OTOIKU Research

Investigating Improvements in Listening Using Levelled Shadowing Materials

Leslie Chivers, Judith Saburi, Shannon Kiyokawa

Abstract

The improvement of listening comprehension skills is often a challenging task for language teachers and learners. Over the years many methods have been developed to help learners and instructors alike. Recently, in Japan especially, shadowing is generally accepted as an effective and efficient means for developing learners' listening skills. Many studies have reported learners' listening competence advancing after being given short-term shadowing training. However, an exact method for implementing shadowing is still yet to be determined. Therefore, this study was conducted to further explore effective procedures to implement shadowing. Shadowing often uses "easy" material the learner can mimic repeatedly, but a study conducted at Akita University indicated that using a combination of two levels of materials was more effective than just using materials of a similar difficulty level for students of the same intermediate level (Hamada, 2012). This study further explores those results, using a variation on the training materials. The purpose of the study was to determine if alternating two levels of materials showed similar improvement for students of varying levels. The results show that a combination of two different difficulties of materials improved learners of varying levels of proficiency listening comprehension skills.

Introduction

The teaching of listening techniques has often proven difficult for instructors in the English as a foreign language (EFL) field. Over the years many methods have been developed to help learners and instructors alike. Repetition is widely accepted as an effective technique, among others, for language acquisition and is integrated into several different tasks. Other common tasks include parallel reading and translation. Interpreters have trained with a technique called "Shadowing" for decades. This has proved effective for advanced language learners but is also effective for the lower-level learner. Shadowing was introduced into the Japanese English classroom 25 years ago by Tamai (1997). Numerous studies have been conducted to determine optimal shadowing duration, length of materials used, and accompanying tasks in the classroom; however, a precise methodology has yet to be determined.

What is shadowing?

Shadowing is generally accepted as an effective and efficient means for developing learners' listening skills. Shadowing is defined as "a paced, auditory tracking task, where the

language learner will repeat, with a slight delay, the audio" (Tamai, 1992 as cited in Hamada, 2012). This lag between the listening audio and the parroting from the learner is like a person walking behind another like a shadow, hence the coined term "Shadowing". One key point of shadowing is, there is no printed script from which to read while parroting the audio.

Effectiveness

Shadowing triggers different cognitive processes which vary greatly depending on the demand of the tasks. Reading or listening to audio with long pauses for repeating, are off-line tasks, whereas shadowing does not allow for pauses for the learner to repeat. Students must parrot with the audio. This is an online cognitive process (Shiki, Mori, Kadota, & Yoshida, 2010). Shadowing allows the learner to activate different levels of cognitive learning more than parallel reading does. Parallel reading utilizes top-down processing. Shadowing does the opposite. The learner is able to discern and recognize content (Rost, 2011 as cited in Hamada, 2018).

Over the last few decades, shadowing has become a popular method for listening practice in Japan. Many studies have been conducted to test its effectiveness. Although shadowing, as a whole, is considered effective the tasks paired with shadowing vary greatly. Hamada (2012) combined dictation with pre- and post-cloze at the beginning and end of each shadowing session. Shadowing uses material of an appropriate level; the learner can mimic repeatedly. Hamada's hypothesis that two levels of difficulty would show more improvement than one was supported to some extent (Hamada, 2012). This study uses a similar approach. This study posed the question - Will using materials of two levels of difficulty improve learners' listening comprehension skills across different ability levels?

Method

Participants

Sixteen Japanese university 2nd year and three 3rd year (female) students majoring in English participated in this research. One 2nd year student's data was excluded from analysis due to absence from the post-test. The total number of participants for the purpose of analysis was eighteen. In April, students were placed in class groups (Beginner, Intermediate and High) based on their placement in their 1st year Listening & Speaking classes, their GPA and most recent TOEIC test (January, 2022) scores. The cohort of all students scored between 150 and 375 on TOEIC Listening. The Advanced class all had TOEIC Listening scores of 255 or higher (and the two lower classes were more mixed with Listening scores of 280 or lower. The participants, in eight of the semester classes, were trained using easier and more difficult materials alternately. (Day 1, 3, 5, 7: Easy materials; Day 2, 4, 6, 8: Difficult materials).

Materials

This experiment is a partial replication study. Similarly, to Hamada (2012), *The TOEIC test new official version* (2009) was used. As Hamada notes, the use of items constructed using a TOEIC format "maintains reliability" (2012). Additionally, students' familiarity with TOEIC as a measure of English ability has broad acceptance and contributes significantly to the face validity of classroom activities using TOEIC materials aimed at improving listening skills.

Performance on listening exercises and tests is influenced by a number of variables, e.g., executive functioning and vocabulary size (Wallace & Lee, 2020). However, as far as characteristics of listening measures are concerned, text length of materials is a major variable in difficulty level and measuring learners' ability to process and store a passage. Text length data, and the Flesch-Kincaid index, providing readability ease and grade level, were used in preparing training materials. Table 5 shows the odd (Easy) and even (Difficult) days' data and the levels of difficulty that were alternatively used. Based on mean number of words each Easy text was adapted to provide 12 cloze gaps (approximately 20% of the total words) and each Difficult text was adapted to provide 17 cloze gaps (approximately 20% of total words).

Table 5

Means of Odd and Even listening measures

	Odd Days (1, 3, 5, 7)	Even Days (2, 4, 6, 8)
Mean Words	65	94.5
Mean No. of Sentences	11.25	9.5
Mean Words/Sentence	5.825	10
Mean Flesch Readability Ease (FRE)	81.75	68.75
Mean Flesch-Kincaid Grade Level	3.25	6.1
Mean Passives	0	12

To assess improvements in listening skills, pre- and post-tests consisting of the full, 45-minute sample TOEIC Listening Test 4 were administered. The same test was administered on both occasions about 2 months apart. However, only Sections 3 and 4 were used for analysis because these were the sections used in the training materials. Students were given no detailed feedback about their responses after the pre-test. Each section of the TOEIC Listening test is detailed in Table 6.

Table 6

TOEIC (2009) listening test sections (adapted from ETS, 2022)

Section	Task
Part 1 **Photographs:** **6 questions**	Four short statements regarding a photograph will be spoken only one time. The statements will not be printed. Of these four statements, select the one that best describes the photograph and mark your answer on the answer sheet.
Part 2 **Question-Response:** **25 questions**	Three responses to one question or statement will be spoken only one time. They will not be printed. Select the best response for the question and mark your answer on the answer sheet.
Part 3 Conversations: **39 questions**	Conversations between two or three people will be spoken only one time. They will not be printed. Listen to each conversation and read the questions printed in the test book (the questions will also be spoken), select the best response for the question, and mark your answer on the answer sheet. Some questions may require responses related to information found in diagrams, etc. printed on the test book as well as what you heard in the conversations. There are three questions for each conversation.
Part 4 **Talks:** **30 questions**	Short talks such as announcements or narrations will be spoken only one time. They will not be printed. Listen to each talk and read the questions printed in the test book (the questions will also be spoken), select the best response for the question, and mark your answer on the answer sheet. Some questions may require responses related to information found in diagrams, etc. printed on the test book as well as what you heard in the talks. There are three questions for each talk.

Procedure

The shadowing training sessions consisted of eight in total and were conducted mid-course. Two parts of the TOEIC Listening test were used. Part 3 texts (Conversations) were used for Days 1, 3, 5 and 7. Part 4 (Talks) was used for Days 2, 4, 6 and 8. The training steps were based on revised instructions used by Hamada (2012) and are shown in Table 7.

Table 7

Procedure of each training session

Step	Procedure
1. Dictation cloze	Fill in the blanks of the written scripts.
2. Mumbling	Silently shadow the incoming sounds without texts
3. Parallel reading	Shadow while reading the text of the passage.
4. Check understanding	Check with the texts written both in English and Japanese for three minutes.
5. Shadowing	Shadow three times.
6. Check details	Check with the written texts for three minutes for sounds one could not hear or shadow, and meanings one could not understand.
7. Content shadowing	Concentrate on both shadowing and interpreting the meaning of the passage.
8. Dictation cloze	Same as Step 1.
9. Check answers for dictation	Check the answers for steps 1-8.

In understanding these steps, it should be noted that in Step 4 students checked the texts without sound. This was considered important to allow students a period of time in which to focus solely on the written text. This break-up of the shadowing was considered important to avoid the ceiling point in successful reproduction identified by Shiki, Mori, Kadota and Yoshida (2010). As Hamada (2012) noted, Steps 1 and 8 were included in order to provide students with an opportunity to compare their initial and final scores, enhancing chances for "repeated success" (p. 7). These "experiences of mastery" in turn can promote self-efficacy (Bandura, 1994).

The pre- and post-TOEIC tests were conducted immediately prior to and after the 8 training sessions on the same day of the week as the training sessions.

Analysis

To measure student improvements two main bodies of data were analysed. The pre-and post-test scores were analysed. For the purpose of analysis and discussion of results, students were grouped into 3 groups (High (7), Mid (6) and Low (6)) according to their January TOEIC Listening Scores. Scores in Steps 1 and 8 (pre- and post-cloze) of each training session were also analysed.

Results

Pre- and Post TOEIC Tests (Sections 3 and 4)

As can be seen in Table 8, the High group experienced an increase in mean score of 5.57, the Mid group an increase of 3.17 and the Low group an increase of only 1.5. All groups increased their minimum Post-test scores (by 5, 6 and 3 points respectively). Maximum Post-test scores increased for the High and Mid group but decreased by 1 point for the Low group.

Table 8

Statistics of TOEIC Pre- and Post-Listening tests for High, Mid and Low groups

	Mean	SD	Min	Max
Pre-test High group	27	7.61	19	40
Post-test High group	32.57	7.95	24	46
Pre-test Mid group	25.83	6.3	17	34
Post-test Mid group	29	4.28	23	35
Pre-test Low group	24.5	4.84	19	32
Post-test Low group	26	3.68	22	31

Figure 1 shows the pre- and post-cloze increases in correct responses for each group in each of the eight training sessions. There were no decreases in correct responses in any sessions. Increases ranged from 14.71% (Low group, Session 2) to 46.34 % (High group, Session 1). Group average increases indicate similar increases for the High and Mid groups.

However, the Low group shows a significantly lower average.

Figure 1

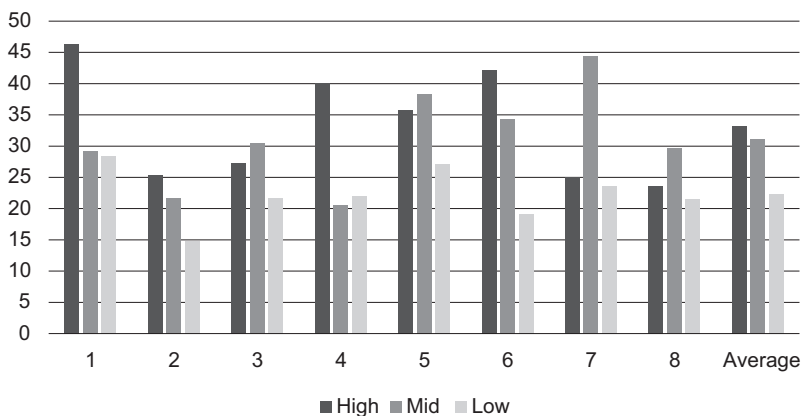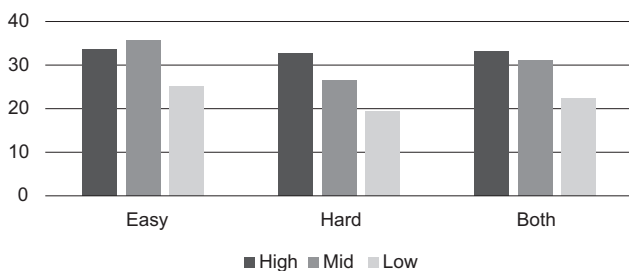*Pre- and Post-Cloze Increases in Correct Response rate (%)*

Figure 2 shows the average increase in correct responses for each group based on type of text (Easy/Hard). Increases for Easy texts ranged from 25.17% (Low) to 35.6 %(Mid) and increases for Hard (Difficult) texts ranged from 19.36% (Low) to 32.73% (High).

Figure 2

Average Increases in Correct Responses Based on Text Difficulty (%)

Discussion

The results of this study show that a combination of two different levels of materials improved the listening comprehension skills of learners of varying degrees of proficiency. The greatest increase can be seen in the highest-level group. The mid and low-level groups showed smaller improvements. Several reasons are examined for these results. First, as Hamada (2012) states in his article, "A combination of two levels can deal with individual differences of listening proficiencies" (p. 8). The students in our study scored between 150 and 375 on TOEIC Listening thus showing a definite range of listening competence. The different levels of materials allowed for a wider range of competences to be addressed. In the study,

all groups showed a decrease in correct responses on Hard texts compared to Easy texts, but the High group showed the slightest decreases while Mid and Low groups show substantially greater decreases compared to their scores on easy texts. Logically this follows with the easy materials being easy for all and the more difficult material being more challenging for students with lower listening competence. However, the difficult material was of a sufficiently comprehensible level for all students to show some improvement. One of the challenges of implementing this technique is to find ideal levels of practice and assessment materials targeting both the level of actual development and the level of potential development (Vygotsky, 1962). What makes listening difficult varies and teacher decisions in selecting levels of materials are complex.

Second, the repetitive aspect of the shadowing technique has been thought to respond to a basic human drive to imitate and repeat according to Tannen (2007). Sasaki (2007) states that the mechanical characteristic of shadowing practice has been considered appropriate for Japan's context, where traditional, mainstream education is a memorization-centered pedagogical culture, which could be why shadowing has gained acceptance in Japan's L2 learning context.

Third, as Hamada (2012) points out in his discussion, the different difficulty level combination has "a positive influence on learners' psychology, especially on anxiety".

The alternating of difficult and easy material allowed students to still experience success despite the variation in the level of achievement. From anecdotal observation by teachers, the more challenging material caused students to have higher anxiety levels one week, but this was then followed by easier materials the next week, thus lowering their anxiety level. The differences in difficulty levels allowed students to regain their confidence from one week to the next and possibly helped to sustain a high motivation level.

Students were taught the function and benefits of shadowing during this process. This allowed students to fully understand the why and how of doing this activity. Knowing the specific purpose of the activity helped students to activate the cognitive processes necessary for the student to succeed. In addition to this, the training session procedure allowed students to see their pre- and post-test results. Students could see improvement in their test scores for both difficulty levels which was encouraging and motivating. Motivation consists of several factors; among them, self-efficacy. Self-efficacy is referred to as the strength of one's belief in a specific type of activity (Hamada, 2017) or one's confidence in being able to complete an assigned task or activity (Matsunuma, 2006). Bandura (1993) defines self-efficacy as belief in one's capability concerning specific tasks.

If one believes strongly in their self-efficacy, it can determine their success in actual abilities, skills, or knowledge (Schunk 1991). How a person perceives their capability can help decide what that person does with the knowledge and skills they have.

Limitations of this study

There are several limitations to this study. The number of participants was small therefore the data set is minimal. In addition to this, not all the participants attended each

session. There were a few students absent for several sessions. Also, the sessions were on Thursday afternoon from around 3:00. This was not ideal timing with it being later in the week and later in the afternoon. Students often said they were tired and appeared unable to give the task their full attention. Six out of 19 students did not improve their post TOEIC scores. Although checking the results of the pre-and post-cloze for those 6 students on all eight days shows an increase in scores every time. Yet these students did not improve their post TOEIC score. This could be due to these students not being in their best condition on the TOEIC post-test day.

Conclusion

Although it is difficult to deduce clear evidence from this study due to the very small-scale, the findings point to some useful observations that may be of benefit and interest to teachers. The data gathered indicate that a combination of two different difficulties of materials improved the listening comprehension skills of learners of varying levels of proficiency. This approach to shadowing can help encourage and advance individual learning regardless of level. The differences in difficulty levels allowed students to regain their confidence from one week to the next and possibly helped to sustain a high motivation level. This in turn could promote a stronger sense of self-efficacy. Although this study's results are very limited, these findings may contribute to the growing body of evidence supporting shadowing as an effective technique to improve students' listening comprehension.

References

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, 28, 117–148.

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), Encyclopedia of human behavior (Vol. 4, pp. 71–81). New York: Academic Press. (Reprinted in H. Friedman [Ed.], Encyclopedia of mental health. San Diego: Academic Press, 1998).

Educational Testing Service. (2022). https://www.iibc-global.org/english/toeic/test/lr/about/format.html

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

Hamada, Y. (2017). *Teaching EFL learners shadowing for listening: Developing learners' bottom-up skills*. London and New York: Routledge.

Hamada, Y. (2018). Shadowing: What is It? How to Use It. Where Will It Go?. *RELC Journal*. https://doi.org/10.1177/0033688218771380

Matsunuma, M. (2006). Eigojikokouryokukanshakudo no sakusei [Development of SES scale]. *Bulletin of the Graduate School of Education, Waseda University*, 14, 89–97.

Sasaki, M. (2007). The 150-year history of English language assessment in Japanese education. Language Teaching, 25(1), 63–83.

Schunk, Dale H. (1991). "Self-Efficacy and Academic Motivation." Educational Psychologist, 26: 207–231 http://dx.doi.org/10.1080/00461520.1991.9653133

Shiki, O., Y. Mori, S. Kadota, S. Yoshida. (2010). Exploring Differences Between Shadowing and Repeating Practices : An Analysis of Reproduction Rate and Types of Reproduced Words. *Annual Review of English Language Education in Japan*. Volume 21. https://doi.org/10.20581/arele.21.0_81

Tannen, D. (2007). Talking Voices. Repetition, Dialogue, and Imagery in Conversational Discourse (2nd ed.). Cambridge, UK: *Cambridge University Press*.

Vygotsky, L. S. (1962). Thought and language. Cambridge, MA: MIT Press.

Wallace, Matthew, P. and K. Lee. (2020). Examining Second Language Listening, Vocabulary and Executive Functioning. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2020.01122