</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td>Part 1</td>
<td></td>
</tr>
<tr>
<td>3 &lt; 4,5,6,</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td>4 &lt; 6</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td>5 &lt; 6</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>1 &lt; 2,3,4,5,6</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td>2 &lt; 5,6</td>
<td>p &lt; .05</td>
</tr>
<tr>
<td>3 &lt; 5</td>
<td>p = .08</td>
</tr>
</tbody>
</table>

Figure 3. Reproduction rates for Part 1 & Part 2 over six repetition trials

These results of the statistical analyses indicate that the pattern in the improvement of reproduction rate over the six repetition trials of SH were quite similar to that of RE. Thus, it can be concluded that in terms of the improvement of reproduction rate the data did not show any signs of processing differences between SH and RE performances, which gives a negative answer to research question 1. However, these data imply that practice of 4 or 5 repetitions enables learners to improve reproduction rates successfully to a ceiling point when the model English being shadowed or repeated is easy enough for them to carry...
out SH and RE as in this study. Moreover, Figure 3 shows that in the line of Part 2, the reproduction rate at the 2nd trial quickly reached a level near the ceiling point obtained at the 4th trial of Part 1 after the sharp gain from the 1st trial. The implication of this high reproduction rate at the 2nd trial is that the practice effects gained from the trainings at the first half (Part 1) were sustained to the second half (Part 2) even after the new material was used and the sharp gain between the 1st and 2nd trials (Part 2) could be due to the learners' efforts to get used to repeating the new material. This means that throughout the 12 trials (from Part 1 to Part 2 in total) five or six repetition trials were enough for the participants' performances to approach the ceiling effects in the reproduction rates. This was also supported by the data that the rate at the 6th trial of the second half (Part 2) was almost identical to the 6th trial of the first half (Part 1). These findings could be a useful for the classroom implementation of SH and RE.

3.2 Types of Reproduced Words

The next question to be asked was whether there were any differences in the type of words correctly reproduced by the participants between SH and RE. English words (or words of many languages) can be categorized into two groups: content words and function words. The former serves to carry information or meaning whereas the latter functions as a connector of content words, indicating the relationship between them. This vocabulary categorization should be taken into account when analyzing the effects of SH vs. RE since in connected speech, content words are usually stressed and prominent enough to be easily heard, whereas function words are weak, short, and low in pitch in a neutral context. In addition, the reproduction rates for these two types of words may reflect the language processing that the students were engaged in during SH or RE practice.

Categorization was done only in sentences whose reproduction rates were analyzed. Syllables that belonged to content or function words were counted and categorized into their corresponding groups except for "ed" in "wanted" and "communicated," which was grouped into the function word category in this study. These groups of syllables that constituted content or function words are referred to as Content Group and Function Group, respectively. In Part 1, 99 syllables belonged to Content Group, and 48 to Function Group; In Part 2, 85 syllables to Content Group, and 50 to Function Group.

The 1st and 6th trials were selected for this vocabulary analyses in order to examine the effect of the six repetition trials on the reproduced vocabulary. Figures 4 and 5 show how the mean reproduction rates of Content and Function Groups changed from the 1st to 6th trial through SH or RE training, for Part 1 and Part 2, respectively. The error bar in the figures shows a standard error above or below the mean. In both figures, reproduction rates are higher for Content Group than for Function Group. When a comparison is made among each vocabulary group, reproduction rates for RE in the 1st trial tend to be lower than their
corresponding rates for SH both in Parts 1 and 2. In the 6th trial, however, reproduction rates for RE are slightly higher than those of SH except that for Content Group in Part 1.

Figure 4. Reproduction rates for Part 1

A three-way ANOVA was performed on the mean reproduction rates of all the syllables of Content and Function Groups with Trial (1st vs. 6th) and Method (SH vs. RE) as within-subject factors, and Word Group (Content Group vs. Function Group) as a between-subject factor. The test was done for Part 1 and Part 2, respectively. In Part 1, three main effects were detected [Trial: $F(1, 145) = 154.80, p < .0001$; Method: $F(1,145) = 22.19, p < .0001$; Word Group: $F(1,145) = 13.94, p < .001$]. In addition, there was a significant interaction between Method and Word Group [$F(1,145) = 10.71$]. Post hoc tests showed that for Function Group, reproduction rates did not differ significantly between SH and RE both in the 1st and 6th trials, whereas for Content Group, the reproduction rate of SH was significantly higher than its corresponding rate of RE both in the 1st and 6th trials. In Part 2, only the effects of Word Group and Trial were significant [Word Group: $F(1,133) = 35.43, p < .0001$; Trial: $F(1,133) = 67.00, p < .0001$]. There were significant interactions between Trial and Word Group [$F(1,133) = 4.27, p < .05$] and between Trial and Method [$F(1,133) = 4.25, p < .05$]. Post hoc tests showed no significant difference in reproduction rates between SH and RE when comparison was made separately for Content Group and Function Group in the 1st trial and 6th trial, respectively.

The results showed that a significant effect of method was restricted to the reproduction rates of Content Group in Part 1. This means that in Part 1, SH is a more efficient method than RE in terms of reproducing content words, but not function words. In Part 2, no effect of different training methods was found for either Content or Function Groups.

Figures 4 and 5 make clear that in the 1st trial of Part 1, the four reproduction rates vary most widely. Let us now examine in more detail what kinds of words were reproduced there. In the following text (1), syllables whose reproduction rates were less than 60% in the 1st trial of SH with Part 1 were indicated with a square, and those in RE, with bold
letters.

(1) The 1st trial of Part 1
She did not understand the culture. / Her family and her friends were eleven thousand kilometers away in Tokyo. / Yamaguchi went to Florida State University to play soccer. / There, she joined a team called the Seminoles. / She quickly became a major soccer player. / She has scored 17 goals and she is at the top of Division I. / She also has one of the highest points-for-game in the United States. / "From when I was 15, I wanted to come to the United States, and I wanted to play soccer in college." This is mainly because of Yamaguchi. / She is not the strongest nor the quickest player, but technically she is one of the best. / But going to the United States was hard for Yamaguchi. / Because she could not speak English, she could only smile and nod when she was with her teammates. / She communicated by writing words on paper. / Her teammates also wrote on paper because Yamaguchi was better at reading than speaking. / But Yamaguchi can speak English easily now and her grades are good, too.

"Young Japanese girl a U.S. soccer star" Shukan ST (Nov. 16, 2007) [www.japantimes.co.jp/shukan-st/]

It can be seen from the text (1) that participants failed to reproduce similar words in the two trainings. Still there were some noticeable differences between them. Figures and nouns such as "eleven thousand", "Tokyo", "Florida State University", and "Seminoles" were reproduced by 60% or more of the participants only in SH. It is generally known that listening to numbers and proper nouns are among the most difficult skills to acquire for learners of English as a foreign language. It can be assumed that these kinds of words were successfully reproduced from the very 1st trial exclusively in SH training since SH was a task to repeat incoming sound as soon as possible, usually after a few syllables/words.

It is also clear from (1) that 60% or more students in SH failed to reproduce function words "From", "I", and "was" in "From when I was", and "when" and "she" in "when she was with her teammates", all of which were repeated by 60% or more participants in RE. These unstressed function words are phonetically reduced and usually difficult to hear for Japanese learners. It is likely that the participants engaged in RE produced these function words probably because they thought the words were necessary for the sentences to be grammatically correct. In other words, they added these function words as they reproduced sentences through syntactic and semantic processing.

These findings suggest that the differences in the type of reproduced words between the two training methods reflected differences in their language processing. This suggests that in RE, participants processed input stimuli syntactically and semantically. In SH, by contrast, it may be possible for participants to reproduce incoming sound without syntactic
or semantic processing, though as training sessions proceeded, the participants could be engaged in syntactic and semantic processing of input speech while shadowing it.

4. Conclusion

This study gives both practical and theoretical insights into the implementation of shadowing and repeating practice in classrooms. As for the practical point of view, the data of the reproduction rate imply that it is not effective to continue shadowing or repeating more than 4 or 5 times in terms of the improvement of reproduction rate, and even after new material is introduced after the 4th or 5th trial, the effects of the practices are likely to be weak. This finding will encourage teachers to implement shadowing and repeating in class since the effects of the methods can be obtained without sacrificing a large portion of regular class hours. With the second data (types of reproduced words) details are given to explore the interesting cognitive aspects of shadowing and repeating. Shadowing is an on-line type of processing while repeating is an approach defined as rather top-down and off-line types of processing. The data obtained here supported the theoretical basis of how these two tasks can be used effectively to take full advantage of the characteristics of each task; for example, shadowing is for gaining student sensitivity to phonological aspects of English while repeating is for internalizing learned language items to long-term memory to facilitate acquisition of them. The findings in this study, however, need to be tested again with learners from various proficiency levels and with various types of texts. Further investigations are also necessary to determine to what extent the reproduction rates can predict the effects of shadowing and repeating practices; thus, it should be pointed out with caution that as the focus of shadowing or repeating trainings is placed more on improving subtle aspects of English phonology, the number of trials should be increased more than the number of repetitions (4 or 5 times) supported in this study.

References


