

The Effect of Shadowing Practice:
A case study of Japanese language learners
in an Australian University

Hideki Sumiyoshi

MA in Arts in International Communication, Macquarie University
MA in Applied Linguistics (TESOL), Macquarie University

BA in Human Science (International Studies), Bradford College
BA in Foreign Studies (English), Kansai Gaidai University

October, 2014

Table of contents

ABSTRACT	V
CHAPTER 1	1
1.1. Literature Review	2
1.2. Research Questions	11
CHAPTER 2. RESEARCH METHOD	12
2.1. Participants	12
2.2. Materials	15
2.2.1. Shadowing materials	15
2.2.2. Pre-/ Post-shadowing tests	17
2.2.3. Pinpoint Dictation	17
2.2.4. Anonymous Survey	19
2.3. Procedure	21
2.3.1. Shadowing period	21
2.3.2. Shadowing submission and feedback	21
2.4. Data Analysis	23
CHAPTER 3. FINDINGS AND DISCUSSION	24
3.1. Shadowing Performance	24
3.1.1. Overall review of shadowing performance	24
3.1.2. Shadowing performance by initial competence level	27
3.2. Pre-/Post-Shadowing Tests	31
3.2.1. Overall review of pre-/post-shadowing tests	31
3.2.2. Pre-/post-shadowing tests by initial competence level	32
3.2.3. Pre-/post-shadowing tests by submission frequency	33
3.3. Pinpoint Dictation	37
3.3.1. Overall review of Pinpoint Dictation	37
3.3.2. Pinpoint Dictation between W7 and W7 Repeat in W11	42
3.4. Anonymous Survey	47

3.4.1. Participants' perceived improvement in listening and speaking skills	47
3.4.2. Participants' perceived effective speed of shadowing	48
3.4.3. Participants' preference of script availability.....	49
3.4.4. Participants' perceived positive aspects of shadowing.....	50
3.4.5. Participants' perceived negative aspects of shadowing	57
 CHAPTER 4. CONCLUSION: IMPLICATIONS AND LIMITATIONS.....	62
 4.1. Summary of key findings	62
4.1.1. Research Question (1)	62
4.1.2. Research Question (2)	62
4.1.3. Research Question (3)	63
 4.2. Implications for implementing shadowing practice in Japanese classes.....	64
 4.3. Implications for further study.....	69
 REFERENCES	71
 APPENDICES.....	75
 Appendix A.....	75
 Appendix B.....	77

List of Tables

Table 1. Six Steps of Shadowing Training for Beginners	6
Table 2. Summary of Participants.....	13
Table 3. Progression of Weekly Shadowing Model Audio Speed	16

List of Figures

Figure 1. The current version of the multi-component working memory model	3
Figure 2. Phonological Loop Construction	5
Figure 3. Shadowing Score by Accuracy (%).....	25
Figure 4. Shadowing Score by Proficiency Level (Advanced Level)	28
Figure 5. Shadowing Score by Proficiency Level (Intermediate Level)	29
Figure 6. Pre-/Post-Shadowing Test Mean Scores	31
Figure 7. Pre-/Post-shadowing Test Scores by Initial Competence Level	32
Figure 8. Pre-/Post-Shadowing Test Scores by Submission Frequency	35
Figure 9. Pinpoint Dictation Mean Scores.....	37

Figure 10. W09 & W10 PD Mean Scores by W09 Shadowing Submission	40
Figure 11. PD Mean Scores by W07 & W07 Repeat in W11	42
Figure 12. W07 & W07 Repeat PD Scores by Submission Frequency (Part A: Slow). 44	
Figure 13. W07 & W07 Repeat PD Scores by Submission Frequency (Part B: Fast) . 45	
Figure 14. Participants' Perceived Improvement in Listening and Speaking.....	47
Figure 15. Participants' Perceived Most Effective Shadowing Speed	48
Figure 16. Participants' Preference for Shadowing Script Availability	50

Acknowledgements

I am heartily thankful to my supervisor, Dr. Chavalin Svetanant, who has given me invaluable advice throughout this project. This study would not have been completed without her patient guidance, encouragement, and support, which enabled me to develop deeper understanding of this project. My appreciation also goes to the Head of Japanese Studies, Dr. Mio Bryce, for her valuable advice and support at the commencement of this project.

Declaration of originality of research

I certify that the research described in this thesis has not already been submitted for any other degree. I certify that to the best of my knowledge all sources used and any help received in the preparation of this thesis have been acknowledged.

Signature

Date

Abstract

The aim of this study is to investigate the effective implementation of shadowing practice in Japanese language units at an Australian university. This study explores how shadowing can be effective for improving listening skills based on Kadota's (2007) theoretical point of view with a focus on the gradual speed progression of the model audio. This study addresses the research questions regarding (1) the most effective speed for shadowing model audio; (2) differences between advanced and intermediate level Japanese learners in the relationship between shadowing and listening ability improvement and (3) participants' perception of shadowing as an effective practice in improving listening ability. The participants in this study were 62 intermediate and 35 advanced level Japanese learners. Two listening comprehension tests (pre-shadowing and post-shadowing test) and weekly pinpoint dictation quizzes were administered to measure the effect of shadowing on improving listening ability. An anonymous survey was collected at the end of the semester to assess the qualitative implications of the participants' shadowing experiences. The results and findings are analysed and the implications for classroom applications are discussed.

Chapter 1.

In the field of second language acquisition, shadowing has increasingly been recognized as a very effective practice for developing listening skills in L2 learning in Japan. Shadowing refers to ‘a paced, auditory tracking task which involves the immediate vocalization of auditorily presented stimuli, *i.e.*, word-for-word repetition, *in the same language*, parrot-style, of a message presented through headphones’ (Lambert, 1992, p. 266). It was originally used as a training method for simultaneous interpreters as the shadower is required to listen and speak at the same time. Therefore, this simultaneous task of listening and speaking has been suggested as an effective means of improving listening skills in L2 learning. However, Lambert’s metaphor of ‘parrot-style’ may give an impression that shadowing itself does not require much brain work since simultaneous interpreters must perform a series of cognitive tasks while listening (Spiller et al, 1990). According to Tamai (2002), ‘shadowing is an act or a task of listening in which the learner tracks the heard speech and repeats it as exactly as possible while listening attentively to the incoming information’ (p. 181). This definition of shadowing highlights the importance of active attention to in-coming sound in terms of linguistic processing in the human brain; otherwise, without paying attention the speech becomes as good as background noise. Therefore, Tamai’s version of shadowing definition is more suitable for implementing shadowing practice into language classes. The present study is conducted as a requirement for the degree of Master of Research (pathway to Ph.D.), and as a pilot study for the author’s future doctorate degree. It attempts to explore the most effective way to implement shadowing practice in Japanese language classrooms in an

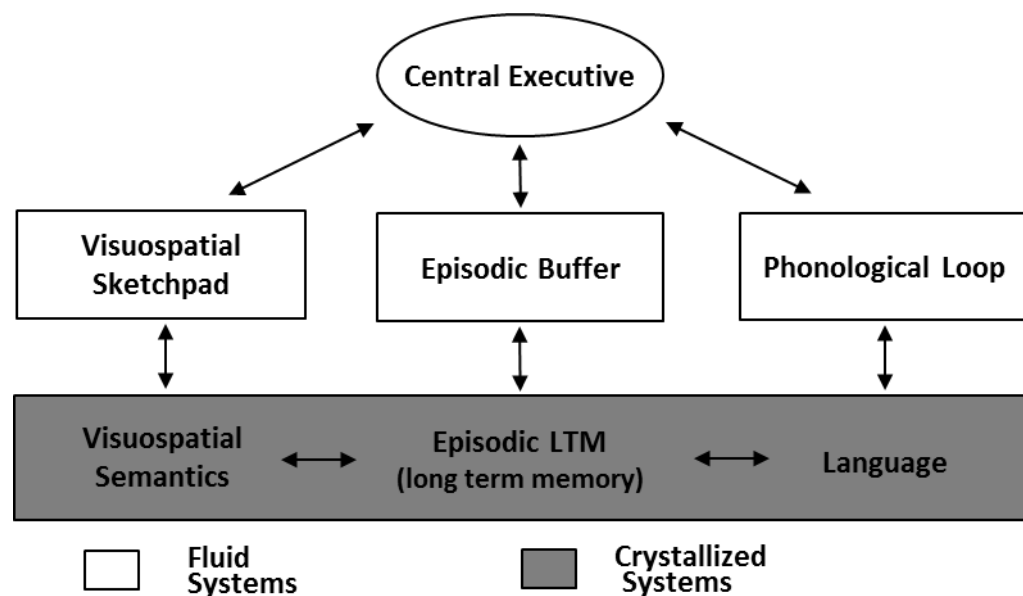
Australian university. Firstly, Baddeley's (1992, 2000, 2002) working memory theory is discussed in order to highlight the cognitive mechanism of shadowing. Secondly, in relation to the shadowing effect and listening skills improvement, Kadota's (2007) theoretical approach to second language acquisition is referenced to explore how to implement shadowing practice in classroom language teaching. And lastly, in reference to previous shadowing studies in both English and Japanese language teaching, this study attempts to develop a shadowing practice with a focus on the speed of the progression throughout the study period. The reason for the focus on the speed of the progression is because the author believes that the listening skill is the only language skill which the language users cannot control the speed by themselves, unlike the other three areas of language skills (reading, writing and speaking) that allow speed control by the language users. Although the listening skill is the most common area that learners of Japanese language wish to improve (Yang, 2008), there is no study yet focusing on speed of progression in the field of shadowing research.

1.1. Literature Review

The mechanism of the human brain often used to be explained using the metaphoric expression of a computer's memory system and the central processing unit (Lakoff & Johnson, 1980). From the point of view of cognitive science, this study refers to Baddeley's (1992) working memory theory in explaining the shadowing practice mechanism as listening and speaking simultaneously requires a cognitively heavy workload for shadowers to perform since this task rarely happens in daily conversation. As shown in Figure 1, Baddeley (1992) proposes the term 'working memory' (hereafter,

WM) of the human brain in terms of information processing of incoming visual and sound stimuli. He describes the following three components: the central executive operates as an attentional-control system; the visuospatial sketchpad processes visual images and the phonological loop is for incoming sound storage.

Figure 1. The current version of the multi-component working memory model



Note: Created by the author with reference to Baddeley (2000, p. 93)

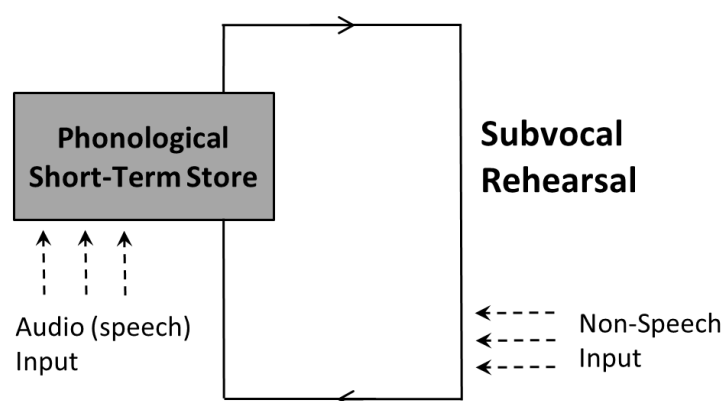
The central executive is solely an attentional system with no storage capacity of its own, and it functions as a processing device between the two ‘slave systems’ with their own temporary short-term memory function (Ibid, p. 556). To illustrate the process of listening comprehension, the heard speech is temporarily stored in the phonological loop, which is considered to last only one to two seconds (Ibid, p. 558). The central executive then runs a semantic analysis of the phonetic signals by referring to the long-term memory (hereafter, LTM), after which comprehension is achieved after examining the pragmatic and contextual usage of the language. Here, the crucial point in this listening comprehension processing is that the capacity of the central

executive is limited. Furthermore, if the central executive mobilizes a large portion of its resources to focus on visuospatial sketchpad or episodic buffer¹ linkage instead of the phonological loop, listening comprehension may not occur as the link between the central executive and the phonological loop becomes very limited or even non-existent. In other words, people do not always comprehend what they hear even in their native language, especially when they are in an absent-minded or multi-tasking state because the heard speech is not processed for comprehension.

One of the key functions explained above is the phonological loop. With respect to the listening process, it is assumed to have two additional functions: phonological short-term storage and subvocal rehearsal (see Figure 2). The former is the memory function of sound stimuli, and the latter is the mental repetition of the heard speech stored inside the phonological loop. This subvocalization is assumed to last for about two seconds (Darro & Fabbro, 1994), but Kadota (2007) suggests that the length of subvocal rehearsal can be extended with active subvocalization, and Klingberg et al (2002) also suggests the possibility of improving WM capacity. Namely, in order to retain the heard speech longer, the role of the central executive becomes crucial as it will require larger processing resource power for more active subvocalization.

¹ Episodic buffer is assumed to represent a storage system that holds integrated episodes or scenes; therefore, it is purely mnemonic in character (Baddeley, 2000, 2002).

Figure 2. Phonological Loop Construction



Note: Created by the author with reference to Kadota (2007, p. 150), reproduced from Gathercode and Baddeley (1993)

From the second language acquisition point of view, Kadota (2007) refers to the listening comprehension process in terms of two opposite approaches that L2 learners typically take: bottom-up processing and top-down processing. The former refers to the processing starting from the minimum element of the speech, such as sound, which forms a morpheme, then a word, then a sentence and then finally the context and pragmatic use of the language, which lead to comprehension. On the contrary, top-down processing refers to processing starting from the context and only recognized words to guess the meaning by compensating for the missed words. This listening technique is naturally practiced by native and advanced L2 speakers, especially in daily conversations where incomplete sentence exchanges are likely to occur. For example, linguistically speaking, functional words, such as the English articles 'a' and 'the', are often pronounced much weaker than other words (Nakayama, 2011), and Japanese particles are often omitted in spoken language (Maeda, 1998). However, for novice and intermediate L2 learners, it is of crucial importance to train sound recognition skills for bottom-up processing in order to maximize the number of recognizable words from their

limited lexicon.

In implementing shadowing practice in the language teaching context, Kadota and Tamai (2004) recommended the following six shadowing steps (p. 62) (as shown in Table 1 below).

Table 1. Six Steps of Shadowing Training for Beginners

Steps	Procedure	Details
1	Listening	Listening to the audio without the script, and trying to roughly grasp the content and the speech style.
2	Mumbling	Shadowing without the script focusing on the heard sound rather than reproducing pronunciation.
3	Synchronized Reading (content understanding)	Shadowing with the script focusing on the meaning of the script.
4	Prosody Shadowing	Shadowing focusing on prosodic features, such as the stress, rhythm, intonation, speed, pause, etc.
5	Synchronized Reading (difficult points)	Shadowing with the script focusing on the parts listeners find difficult.
6	Content Shadowing	Shadowing focusing on the content without reading the script.

This set of shadowing steps has been widely referred by teachers and researchers (Hamada, 2012; Kyo, 2012; Saito, Nagasawa, & Ishikawa, 2011, etc.). Since Kadota recommended that the shadowing scripts should be easier than the learners' current language level, e.g. using less than 3% of unknown vocabulary, steps 3 and 5 (Synchronized Reading) reasonably fit in the process of shadowing development as they play the important role of eliminating the gap between unfamiliar pronunciation and identification of vocabulary in the script. However, Kondo (2012) pointed out the possibility of a negative effect of this reading process during the shadowing practice; that is, whether the shadower neglects the phonological attention by using the recognized

vocabulary from the LTM through the reading process. She also implied a positive effect of not showing the script during the shadowing practice, arguing that it is closer to authentic listening in actual conversation in the target language. In fact, the timing of the shadowing script use has become controversial as some researchers have found it more effective when the script was shown after the shadowing (Hamada, 2014; Kondo, 2012). Hence, in this study, the script of the shadowing practice is not shown before or during the practice. This post-shadowing timing of the script is also considered crucial to the research design of this study to mark the learner's shadowing performance by its accuracy, which is an attempt to measure the sound recognition ability without visual assistance or preliminary knowledge.

Among other findings which influenced this study is Shiki et al.'s (2010) report that the shadowing effect likely hits the ceiling point of the reproduction rate after four or five trials. Although Shiki et al.'s study was conducted during a one-day session, and the implication may be limited within the conditions of the study, this ceiling phenomenon was observed in Tamai's (2002) studies as well. However, the shadowing materials are set at the average of the participants' L2 proficiency in these two aforementioned studies, and different materials (but the same difficulty level and speed) are used throughout the study period. Therefore, it is reasonable to assume that those at lower L2 proficiency show more improvement than those at a higher proficiency level as they are assumed to have more room for improvement. However, would those at a higher proficiency level improve more if they use more difficult shadowing material? Since it appears that it is the audio speed that the shadowers feel is the most difficult to follow (Chang, 2008), it is of vital importance to consider

the audio speed as one of the variables; however, most of the shadowing studies seem to leave this out. In this respect, Nakayama's (2011) study in an English as a Foreign Language (EFL) context in Japan presents a very important account addressing the speed difference of shadowing material. The objective of the study was to investigate whether the speed progression affects the shadowers' sound recognition of weak forms of function words (e.g. prepositions and articles). Nakayama's study used one shadowing script read at three different speeds for the experimental group, and the same script at the fastest speed for the control group. The study reported that the experimental group performed better than the control. However, as Nakayama pointed out himself, the result may be limited to the study context as the same script is used by showing the weak form of function words more clearly at slower speed, and thus the shadowers may have developed better recognition of the script than the control group. Therefore, the result may not imply that the shadowers' overall sound recognition skills have improved (Ibid, p. 30).

The present study attempts to implement shadowing practice in intermediate and advanced level Japanese language courses² at an Australian university during the first academic semester. In implementing shadowing practice in university language units, it is of vital importance to note that it is virtually impossible to create a control group as the university regulation for this study does not allow conducting a different curriculum within the same unit in order to ensure that every student receives the same learning materials. With regards to this limitation, it is understandable to see most shadowing studies

² For the intermediate level, students are required to have passed introductory Japanese units with a total of 104 contact hours of study, and advanced level students are required to have 208 contact hours as a prerequisite requirement for enrollment.

conducted in Japanese as a second language (JSL) as well as Japanese as a foreign language (JFL) contexts using a fairly small number of participants as they must recruit volunteers in order to create control and experimental groups (Iwashita, 2008; Kondo, 2012; Kurata, 2007; Takahashi, Fukuda, Iwashita, & Sakoda, 2010). In some cases, researchers might possibly conduct a joint project to recruit participants from different universities. Another important aspect that determines the conduct of shadowing practice in this study is that the classrooms are equipped with an interactive white board (IWB) that allows instructors to use computer software as well as audio and visual presentations; however, not all students have their own headphone jacks. Nonetheless, while it is possible to play the model audio and shadow along with it, the noise level produced by other students' shadowing voice easily exceeds the distractive level, which ruins the primary objective of the shadowing practice of only listening to the model audio. Therefore, it is extremely difficult to conduct shadowing in a classroom where students cannot use headphones all at the same time because hearing one's own voice is already disturbing; listening and speaking simultaneously is a highly cognitive-loaded task (Broadbent, 1952). Therefore, this study conducts shadowing practice as a homework assignment where distribution of model audio and collection of students' shadowing recording are done via an online unit delivery system called 'iLearn' (online unit module for course material delivery, assignment file submission, online discussion, etc.). This homework format of shadowing is considered to have both advantages and disadvantages. The advantages are that students can practice shadowing flexibly at their own time, pace and frequency with least possible foreign language anxiety in speaking the target language in front of other students in the classroom

(Horwitz et al, 1986), and that their audio submission can be assessed and feedback can be provided for each submission. However, the disadvantage may be that the instructor has no control over how students conduct shadowing in terms of the depth of their engagement; that is, the amount of practice time, the use of dictation method for visual assistance and the submission rate, all of which depend on the level of the student's motivation in shadowing practice. Therefore, it is of crucial importance for the instructor to give a full and appropriate explanation of the objective of shadowing practice and its theoretical background (Mochizuki, 2006).

1.2. Research Questions

Considering the above discussion, this study firstly investigates the effect of shadowing in relation to the speed of the model audio, followed by the difference between the advanced and intermediate levels. The research questions are as follows:

- 1) What is the most effective shadowing speed range?
- 2) Are there any differences between advanced and intermediate level Japanese learners in the relationship between shadowing and listening ability improvement?

And lastly, the shadowing survey will be qualitatively analysed in order to judge whether the implementation of shadowing practice in Japanese language units in an Australian university was successful. The final research question is thus as follows:

- 3) Do participants perceive shadowing as an effective practice in improving their own listening ability?

Chapter 2. Research Method

2.1. Participants

The target participants of this study were Australian university students who were enrolled in intermediate and advanced level Japanese language units. Since shadowing practice and other listening-related tasks were part of the unit assessment, it was compulsory that students complete those tasks as unit requirements. However, students were told that their participation in this study was voluntary, and their decision would not affect their grades. They were asked to sign the consent form and answer the anonymous survey on the last day of the semester. Since the instructor and the author of this study is the same person, in order to assure complete anonymity, another teaching staff member collected these documents on behalf of the author by using a drop box for students to submit their consent form and anonymous survey while the author was absent from the classroom. The documents were kept in sealed envelopes until the release of the final grades in order to ensure that the students' decisions did not influence their grades. This participant consent procedure was approved by the university's Ethics Committee before the commencement of this study. Out of 43 advanced and 67 intermediate level enrolled students, 35 advanced level and 62 intermediate level students agreed to participate in this study. Table 2 summarizes the participants of this study.

Table 2. Summary of Participants

		Advanced		Intermediate	
Gender	Male	18	51%	23	37%
	Female	17	49%	39	63%
Enrolment	Domestic	29	83%	54	87%
	International	6	17%	8	13%
Native language	English	24	69%	37	60%
	Chinese	7	19%	20	32%
	Korean	3	8%	2	3%
	Vietnamese			2	3%
	Indonesian			1	2%
	Thai	1	3%		
Shadowing experience	Yes	1	3%	4	6%
	Yes (previous semester)	23	66%	0	
	Never	11	31%	58	94%

In the advanced level, there are almost an equal number of males ($n = 18$) and females ($n = 17$). The majority of the advanced level students are domestic students ($n = 29$) and several are international students ($n = 6$). Nearly 70% of this level speak English as a native language ($n = 24$); other native languages include Chinese ($n = 7$), Korean ($n = 3$) and one Thai speaker ($n = 1$). On the other hand, the intermediate level has more female students (39) than male students ($n = 23$), but a similar demographic pattern in the enrolment mode as the majority are domestic ($n = 54$) and international students ($n = 8$), and 60% are native English speakers ($n = 37$). There are also (relatively) many native Chinese-speaking students ($n = 20$) and other Asian language speakers, such as Korean ($n = 2$), Vietnamese ($n = 2$) and one Indonesian language speaker ($n = 1$). With regard to previous experience in shadowing practice before this study, very few participants (advanced = 1, intermediate = 4) had experienced shadowing, while 67% of the advanced level participants ($n = 24$) experienced

shadowing in the previous semester under the same instructor as a unit assessment. However, 31% of the advanced ($n = 11$) and 94% of the intermediate participants ($n = 58$) had never engaged in shadowing before this study.

2.2. Materials

2.2.1. Shadowing materials

For the shadowing materials, scripts contained approximately 5 to 10% of unfamiliar vocabulary, which was more than Kadota's (2007) recommendation of less than 3%. However, it was expected that the shadowers would rely less on their LTM for understanding of the script, instead focusing more on the audio speech because of the unfamiliar vocabulary. For the intermediate level unit, shadowing materials used the following week's course reading passages, which inevitably included unfamiliar grammar points but mostly familiar vocabulary and expressions, ranging between 150 and 200 mora³. These unfamiliar grammar points, however, can be considered relatively manageable according to Krashen's (1985) Input Hypothesis: the 'natural order' of language acquisition moves from the current language level of i to $i + 1$ (Ibid, p. 81). For instance, past-tense marker [-ed] in English language must be taught before introducing irregular past-tense conjugation verbs (e.g. to read, to eat, etc.) because the learner must comprehend the concept of 'past tense' before becoming able to handle different types of past tense conjugations. On the other hand, the advanced level unit used authentic reading materials available on Japanese websites, ranging between 250 and 400 mora. Both reading materials were recorded by a male native Japanese speaker, who was the instructor of all the tutorial classes of the units and the author of this study. The speed of the model audio began very slowly in the first week of shadowing (approximately 240 mora per minute in Week 6); the speed gradually increased towards native-like speed at the end of the semester

³ Japanese is a mora-timed language in which each mora is spoken at a roughly constant rate, while English is a stress-timed language with syllables (Ishihara, Tsurutani, & Tsukada, 2011).

(approximately 380 mora per minute⁴ in Week 11). The length of the model audio recording was approximately 30 seconds for the intermediate level and 60 seconds for the advanced level. Table 3 summarizes the shadowing speed common for the advanced and intermediate level.

Table 3. Progression of Weekly Shadowing Model Audio Speed

Week	W06	W07	W08	W09	W10	W11
Speed (mora/minute)	240	270	300	320	350	380

This adjustable speed of shadowing materials is particularly unique in this field of shadowing research. The greatest advantage of applying this increasing speed of the shadowing audio is assumed to be the scaffolding effect for almost all the participants since the first material starts at 240 mora per minute (m/m). However, the speed increased by approximately 30 mora every week, and eventually reached 380 m/m by the end of the study period. During this six-week speed range, it is reasonable to assume that most participants reach their ceiling speed of shadowing in earlier weeks than listening alone since the shadowing task itself is already cognitively challenging as no one listens and speaks at the same time (Kadota, 2007). Therefore, starting at a slower speed than necessary contributes to allowing the learners to practice shadowing, especially for those who are experiencing shadowing for the first time. It is also assumed to mitigate learners' anxiety towards spoken language if the speed is manageable for the shadowers (Gregersen & Horwitz, 2002). However, when the speed increases to the point where the shadower feels challenged in processing the in-coming sound, the learner needs to practice

⁴ Data based on NHK news and weather forecast (five samples each). Calculation conducted by the author.

more in order to be able to perform the shadowing at their satisfactory level; such practice is considered vital for improvement. This planned speed progression was explained to students along with the introduction of shadowing practice in the commencing week.

2.2.2. Pre-/ Post-shadowing tests

As for the pre-/post-shadowing test materials, Japanese Language Proficiency Test (henceforth, JLPT) mock examination collection books⁵ were used. At the beginning of the academic year during the first semester, N4 for intermediate level students and N3 for advanced level students were considered as appropriate levels since the end-of-academic year target levels are N3 and N2, respectively. Therefore, in order to cover the current proficiency level among participants, N5, N4 and N3 mixed levels materials were used for intermediate level students, and N4, N3 and N2 mixed levels materials were used for advanced level students. The content of these tests are designed to measure students' overall listening skills, which is based on comprehensive linguistic competence 'to use the knowledge in actual communication' (JapanFoundation, 2012).

2.2.3. Pinpoint Dictation

As a counterpart to the pre-/post-shadowing tests, the author devised a dictation quiz called pinpoint dictation (PD) in order to measure the specific listening skill focusing solely on sound recognition, where students are asked to write down only certain information from a short sentence. This quiz was

⁵ *Japanese Language Proficiency Test – Complete Mock Exams* series from J Research Publisher (Watanabe et al., 2013)

inspired by the cloze test⁶ and the listening span test⁷ format. In this test, 10 short Japanese sentences were read, followed by a question in English. The sentence included two to three pieces of random information containing either random numbers or names that L2 learners are likely unfamiliar with (e.g. celebrities and popular travel destinations are not used in this test).

Consider the following example:

Sentence: このねぎとろ井は 560 円だ。

(kono negitorodon wa gohyakurokuju en da.)

(This spring onion and raw tuna mince on rice costs 560 yen.)

Q (pattern 1): How much is the dish?

A (pattern 1): 560 yen.

Q (pattern 2): What is the name of the dish?

A (pattern 2) ねぎとろどん (negitorodon).

Each question was read only once with the interval slightly longer than the estimated time for writing the answers (answers can be written in any form of phonetic writing systems, e.g. kana scripts, alphabet, numerals, etc.). The whole task took about three minutes. Each answer is marked by the portion of accurate sets of phonetic representation in 4-scales (with 4 points being a full mark) so that the test score will represent greater accuracy than the all-or-nothing scoring method. PD does not attempt to measure listeners' comprehension, instead measuring sound recognition and retention as the

⁶ A widely used language test to fill in the blanks in the text with a number of words that have been deleted (Kobayashi, 2002).

⁷ Listening version of the reading span test which asks the testee to memorize the last words of the three read sentences without taking any notes and then answer after reading all the sentences (Kadota, 2007).

answers are phonetic and thus do not require complex linguistic processing or previous knowledge to answer (Kadota, 2007). Therefore, PD is considered to specifically measure the bottom-up processing, unlike most other listening tests where questions are available before the audio is played and is normally repeated two to three times; by then, the listeners' focus has already shifted towards the content of the speech, which can be processed top-down. In order to measure participants' listening skill at different speeds, PD is divided into the following two parts: the first five sentences (part A) for relatively slow speed and the second five sentences (part B) for fast speed. The advanced and intermediate levels use the same PD sentences with the same set of English questions, but the speed was set at different rates as advanced level students were considered more accustomed to a faster spoken speed than intermediate students since they have longer study experience with more exposure to listening to spoken Japanese. Initially, W6 PD was recorded in three different speeds that were equivalent to W7, W9 and W11 shadowing speed. For the advanced level, W9 speed (320 m/m) was used for part A and W11 (380 m/m) for part B. For the intermediate level, W7 (270 m/m) was used for part A and W9 (320 m/m) for part B. In order to maintain the consistent speed difference throughout the later weeks, W6 PD audio was always reviewed carefully by the instructor who recorded the audio focusing on the different speeds right before the week's PD audio recording.

2.2.4. Anonymous Survey

In order to obtain a source for qualitative analysis, an anonymous survey was conducted at the end of the semester (see Appendix), which contained questions regarding students' perceived improvement in language skills in

5-point Likert scales as well as other shadowing experiences, such as the frequency of practice and the most effective speed range of model audio. This was followed by open questions addressing positive and negative aspects of shadowing practice. The collection of the survey was conducted carefully in order to secure anonymity; furthermore, the collection procedure met the university's Ethics Committee's protocol (see *Section 2.1.* for more details).

2.3. Procedure

2.3.1. Shadowing period

Out of the 13-week semester period, the shadowing practice was designed to be conducted for six weeks between weeks W05 and W10, as there were other assessment tasks, such as a listening comprehension quiz between W01 and W04, a major test in W12 and a speaking test in W13. The theoretical background of the shadowing was explained to the students during the W05 tutorial class before the commencement of weekly shadowing assessment as it is considered crucial to share such information about the exercise in order to help students better understand the objectives (Mochizuki, 2006). Also, it was explained to the students that the script was not made available during the period of shadowing because the purpose of the exercise was to train their sound recognition ability. At this stage, the increase of the shadowing model audio speed was also mentioned in order to raise students' awareness about the challenges they could expect in the future weeks.

2.3.2. Shadowing submission and feedback

During the six-week shadowing period, students were required to conduct shadowing as a homework assignment. The model audio was made available via iLearn on the day of the tutorial class⁸, and the submission due date was set four days later in order to maintain the weekly cycle. Feedback was given during the following week's tutorial class. During the tutorial, the shadowing feedback took the following steps: 1) the instructor played the model audio

⁸ The language units consist of a lecture and a tutorial with two contact hours each per week. The former introduces new grammar points and is conducted in a theater-type classroom with a maximum of 100 students, while the latter is conducted with a maximum of 20 students and consolidates the lecture content using various activities, such as workbook reading and writing exercise, conversation practice, etc.

once while everyone is shadowing altogether; 2) the instructor played the model audio again but shadowing only one sentence by a randomly assigned student to demonstrate their shadowing performance (different students took turns by sentence until the end of the week's model shadowing); 3) the instructor points out the common mistakes and provides general comments on the week's shadowing performance; 4) show the average score of all the students; and 5) return the marked feedback sheet, which is a printed script with mistakes highlighted in red with the spelling of wrong pronunciations provided above the highlighted part. This shadowing feedback normally took approximately 15 minutes each week.

2.4. Data Analysis

This study uses two types of quantitative and qualitative data: the former includes participants' shadowing performance, pre-/post-shadowing test scores, pinpoint dictation scores and some questions from the anonymous survey; the latter includes the anonymous survey's open questions. Since the current study lacks control groups, the possible comparative analysis between groups shrank the sample size dramatically, leaving the large size differences between the groups. No statistical analytical tools are considered appropriate for such large size differences. Therefore, a comparative analysis of the mean scores between groups was used. As for the survey's open-ended questions, qualitative analysis was used to examine the participants' comments on shadowing experiences.

Chapter 3. Findings and Discussion

The first half of this chapter presents the descriptive statistics of the participants' shadowing and tests performance as well as analysis of the data in relation to possible variables that are considered influential on the results. This chapter firstly reports participants' overall shadowing performance, followed by the analysis of pre-/post-shadowing tests. Finally, it discusses the pinpoint dictation (PD) results and key findings that are considered relevant to the effect of shadowing on learners' listening skill. The second half of this chapter presents the results of the anonymous survey in light of the participants' perceived response to shadowing experiences. It firstly reports perceived improvement in listening and speaking skills through shadowing, and secondly, the participants' perceived most effective shadowing speed, followed by their preference regarding the shadowing script availability. Finally, it discusses the positive and negative aspects of shadowing by including relevant excerpts from the participants' comments.

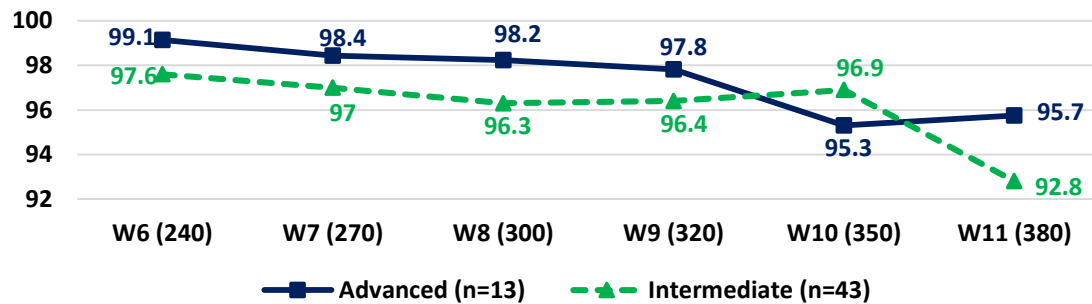
3.1. Shadowing Performance

3.1.1. Overall review of shadowing performance

Figure 3 shows the overall shadowing mean scores in the advanced and intermediate level conducted between week 6 and week 11. The numbers in brackets next to the weeks are the mora per minute (m/m), which indicates the speed of the shadowing audio. The number of participants in the advanced level (37%, $n = 13$) and the intermediate level (69%, $n = 43$) are the result of selecting only the participants who submitted shadowing homework from all 6 weeks in order to obtain only the mean scores by the same individuals who conducted shadowing practice all six weeks. By doing so, the mean scores

represent a genuine shift among the same participants. This filtering method is applied throughout the following sections in order to make genuine comparisons of the mean scores. Therefore, the number of participants varies depending on the variables used for each section.

Figure 3. Shadowing Score by Accuracy (%)



In general, both levels show a decline from W6 to W11 (advanced = -3.4%, intermediate = 4.8%). The advanced level students showed a constant decline except for W10 where a sharp drop is shown, and then in the final week (W11) a very slight increase of 0.4%. The intermediate level also showed a steady decrease towards W8, but slight improvement in W9 and W10. They then dropped suddenly in W11 by 4.1%, which is the biggest decrease in all the weeks in both levels. It appears as if the fluctuation pattern can be divided into two periods: the average maintains at a certain level ('holding period') and drops ('dropping period'). The 'holding period' is one week shorter in the advanced level between W06 and W09 than in the intermediate level between W06 and W10. One of the main factors that can be considered as the reason for this difference may be that the length of the shadowing model audio is significantly longer for the advanced level (approximately 60 seconds) than the intermediate level (approximately 30 seconds), which may have affected the performativity at different speeds as it is quite natural to assume that

3.1. Shadowing Performance

students would make more mistakes during a longer shadowing time. Another factor can be explained by the difference of the text type: the advanced level uses unlimited text range from authentic texts from web pages that are likely to involve more unfamiliar vocabulary and expressions, while the intermediate level uses limited texts within curriculum reading passages. One important note to make regarding the shadowing scores is that the increase or decrease does not simply mean an improvement or decline in the shadowing skills; considering the steady speed increase week after week, the same scores may well suggest improvement in shadowing performance since it is fairly natural to assume that the faster the audio speed gets, the more difficult it is to shadow and therefore the lower the score would become if there is no improvement. For example, a W11 score of 92.8% in the intermediate level is arguably more valuable and skilled than a W06 score of 97.6% because the speed is significantly faster in W11 (380 m/m) than W06 (240 m/m). In this regard, advanced W11 increase (by 0.4%) may be the result of more significant improvement in shadowing skill than it appears.

Such differences in shadowing materials may have affected the lower submission rate in the advanced level (37%) more than the intermediate level (69%), which means that 63% of advanced level participants had less shadowing experience than those with all six submissions. The same is true for 31% of the intermediate level participants. This difference in shadowing practice frequency among participants was the result of the nature of the homework assignment in which the level of engagement was determined by how the students decided to prioritize their assignments. For instance, students naturally prioritized major assessments, such as the mid-term (15%)

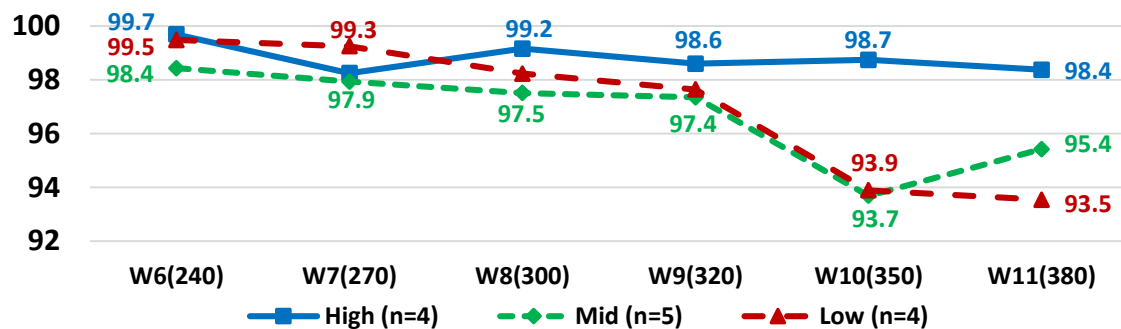
and final exams (25%), over minor assessments, such as written homework, online quizzes and shadowing (grouped together as one category of Homework which is worth 15% of their final grade, and therefore each weekly completion is only worth between 0.5 and 1%). Another factor that may have affected the difference of the shadowing submission rate is the ‘newness’ of this exercise: almost all of the intermediate level participants (94%) had never done shadowing practice before, whereas almost 70% of the advanced level participants had already experienced it during the previous semester (see Table 2 in *Section 2.1*). This ‘newness’ of the exercise may have resulted in greater motivation among intermediate level participants, resulting in a significantly higher submission frequency than the advanced level participants. On the other hand, the length of the shadowing audio for the advanced level is twice as long as the intermediate level, which may have demotivated those who submitted less frequently. Also, this more likely-hood of making mistakes than the intermediate level shadowing may have lowered the submission rate of participants who possess excessively high standards for self-evaluation, or so-called perfectionist students (Gregersen & Horwitz, 2002). Since shadowing feedback was based on the accuracy, perfectionist participants may have seen their shadowing performance more from the fear of failure than from the success because 95% accuracy is still 5% deduction from their ideal result.

3.1.2. Shadowing performance by initial competence level

Figure 4 illustrates the shadowing performance of the advanced level ($n = 13$) in three groups by examining the listening proficiency level during the first week when shadowing commenced. When dividing into groups, the

pre-shadowing test score and the W06 PD score were put together and then divided into three groups. If not evenly divided with the extra number, the *Mid* group was formed as the largest group followed by the *Low* group (*High*: n = 4, *Mid*: n = 5, *Low*: n = 4). This descriptive analysis of the mean scores within one level is considered beneficial in analysing the relationship between participants' listening competence and the weekly increase of speed progression.

Figure 4. Shadowing Score by Initial Competence Level (Advanced Level)

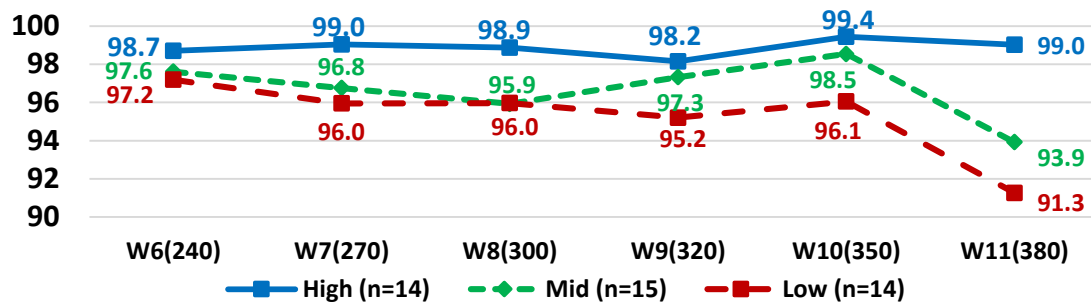


The listening competence level variable reveals that the *High* group does not have the 'dropping period' as shown in Figure 3 as it maintains above the 98% level throughout the entire period of analysis. In contrast, the *Mid* and *Low* groups follow a very similar pattern of a 'holding' and 'dropping period'. It is worth noting that the *Low* group's performance was higher than the other two groups almost throughout the entire period until the final week when the *Mid* group increased by 1.7% during W11. This may suggest that the difference in listening competence level becomes salient when the shadowing speed is faster than 350 m/m and an authentic script is shadowed for 60 seconds.

In the same comparison at the intermediate level, Figure 5 shows the

shadowing performance of the three groups at different listening competency levels (*High*: $n = 14$, *Mid*: $n = 15$, *Low*: $n = 14$).

Figure 5. Shadowing Score by Initial Competence Level (Intermediate Level)



Similar to the advanced level, the *High* group maintains consistently high scores throughout. Other groups perform according to their initial level: the *Low* group scores the lowest each week while the *Mid* group remains in the middle, and both groups drop sharply in the final week. As for the *Mid* group's scores during the 'holding period', it was observed in the mixture of two groups that the first three weeks (especially W06 and W08) are almost the same as the *Low* group. The following two weeks (W09 and W10), the mixture of two groups came closer to the *High* group, maintaining the margin from the *Low* group even during the 'dropping period' (W11). It is reasonable to assume that the shadowing speed affects the accuracy such that in the final week all the groups' scores dropped, including the *High* group (-0.4%). However, W10 scores increased from W09 in all groups at a very similar rate (*High* = 1.2%, *Mid* = 1.2%, *Low* = 0.9%), which is the only time this occurred during the entire period of study. As an instructor of the participants, the author wishes to believe that their efforts bore fruit and genuinely improved during this week, but it is more plausible to assume that there were other factors contributing to

the increase of the score that outweighed the speed increase, such as more manageable vocabulary, expressions, grammar usage and/or the pronunciation sequence in the script. In this respect, the drop of W11 performance in all groups may not have been only due to the speed increase.

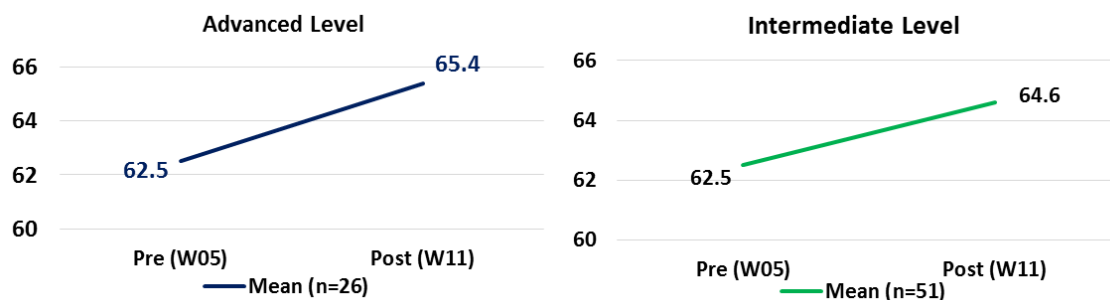
In sum, it was observed in the advanced and intermediate level that participants' listening competence is one of the indicators that affect shadowing performativity, and the gap becomes salient with respect to fast speed at 350 m/m for the advanced level learners and at 380 m/m for the intermediate level learners.

3.2. Pre-/Post-Shadowing Tests

3.2.1. Overall review of pre-/post-shadowing tests

Figure 6 shows the comparison of mean scores between pre-shadowing (W05) and post-shadowing (W11) in the advanced (n = 26) and intermediate level (n = 51). As discussed in the previous section, the number of participants in this section is different as a result of filtering for genuine group scores by the same individuals who took both tests for valid comparison of the mean scores. The results for pre-shadowing and post-shadowing tests show improvement in both levels. Advanced level learners improved their score by 2.9 points and intermediate level learners improved by 2.1 points.

Figure 6. Pre-/Post-Shadowing Test Mean Scores



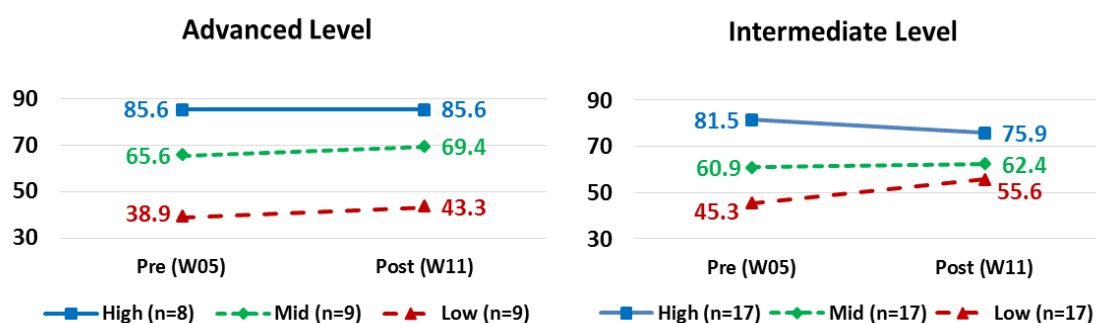
The pre-/post-shadowing tests are JLPT format tests whose characteristics are comprehensive in nature and include general linguistic competence which aims to measure the test takers' overall language proficiency in terms of knowledge and practice. In fact, the JLPT listening test aims to measure the comprehension of spoken Japanese with a focus on specific contextual language use (JapanFoundation, 2012). With regards to this testing content, the pre-/post-test results may likely lack construct validity in measuring the effect of shadowing alone which this study aims to investigate. Hence, the

improvement in the test scores may indicate the participants' overall improvement in Japanese language in the listening mode. In other words, the improvement is likely the result of participants' acquisition of new vocabulary and grammar items or even the development of test-taking strategies for this kind of test, and not necessarily the improvement of sheer listening skills based on the WM functions.

3.2.2. Pre-/post-shadowing tests by initial competence level

Figure 7 illustrates the listening competence level grouping in order to provide a more detailed analysis of the pre-/post-test results. The same grouping method as the previous section was used for dividing into three groups with reference to the participants' W05 pre-shadowing test and W06 PD scores.

Figure 7. Pre-/Post-shadowing Test Scores by Initial Competence Level



It is natural that pre-shadowing test scores in both levels are aligned almost equally in the order of *High*, *Mid* and *Low* as this score is one of the grouping sources. However, the overall projection appears to differ between the advanced and intermediate level: three advanced level groups remain flat in terms of the gap between the groups, whereas the gap between the intermediate groups narrows down towards the *Mid* group. However, a closer

look reveals that the advanced *High* group did not change its score (85.6 pts), while the *Mid* group improved by 3.8 pts and the *Low* group improved by 4.4 pts. In contrast, the intermediate *High* group decreased by 5.6 pts, while the other groups improved slightly: the *Mid* group improved by 1.5 pts and the *Low* group improved by 10.3 pts. In terms of the comparison between the *High* group (0 at the advanced level and negative improvement at the intermediate level) and the other two groups (both the *Mid* and *Low* groups improved at the advanced and intermediate levels), there seems to be a certain commonality in these levels. This common pattern of the different proficiency levels is known as the ‘ceiling effect’. This is considered a natural phenomenon in that already proficient learners have little to gain whereas less competent learners have more room for improvement (Dewey, Brown, & Eggett, 2012). This effect was also observed in Tamai’s (2002) study which reported that less proficient students improved more in a listening test than already proficient students.

3.2.3. Pre-/post-shadowing tests by submission frequency

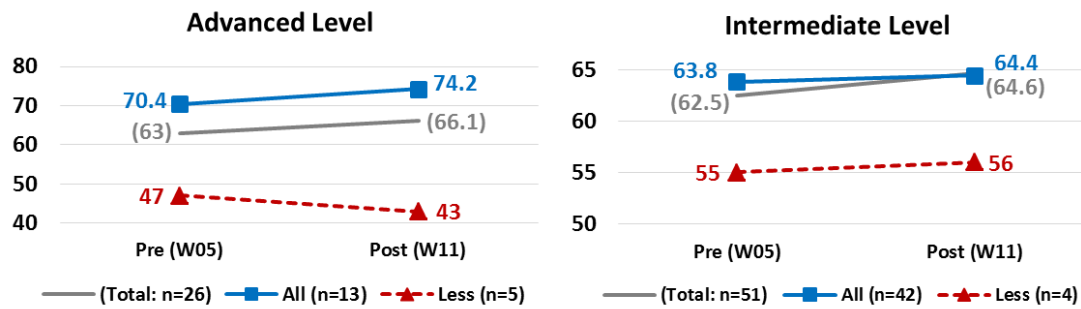
Since the current study was conducted based on the shadowing practice as part of a university language unit assessment, the creation of control groups (conducting different curriculum contents in different classes in the same unit, e.g. shadowing for class A, dictation for class B, no listening for class C, etc.) was impossible. Therefore, this study design lacks the very crucial comparative statistical analysis of the shadowing effect between control and experimental groups. However, the difference in the shadowing practice time – i.e. the difference in submission frequency – can be considered as one inferential indicator in measuring the effect of shadowing. Since shadowing was conducted as homework, there were several absent submissions each

week; there was no week in which all of the students in either level turned in their submissions.

In dividing into two groups between *All* submissions and *Fewer* submissions, participants that submitted their shadowing homework all six weeks are grouped as *All*, whereas participants with less than three submissions are labelled as the *Fewer* group (participants with between four and five submissions were excluded from the *Fewer* group as they may have sufficient shadowing experiences as a result of completing more than half of the submissions). Although the difference in submission frequency allows the measurement of the effect of shadowing alone, there is an inevitable shortcoming in that the *Fewer* group significantly lacks the number of corresponding participants (a maximum of six participants in the advanced level and seven in the intermediate level) even before conducting the filtering for the variables. However, the author believes that the difference in the submission frequency is a valuable analysis option to help reveal the effect of shadowing. Therefore, the following results are limited in terms of descriptive purposes.

Figure 8 shows the comparison of the mean scores of the pre-/post-shadowing tests with respect to the variable of shadowing submission frequency in two groups in the advanced level, *All* ($n = 13$) and *Fewer* ($n = 5$, submission average = 2.6 times), and in the intermediate level, *All* ($n = 42$) and *Fewer* ($n = 4$, submission average = 2.5 times).

Figure 8. Pre-/Post-Shadowing Test Scores by Submission Frequency



The common trend in both the advanced and intermediate level is that the *All* groups retain much higher scores than the *Fewer* groups. It is rather surprising to see this phenomenon as the grouping was made based on the shadowing submission frequency and not the initial listening competence levels, as shown in Figure 7 in the previous section. One possible explanation for this may be that the less competent participants are likely to struggle in shadowing, and as a result of this struggle, there were fewer submissions. This may have been the case especially for those who scored much less than average in the pre-shadowing test (advanced = -16 pts, intermediate = -7.5 pts). On the other hand, the *All* groups in both levels supersede the average, especially the advanced *All* group, which increased the gap from 7.4 pts in pre-shadowing to 8.1 pt in post-test. The intermediate *All* group narrowed down the gap from 1.3 pts in pre-test to -0.2 in post-test, but considering that the intermediate *All* group is the dominant population in the intermediate participants ($n = 42/51$), it is quite natural to remain close to the average score. Parallel to this group, the intermediate *Fewer* group ($n = 4$) maintains the increase by 1 pts. It is the advanced *Fewer* group that remarkably deviates from this increase trend, decreasing 4 pts and thereby widening the gap by 23.1 pts from the average and 31.2 pts from the *All* group. This massive

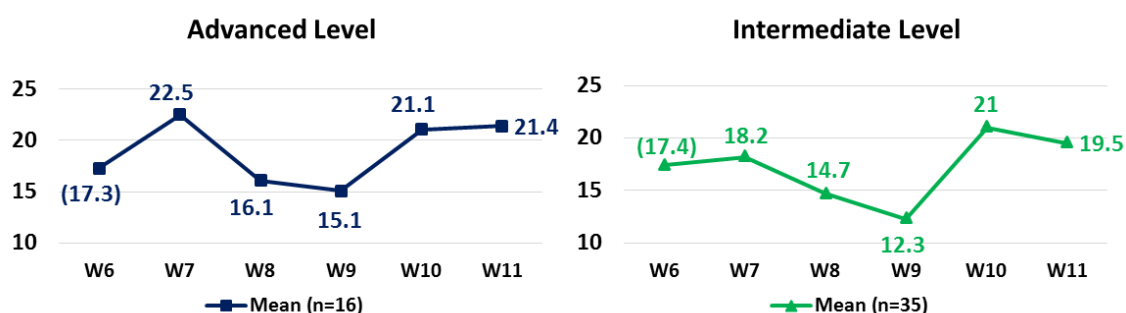
decrease rejects the possibility in the above discussion of the *Low* group ($n = 9$) coinciding with the same individuals as the *Fewer* submission participants ($n = 5$) because the *Low* group increased the score by 4.4 pts in the previous section (see Figure 7). This increase contradicts the 4 pts decrease in the *Fewer* group. Out of five possible groupings in the advanced level – *High*, *Mid*, *Low*, *All* and *Fewer* – only the *Fewer* group's score decreases. The fact that the other groups increased or maintained the scores may indicate that the shadowing practice time is a possible contributing factor to this increase in test scores.

3.3. Pinpoint Dictation

3.3.1. Overall review of Pinpoint Dictation

Figure 9 summarizes the mean scores of pinpoint dictation (PD) conducted between W06 and W11 in the advanced ($n = 16$) and intermediate ($n = 35$) levels. PD was conducted during the tutorial classes as a listening activity (10 sentences = 40 pts), and no supplementary opportunity, such as online submission, was made available in order to maintain the consistency in test-taking conditions. For example, online access to the audio would allow the test takers multiple trials and they would also be able to pause the audio recording, both of which are considered to contaminate the PD results as it is designed to measure students' sound recognition ability with a one-off trial of listening to the audio sentence. Therefore, the smaller number of participants than in the pre-/post-test is the result of the filtering method of selecting only those who attended the classes and took the PD in all six weeks in order to obtain genuine mean scores from the same individuals.

Figure 9. Pinpoint Dictation Mean Scores



Overall mean scores fluctuate in both the advanced and intermediate level following common patterns in terms of high and low shifts during the same week: both increase the scores in W07 from W06, then a gradual decline

towards W8 and W9, but then a sharp rise in W10 before almost going flat in W11. The W06 scores are in brackets (advanced = 17.3 pts, intermediate = 17.4 pts) as it was a ‘trial attempt’ to introduce PD to the class. During the first PD attempt in W6, some students showed quick adaptability to this quiz format, but others appeared perplexed as the audio sentences were only read once and the questions were asked only after each sentence. Therefore, the W06 PD performance is considered inapplicable for comparison to the other weeks. In this account, the improvement from W06 to W07 may be the result of becoming accustomed to the quiz format.

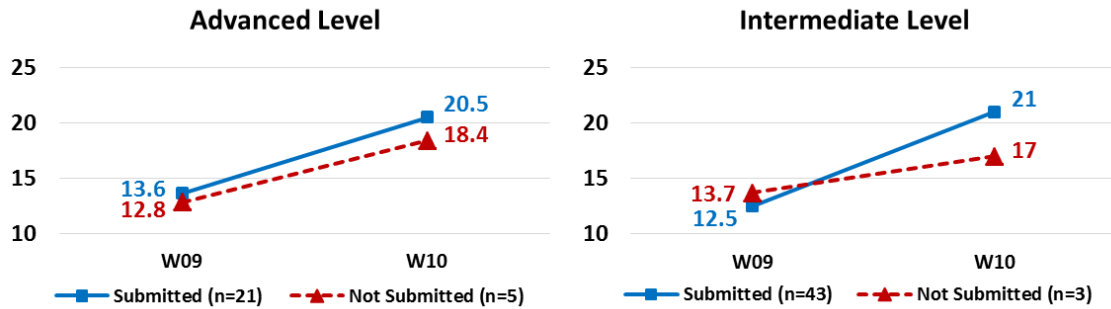
The factor that may explain the decline from W07 to W08 in both levels is that the content of the dictation answers uses *katakana* words, or foreign loan words in Japanese language which are of phonetic remembrance. In W6 and W7, PD sentences contained a few *katakana* words, such as ‘*kyabetsu sarada*’ for ‘cabbage salad’, and answers in English were also given full marks.

Although such *katakana* words sound differently to non-Japanese speakers (Bloch, 1950), they were excluded from W08 and on in order to avoid answers in English, which seemingly made the content more challenging in that it does not allow ‘guessed’ answers with phonetically similar words in English. This modification of the quiz content has likely resulted in the drop of the mean scores as answers required more accuracy in phonetic sequences in Japanese language.

As for the sharp rise from W09 to W10 in both levels (advanced = 6.0 pts, intermediate = 8.7 pts), various factors can be considered to account for this phenomenon. For example, the PD difficulty level may have been more

manageable in W10 than in W09; participants may have improved skills, such as their PD test-taking strategy; or they may have improved their incoming sound processing speed, which is the crucial skill in scoring PD. Furthermore, shadowing is considered to be effective for this improvement as it is considered to stimulate the shadowers' WM phonological loop. However, such miscellaneous assumptions are not convincing instruments for valid analysis since it is too complicated to measure the difficulty level between W09 and W10 PD sentences, and it is virtually impossible to measure participants' test-taking strategy improvement, especially after the data collection has been completed. However, there is one indicator that may as well be used as a treatment that can separate participants into two groups of control and experimental groups between W09 and W10 PD: the submission of W09 shadowing. As shown in Figure 10, W09 and W10 PD scores were compared with the variable of W09 shadowing submission. The number of participants is significantly larger than in the previous figure (Figure 9) since the filtering was applied only to W09 and W10 PD participants (advanced: $n = 26$, intermediate: $n = 46$), though the number of control groups (no W09 shadowing submission) in both levels are very small (advanced: $n = 5$, intermediate: $n = 3$).

Figure 10. W09 & W10 PD Mean Scores by W09 Shadowing Submission



The common trend in both the advanced and intermediate level is that all the groups increased the scores from W09 to W10, with the *Submitted* group (advanced = 6.9 pts, intermediate = 8.5 pts) scoring more points than the *Not-Submitted* group (advanced = 5.6 pts, intermediate = 3.3 pts). The advanced groups increased at a very similar rate, but the intermediate groups' rates of increase are remarkably different: the *Submitted* group scored lower than the *Not-Submitted* group in W09 by 1.2 pts, increasing in W10 by 4 pts. Considering the number of participants in the intermediate *Not-Submitted* group was very small ($n = 3$) and that the individual score can have a significant effect on the mean score, the improvement rate of 3.3 pts appears too small for the majority rate ($n = 43$) of 8.5 pts. As mentioned above, this comparison can only suggest statistically limited implications, but the results may as well highlight the possible effect of shadowing in the improvement of PD scores.

The general assumption of the experimental treatment is that the effect is evident in the experimental group while no effect is evident in the control group. Considering shadowing as a treatment in the above analysis, the increase of PD scores by the control groups confirms the involvement of other

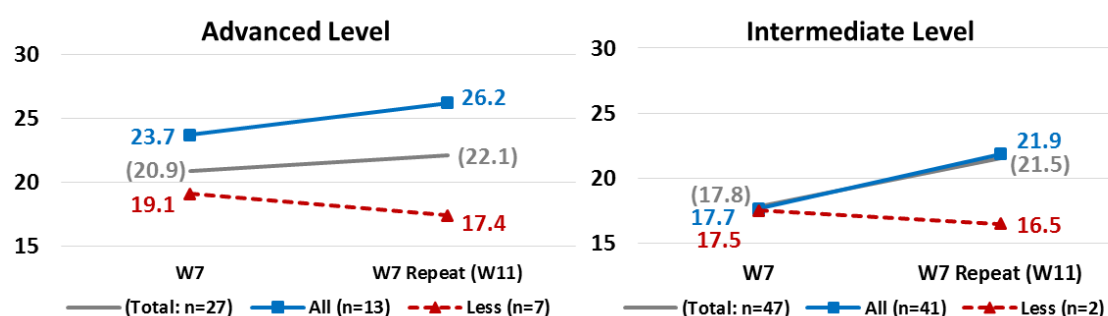
factors contributing to the increase of the scores, such as the aforementioned miscellaneous factors of the differences in PD sentences (e.g. sentence length, pronunciation sequence, familiarity of the names, clarity of the audio or even the speed of each sentence; though the audio was carefully adjusted by the instructor, human errors may occur). Therefore, the difficulty level of each week's PD may vary, and certain participants may prefer certain types of pronunciations, which is likely to occur between different native languages. For example, Korean speakers are known to have a tendency to pronounce [ts] as [ch], and [z] as the [j] sound; people that speak European languages that originated from Latin (e.g. French, Spanish, Portuguese, Italian, etc.) tend to have difficulty in distinguishing [h] sounds as they do not have this sound (JapanFoundation, 2009, p.24, p.29). In order to eliminate such differences between PD sentences, W07 PD was included in W11 PD. Therefore, W11 PD has 20 sentences with W11 sentences in odd numbers and W07 in even numbers. This way, the comparison of the participants' listening skills during the study period can be measured based on the same identical listening material. It is a reasonable question whether the W07 repeated results include LTM influence as participants may recognize the question and recall the answer. The chances are considered very slim as PD sentences consist of only random names and numbers, and each sentence was read for no longer than six seconds. Feedback only contained ticks and scores without releasing answers for students to review the results (PD feedback is more of a score card than learning material from the mistakes) because the repetition of the same week's PD in W11 was planned when PD was devised by the author. Also, there were six weeks apart between W07 and W11 (including a two-week recess between W07 and W08). Therefore, it is reasonable to assume that it

was virtually impossible for the participants to utilize their LTM from W07 PD in answering the W07 PD Repeat in W11.

3.3.2. Pinpoint Dictation between W7 and W7 Repeat in W11

Figure 11 shows the comparison of the mean scores of the W7 PD and W7 PD Repeated (conducted in W11) in the advanced (n = 27) and intermediate (n = 47) levels with a comparison between the submission frequency groups in the advanced level, *All* (n = 13) and *Fewer* (n = 7, submission average = 2.3 times) and the intermediate level, *All* (n = 41) and *Fewer* (n = 2, submission average = 2.5 times).

Figure 11. PD Mean Scores by W07 & W07 Repeat in W11



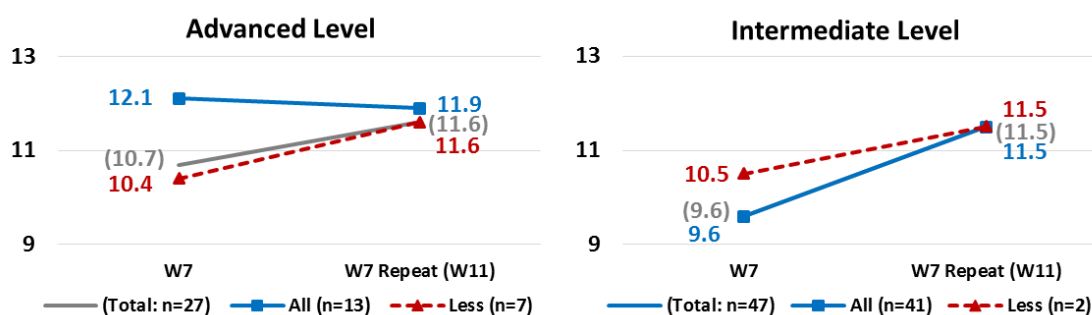
The common feature in both levels is the improvement in the mean scores from W07 PD to W07 Repeat (Advanced = 1.2 pts, Intermediate = 3.7 pts).

The intermediate level increased the mean score much more than the advanced level possibly due to the lower score in W07 (17.8 pts), which left more room for improvement than the advanced level which already had a high score in W07 (20.9 pts). It is safe to assume that this increase of the scores indicates the improvement of participants' listening skills due to the difficulty

of the test is the same. As discussed in the previous section, the possibility of the LTM effect in repeating the identical set of PD can be denied because not all of the groups improved their scores. In fact, there are two completely opposite trends between the groups: improved groups (*All*) and decreased groups (*Fewer*), and the trend is common to both levels. The advanced *All* group increased the score by 2.5 pts and the intermediate *All* group increased by 4.1 pts, whereas the advanced *Fewer* group decreased by 1.7 pts and the Intermediate *Fewer* group decreased by 1.0 pts. This phenomenon alone would be convincing enough to suggest that the amount of shadowing practice affects the improvement rate in sound recognition ability, if only the number of *Fewer* groups were equal to that of the counterpart.

Another characteristic of PD enables analysis of speed difference since the PD 10 sentences were read at different speeds: Part A (sentences 1–5 = 20 pts) was read slowly at the W07 shadowing equivalent (270 m/m) for the intermediate level and at the W09 equivalent for the advanced level (320 m/m), while Part B (sentences 6–10 = 20 pts) was read faster at the W09 equivalent for the intermediate level (320 m/m) and at the W11 equivalent (380 m/m) for the advanced level. This speed difference was fixed throughout the entire time PD was conducted. Although human error may have been involved in the speed adjustment, the first PD audio (W06) was always reviewed when recording the preceding weeks. Figure 12 shows the comparison of the mean scores of Part A (speed = slow) in W07PD and W07 Repeat PD by the shadowing submission frequency groups.

Figure 12. W07 & W07 Repeat PD Scores by Submission Frequency (Part A: Slow)

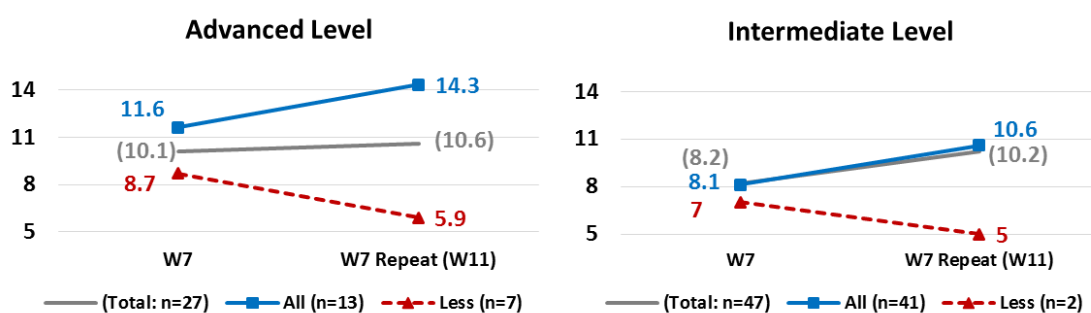


Part A also projects the overall increase of the mean scores in the advanced (0.9 pts) and intermediate level (1.9 pts). At a glance, both levels appear to evidence a similar pattern, i.e. the gaps in the scores between groups in W07 narrow down closely by 0.3 pts in the advance level; the same is also true in the intermediate level (11.5 pts) in W07 Repeat. However, the advanced *All* group remains flat with a slight decrease of 0.3 pts and the advanced *Fewer* group caught up with an increase of 1.2 pts. The intermediate *Fewer* group started 0.9 pts higher than the intermediate *All* group, but both increased and finished with the same score of 11.5 pts. This negative growth in the advanced *All* group may be explained as follows: those who scored high in Part A (speed = slow) of W7 PD were already capable of processing the audio speed, which may not have left much room for improvement. However, this ceiling effect alone does not account for the decline of the mean scores of the advanced *All* group. Another viewpoint that may explain this phenomenon is the type of the test format. As In'nami (2009) suggested, multiple-choice formats are easier than an open-ended format in that the former allows the test takers to apply a test-taking strategy while the latter does not. Test-taking strategy refers to techniques specific to the test format, such as elimination and guessing for the multiple-choice format, which allows the test taker to fill out all the questions

even though they may not be certain of the answers. This allows the possibility of scoring higher than test takers' actual proficiency as those 'guessed' answers have equal probability of getting the right answers (25% probability with four choices and 33% probability with three choices). JLPT follows a multiple-choice format with all the questions featuring either three or four possible answers. However, in an open-ended format such as PD where answers are to be written in blank boxes with no hints or choices to refer to, all answers must be produced as free writing. Under such a test format, there is no convenient test-taking strategy, such as elimination or guessing; test takers can only develop skills for time management, note-taking or even how to better concentrate on the audio recording so that they can hear the information to the best of their ability. Therefore, an open-ended test format often involves the possibility of scoring lower than the actual proficiency level of the test-takers, but it is considered virtually impossible to score higher than their actual proficiency level.

Figure 13 shows the comparison of the mean scores of Part B (speed = fast) in W07 PD and W07 Repeat PD by the shadowing submission.

Figure 13. W07 & W07 Repeat PD Scores by Submission Frequency (Part B: Fast)



A common trend is found in both the advanced and intermediate level: the *All* group improved by 2.7 pts (advanced *All*) and 2.5 pts (intermediate *All*), while the *Fewer* group declined by 2.8 pts (advanced *Fewer*) and 2.0 pts (intermediate *Fewer*). As for the advanced level, both groups scored similar in W07 (*All*= 11.6 pts, *Fewer*= 8.7 pts), but the outcome in W07 Repeat is completely different as evidenced by the gap of 8.4 pts between groups. A similar pattern is seen in the intermediate level as well: the groups start close with a 1.1 pts gap, resulting in the larger gap of 5.6 pts in W07 Repeat PD. To explain the declining scores in the *Fewer* groups, the previous discussion on the possible lower performance of the advanced *All* group in Part A (see Figure 12) can be applied to Part B as well. As the audio speed for Part B is much faster than that for Part A, it is reasonable to assume that there is greater possibility for the test takers to miss the dictation answer chunk in Part B. Therefore, the *Fewer* groups' results may likely indicate that the decline was due to their inadaptability to the speed of the audio sound, and therefore this can imply that there was no improvement in their listening skill, specifically in the area of sound recognition with fast-paced target language. On the other hand, the improvement in the scores by the *All* groups may be the strongest indicator that suggests improvement in their sound recognition of fast-paced spoken Japanese since the PD format does not allow multiple-choice-like fluke scoring. With all of the possible errors discussed above that may have resulted in the difference in the scores between the *All* and *Fewer* groups, it is reasonable to assume that the amount of shadowing practice time is one of the factors that made a difference in the improvement rate in the participants' fast-paced listening skill.

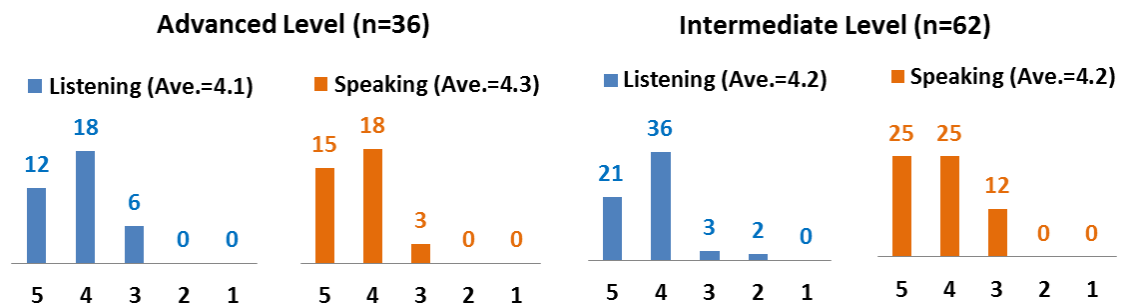
3.4. Anonymous Survey

A total of 36 advanced and 62 intermediate level participants filled out the anonymous survey at the end of the semester (see Appendix). The total number of each item may not be equal to the population as some answers included blank items.

3.4.1. Participants' perceived improvement in listening and speaking skills

Figure 14 summarizes the survey questionnaires regarding participants' perceived improvement in their listening (Q7) and speaking ability (Q8) by shadowing. The answers are based on 5-point Likert scales (5: strongly agree (*SA*), 4: agree (*A*), 3: not sure (*NS*), 2: disagree (*D*), 1: strongly disagree (*SD*)).

Figure 14. Participants' Perceived Improvement in Listening and Speaking



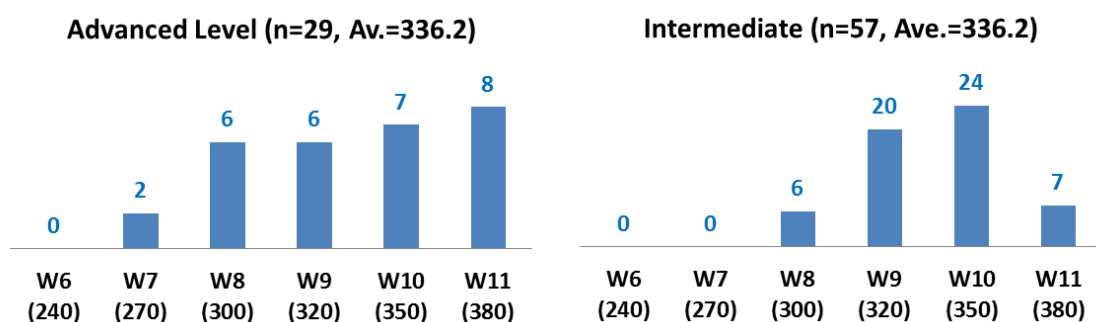
The common feature in all results is that the majority fall in either *SA* or *A*; there are almost no negative answers (i.e. *D* or *SD*) except for 2 *D*s in intermediate listening. These high average points in both levels indicate that the participants' overall shadowing experiences were positive in that they perceive shadowing improved their listening and speaking abilities. However, the Not Sure (*NS*) response may not indicate 'neutral' in that it is rejected as a favourable option by the participants who ticked this option. With this perspective of *NS* being not favoured together with *D*, 16.7% of the advanced and 8.3% of the intermediate participants perceive shadowing as effective in

improving their listening, and 8.1% of advanced and 19.4% of intermediate participants do not think shadowing improved their speaking ability. This is an interesting contrast of participants' perspectives regarding shadowing, i.e. more advanced participants felt an improvement in speaking than listening while intermediate participants felt an improvement in listening than speaking. This is the limit of the analysis with the set of above questions and the data alone. In future studies more elaborate questions should be added that ask in more detail about the aspects of improvement in each category of listening and speaking.

3.4.2. Participants' perceived effective speed of shadowing

Figure 15 shows the participants' perceived most effective shadowing audio speed. The numbers in brackets under the weeks are the mora per minute (m/m), which indicates the speed of the shadowing audio.

Figure 15. Participants' Perceived Most Effective Shadowing Speed

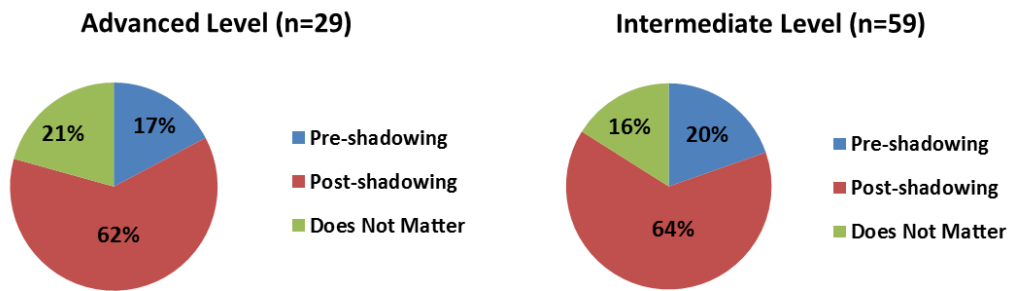


The survey results show different trends between the advanced and intermediate level participants where the former is spread almost evenly between W08 and W11 with a gradual increase towards (and peaking in) W11 (n = 8), while the latter projects a bell-curve like shape between W08 and W11,

peaking in W10 ($n = 24$). The difference in each trend can be explained by referring to the shadowing performance in Figure 3 in *Section 3.1.1*. The mean scores of shadowing in the advanced level start to drop gradually from W08 and on, suggesting that the shadowing speeds of those weeks were more challenging than in other weeks. However, the degree of the decline is rather moderate. Therefore, the gradual increase of participants who felt more effective towards W11 might have been because they felt a sense of success in managing the progressive shadowing speed towards the native-like speed in W11. In contrast, the shadowing performance in the intermediate level remains high until W10 before the sudden drop in W11; the majority of the participants at this level might have felt improvement up until W10 ($n = 24$), the final week of the ‘holding period’. Although the trend in the advanced and intermediate level is different, the average speed coincides at exactly the same speed of 336.2 m/m between W9 and W11. This may have two implications regarding the effective shadowing speed: firstly, for the shadowing model audio speed, approximately 336 m/m may be the common speed in both levels at which most participants feel effective; and secondly, advanced participants may potentially prefer a faster speed than 380 m/m since the number is increasing towards the end in W11, while careful consideration may be required when determining if intermediate participants should use the speed of 380 m/m.

3.4.3. Participants’ preference of script availability

Figure 16 shows participants’ preference for shadowing scripts availability both before and after the conduct of shadowing.

Figure 16. Participants' Preference for Shadowing Script Availability

The common feature in both the advanced and intermediate level is that the majority prefer a post-shadowing script (advanced = 62%, intermediate = 64%) while less prefer a pre-shadowing script (advanced = 17%, intermediate = 20%). Kadota and Tamai (2004) stated that the availability of a pre-shadowing script would enhance the understanding of the shadowing content; in fact, this script learning procedure is included in their recommended steps of shadowing conduct for beginners (Ibid, p. 62). However, the results of the present study indicate that the majority of the participants understand the primary objective of shadowing as being sound recognition: a pre-shadowing script would spoil the opportunity to stimulate the phonological loop for improving bottom-up processing.

3.4.4. Participants' perceived positive aspects of shadowing

In response to the open-ended question (Q15), 'What do you think are the positive aspect(s) of shadowing practice?', most participants responded with comments (advanced: 97%, n = 34; intermediate: 92%, n = 57). More than half of the participants in both levels commented on the improvement in listening skill they felt from their shadowing experiences (advanced: n = 21; intermediate: n = 33). More than one third of the participants in both levels

commented on the improvement in their speaking skill (advanced: $n = 15$; intermediate: $n = 20$), which reflects the results in the previous *Section 3.4.1*. that the majority of participants perceived that shadowing improved their aural and oral skills. Some of the participants' comments are as follows (the participant ID after the quotes uses the abbreviation *A* for the advanced level and *I* for the intermediate level):

It did improve my speaking/listening ability by quite a lot (A-7).

It significantly helps with speech practice (something I otherwise lack) and listening skills as you are actively listening (I-2).

I liked it because it was an activity that I had never done before (I-37).

As previously shown in Table 2 in *Chapter 2, Section 2.1.*, 30% of the participants in the advanced level and 94% in the intermediate level said that they had never done shadowing before this study (67% of advanced participants had experienced shadowing in the previous semester as a unit assessment, but with a much shorter script and a slower model audio speed than this study). As participant I-2 mentions, 'active listening' is the fundamental objective of shadowing practice as it forces shadowers to predominantly focus on sound recognition by making them simultaneously speak what they heard.

Therefore, as the following comments suggest, it is inevitable for shadowers to mobilize their primary attention to the accuracy of pronunciation:

It helped greatly to improve pronunciation and fast articulation as well as listening skills (A-23).

It got me practicing Japanese in a way I hadn't done before. It was challenging but helped my pronunciation and speed (Participant A-32).

Another common comment mentioned by participants is the improvement in processing speed (advanced: 37%, $n = 13$; intermediate: 34%, $n = 21$) and pronunciation (advanced: 23%, $n = 8$; intermediate: 21%, $n = 13$). This study was designed to assure that almost all the participants experience the difficulty in maintaining the speech speed with the shadowing model audio by applying a steady increase in the speed progression. It is natural to assume that the faster one tries to speak L2, the less accurate the language becomes, and this typically occurs in decreasing the accuracy of pronunciation when the speech speed exceeds the speakers' potential speed in the language. This applies to L1 as well, a perfect example of which is a tongue twister. If pronounced by syllable (or mora in Japanese) and taking sufficient time to pause, almost everyone can complete the sentence without any mistakes. However, as syllables are put together with fewer pauses and the speed gradually increases, eventually one hits a certain speed at which they will often make mistakes. However, it is crucial to note that the ceiling speed and the often-mistaken spots can be improved by repeated practice relatively more easily when one is unfamiliar with a certain tongue twister. To demonstrate this, the instructor introduced a few Japanese tongue twisters after introducing shadowing practice in the previous week in order to emphasize the

relationship between speech speed and accuracy as well as the importance of repetitious practice to complete the weekly shadowing along with the native model audio. The above comment by participant A-32 reveals his/her improvement after the ‘challenging’ experience which ‘helped pronunciation and speed’. Shadowing allows this challenge along with the ongoing spoken language as shadowers are forced to monitor how well they are keeping pace with the model audio by simultaneously speaking it; otherwise, like in repeating practice, the gap between the incoming and outgoing speed cannot be recognized. This recognition urges shadowers to engage in faster processing of the phonological loop of WM.

This speed progression is also beneficial as it provides participants with a certain benchmark of standard spoken speed:

It gives me an idea of how fast I should be aiming to speak at (I-14).

It not only improve[d] listening and speaking skills but as well as the flow and pronunciation of speaking Japanese (I-5).

My brain is now more comfortable for handle fast sentence during listening exam (I-30).

The assessment design of accuracy-only marking is considered beneficial in training bottom-up processing ability through shadowing because shadowers can mobilize their full attention to the sound recognition rather than sharing the limited brain resources with higher language processing, such as

comprehension of the text meaning. As the following participant noted, [It] helps focus on pronouncing the words without having to worry about understanding the meaning straight away' (A-2).

Participant A-2 may have developed a certain sense of anxiety if the assessment requirement involved comprehension to 'worry about' on top of the accuracy of sound reproduction. The shadowing task of simultaneous listening and speaking is already a complex, cognitively heavy-loaded task; therefore, it is considered best to limit the shadowers' burden especially at this stage of the 'introduction to shadowing' period. Two participants commented as follows:

It was fun and challenging. It made me dread it every week but at the same time look forward to it (A-15).

I quite liked doing [shadowing] as a weekly task, because it really challenges me (I-43).

It appears to be true that a gradual increase of resistance is necessary for both physical and cognitive development: for example, dumbbell weights for muscles and speed increase for shadowing in the context of this study. If the resistance is too heavy for the learners' current level, they would likely feel too much 'dread', a feeling which would lead to demotivation and may perhaps result in learners not completing the shadowing. Therefore, it is crucial to identify the $i+1$ zone, or Vygotsky's (1978) proximate zone of development, for shadowers to be able to find shadowing appropriately challenging. The above comments by participants A-15 and I-43 suggest the challenge being

manageable in a favourable manner. As the following comment illustrates, this sense of conquering the challenge would help motivate participants to engage in shadowing practice with an individual goal: ‘I was motivated to do the best that I could – it’s a feeling like “I have to beat this”’ (A-4).

As the shadowing assessment in this study was conducted as homework, it is entirely up to the students to choose the level of commitment to the task. Unlike creative tasks, such as composition or essay writing, shadowing shows the goal from the start. It is a positive motivator when the specific path to follow is shown (Edwin et al, 1981), but the challenge is how far they can push themselves to achieve the goal. The following comments express participants’ perceived distance towards the model audio that they aimed for:

Became more used to speaking Japanese at a much more natural speed – not taking easy by myself (A-35).

[It] encourages participants to leave their comfort zone (I-50).

Participant A-35 implies a sense of success after being able to manage a ‘much more natural speed’, otherwise he/she would have not reached the speed had it been ‘by [him/her]self’. Similarly, participant I-50 uses the phrase ‘comfort zone’ to imply that improvement requires discomfort, but as a positive aspect of shadowing as it leads to improvement.

Another perspective to develop motivation is to change the standpoint regarding the achieved point of view of oneself: ‘You sound pretty badass

speaking that fast' (I-52). Participant I-52 clearly used the word 'badass' in a positive manner (as this comment was written in response to the question about 'positive aspects' in the survey questionnaire), and this metaphoric expression conveys the feeling of the participant's desired goal, or 'possible self' (Markus & Nurius, 1986) such that 'badass' can be paraphrased as something like 'neat', 'cool' or 'awesome'. Lastly, the following comments represent the improvement through weekly shadowing:

When I looked back at the first week, I was surprised at how much my listening improved (I-28).

Being able to look back at slower speed and be able to understand every word said (I-48).

During W10 class after the shadowing feedback, the instructor played the first week's shadowing audio (speed: 240 m/m) in front of the class in order to compare to the speed they had just completed that week (speed: 350 m/m) in all the classes in both the advanced and intermediate levels. Students' reactions were remarkable as they listened to the same audio which they used to feel was much faster when they first heard it. It was apparent that everyone reached the same opinion: the shadowing speed of 240 m/m was slow. It is very difficult to 'see' the improvement in listening skills, especially in a short period of time such as six weeks. Objective listening test scores like pre-/post-shadowing and PD may be one of the resources to reference for improvement, but such numerical results are nowhere near as powerful as the sense that one can actually feel and recognize the difference in the comparison

of his/her perception of the same audio they used to feel was more difficult just a weeks ago. After playing the initial week's shadowing audio, many commented that they felt as if they had to stop and wait for the audio to catch up with it. This comment was common to all the classes in both levels.

3.4.5. Participants' perceived negative aspects of shadowing

This section discusses the open question (Q16), 'What do you think are the negative/difficult aspect(s) of shadowing practice?' Although there were fewer comments than in the previous section regarding the positive aspects, many participants responded with comments (advanced: 89%, n = 31; intermediate: 74%, n = 46). Common to both levels, nearly half of the participants commented that the speed increase towards the end surpassed their competence (advanced: n = 15; intermediate: n=14). The following comments illustrate this general sentiment.

[The last 2 weeks' speeds] 350/m and 380/m are too fast in my experience (A-9).

I think one negative aspect is the speed increase too fast over the six weeks. I felt I was not doing great towards the end and made more mistakes (I-31).

As shown in Figure 3 in the previous *Section 3.1.1.*, the overall shadowing score in the advanced level showed a decline from W09 (320 m/m) to W11 (380 m/m). Participant A-9's comment was fairly common among many participants. Similarly, the intermediate level's overall score declined sharply in W11 (380

m/m), and as participant I-31 commented, many participants actually made more mistakes in the final week. However, the speed progression of shadowing audio is an inevitable challenge, which was already announced at the beginning of the shadowing practice. In fact, as shown in Figure 13 in the previous *Section 3.4.2.*, 52% in the advanced level and 54% in the intermediate level perceived that the speeds of 350 m/m and 380 m/m were most effective for improvement. However, it is also true that the other participants perceived the opposite. Since the marking of shadowing is based on accuracy in this study, the common perception of shadowing seems to be based on deduction from the score of 100%. However, this may only apply to those who are competent enough to set their goal at 100%. Therefore, to those who feel the speed is overwhelming, an alternative view should be recommended, i.e. that building more accuracy leads to greater improvement. The first attempt may only be as accurate as 70%, but more practice will result in greater accuracy. If they reach 90%, this means that they have increased their accuracy by 20%, which is a meaningful stimulus for improvement in bottom-up processing.

The second most common comments in both levels (advanced: $n = 6$; intermediate: $n = 8$) were regarding the lack of comprehension of the script while shadowing:

[I was] too focused on trying to keep up with the speed to know the details what I am saying (A-3).

[It was] too fast and I find myself simply listening for each syllable completely abandoning the deciphering of meaning (I-43).

This study was designed not to provide a script before or during shadowing as the focus is on the improvement of the phonological loop in WM solely by sound recognition without visual or LTM assistance. Therefore, it was expected that participants would have more difficulties in comprehending the script than when using the pre-shadowing script method. In fact, there were notable mistakes in the parts with unfamiliar vocabulary, expressions or grammar usage as shadowers cannot refer to their LTM storage because such knowledge is non-existent in their mental lexicon. However, in some cases, mistakes were found where the shadower used different particles that were grammatically applicable in the context; however, the meaning changed as different particles were used. This is what the author believes to be the negative effect of top-down processing while shadowing as the shadower compensated for the unheard parts by replacing them with different grammar functions in order to complete the sentence. This may potentially lead to misunderstanding as the output is manipulated because of the lack of sound recognition competence. As shown in Figure 14 in the previous *Section 3.4.3.*, the fact that the majority of participants' (advanced = 62%; intermediate = 64%) preferred a post-shadowing script indicates that the objective of the shadowing exercise for this study places the development of bottom-up processing as a common goal.

Therefore, the extra difficulty due to the post-shadowing script method is already embedded in this study design. The participants commented on this difficulty as follows:

We can't read script when we do the shadowing was a bit difficult (A-22).

[The] more distracting the faster it gets since there are no scripts (I-36).

Unknown words if any, [were] hard to distinguish in contrast to text (I-24).

The combination of no script, unfamiliar texts and/or excessive speed progression may have a negative effect on participants' engagement in shadowing practice.

The following comments focus on administrative matters rather than shadowing practice itself:

For me it was due when I had work so it was hard for me to find practice time and submit online (A-2).

Not much time to do, and difficult to find a place to record that isn't in front of family or uni[versity] friends (within the time we had to do it) (I-9).

In order to maintain the weekly cycle of conducting shadowing and feedback return, the submission due date was set right after the weekend (e.g. shadowing model audio is released after the tutorial class on Wednesday and the audio submission is due on Monday so that the instructor can mark all the shadowing audio submissions and complete data input by the following Wednesday). Maintaining a weekly cycle was considered important because the feedback should be returned while the shadowing experience is still fresh

among students. This issue of participants not having enough time for shadowing practice can be solved by releasing the model audio earlier than the tutorial class since the model audio is accessible online.

Chapter 4. Conclusion: Implications and Limitations

This study investigated the effect of shadowing practice through multiple sets of testing materials using descriptive statistical analysis and participants' perceived response to their shadowing experiences. As a summary of this study, this chapter outlines the key findings through answering the research questions presented in *Section 1.2*. Then, this chapter suggests implications of shadowing implementation in Japanese language classes, followed by implications for further study.

4.1. Summary of key findings

4.1.1. Research Question (1)

'What is the most effective shadowing speed range?'

Based on the shadowing performance data, this study revealed the most effective shadowing speed from the observation of the participants' weekly shadowing performance (see Figure 3), showing a common trend in both levels – 'holding period' and 'dropping period' – that the end of the 'holding period' speed at 320 m/m in the advanced level and 350 m/m in the intermediate level are considered most effective. Although the advanced level survey results suggest otherwise, i.e. more participants prefer 350 m/m and 380 m/m, the fact that more than half of the participants' scores dropped sharply after 320 m/m (see Figure 4) suggest that 320 m/m should be considered as the most effective speed for the advanced level. The intermediate survey results, meanwhile, match the performance results which indicated that 350 m/m is the most effective speed.

4.1.2. Research Question (2)

'Are there any differences between advanced and intermediate level in the

relationship between shadowing and listening ability improvement?

The answer to this research question is ‘no’ in that both advanced and intermediate level participants behaved in a very similar manner in the comparative analysis in pre-/post-shadowing (see Figure 8) and pinpoint dictation (see Figure 10 and 11). However, there was a remarkable difference in the shadowing submission rate in that the majority of the participants submitted all six weeks in the intermediate level (69%) while well below half (37%) submitted all six weeks in the advanced level.

4.1.3. Research Question (3)

‘Do participants perceive shadowing as effective practice in improving listening ability?’

The survey results in both the advanced and intermediate level report fairly positive feedback in that the majority (advanced = 83%, intermediate = 92%) agreed that they ‘think shadowing improves [their] Japanese listening ability’ (survey Q7). In response to the open question (Q15) which asked participants to comment on the ‘positive aspect(s) of shadowing practice’, the most frequent response was their improvement in listening ability (advanced = 61.8%; intermediate: n= 57.9%). The above results affirm the research question (3) that the participants agree that shadowing is an effective practice for listening ability improvement.

4.2. Implications for implementing shadowing practice in Japanese classes

This section discusses the implication for implementing shadowing practice in Japanese language classes based on the experiences from the current study.

The discussion also refers to some comments in response to one of the anonymous survey questions (Q17) for the suggestions for improvement in order to reflect participants' shadowing experiences.

(1) Shadowing materials

This study used self-recorded shadowing model audio (recorded by the male instructor) in order to be able to adjust the speed increase at exactly the same rates and speeds, and the same speed was used for both the advanced and intermediate levels each week. This adjustability of the shadowing model audio restricted the freedom of using other listening audio materials that may be suitable for the study because it is virtually impossible to find such an audio set with a specific speech speed ranging between 240 m/m and 380 m/m in different levels. Furthermore, the practicalities associated with creating such audio materials would make it difficult to request other teachers of native Japanese speakers to help with the recording.

One participant commented on the lack of a female speaker in the recording, observing that 'It's better to have two gender speakers' version because I'm a female student. When I try to follow the male speaker's Japanese, I feel kind of difficult' (I-45). The variety of shadowing materials would enhance learners' engagement in the practice, and thus it is recommended to utilize shadowing model audio recorded by speakers from different genders if available. As participant I-45 points out, female learners may have found it less beneficial

4.2. Implications for Implementing Shadowing Practice in Japanese Classes

to shadow after a male voice, especially for those female learners with a high-pitched voice as the gap between the input (male voice) and output (female voice) is large.

The speed progression of the shadowing model audio is based on the research question of this study, so it may not be necessary to prepare such a wide range of shadowing audio speeds. The participants commented on the audio speeds as follows:

Keep the speed between 300/m–380/m (I-46).

The maximum speed can be 350/m. When I was doing 380/m, I was quite challenging and I wasn't sure what I was doing (I-42).

As discussed in the previous *Section 3.1.2.*, the 'holding period' differs in the advanced (320 m/m) and intermediate levels (350 m/m), but in both levels the majority of the participants were able to manage the first two weeks (240 m/m and 270 m/m). As participant I-46 suggested, the speed of 300 m/m can be used as shadowing introduction followed by 320 m/m and 350 m/m as actual shadowing practice. However, as participant I-42 pointed out, the speed of 380 m/m should require careful attention depending on the competence level of the participants. Therefore, it is recommended to check the accuracy of learners' performance every time shadowing is conducted since doing so will provide the instructor with a firm grasp of the matching level between the model audio and the participants' current shadowing level.

4.2. Implications for Implementing Shadowing Practice in Japanese Classes

As for the availability of the shadowing script, it is strongly recommended to provide a written script so that the learners can review their mistakes and to help ensure the clarity of pronunciation and comprehension of the context. Whether instructors use the script before or after shadowing would depend on the purpose of the exercise: post-shadowing is recommended for accuracy-focus and pre-shadowing for comprehension-focus. Another script timing known as ‘parallel-shadowing’ can be used for pre-shadowing activity when introducing shadowing activity for the first time (Kadota, 2007, p.27).

For more advanced learners, authentic materials, such as media reports (e.g. news, weather forecast, etc.), movies and anime, can be included in the list of shadowing materials as they are likely to attract learners’ curiosity and interest in Japanese study (Tanaka, 2014).

(2) Administration of shadowing

This study conducted shadowing solely as a homework assignment due to the constraints in the language classroom setting where audio equipment was not available for individual students to use a headphone set for shadowing practice. However, if Computer Assisted Language Learning (CALL) classrooms are available, or if it is possible to provide portable CD players for each student as in Takanishi and Ueda’s (2011) study, in-class shadowing practice on top of homework assignments is highly recommended for several reasons: firstly, it assures certain shadowing practice time for all the students, which homework alone does not guarantee; secondly, shadowing all together at the same time would not only improve the classroom atmosphere, but also enhance active engagement in shadowing activity as opposed to solitary

4.2. *Implications for Implementing Shadowing Practice in Japanese Classes*

practice at home; and lastly, it would help motivate more homework practice if the demonstration of shadowing is conducted in the following week. As some students may feel anxious about demonstrating shadowing in front of the entire class, the demonstration can be arranged as a smaller group activity (e.g. shadowing a paragraph by taking turns), or the presentation may be conducted on a voluntary basis. In any case, sharing the shadowing experience among peers is considered crucial in students' engagement.

Another important aspect in shadowing conduct is the feedback based on individual performance. Among various aspects in spoken language, this study referred only to the accuracy of shadowing performance in order to measure the accuracy of the sound recognition. The benefit of this accuracy-focused marking was that the researcher was able to collect very detailed data related to the students' shadowing performance, which was crucial to this study.

However, the downside of this marking method is that it was a time-consuming and often exhausting task to mark over 100 shadowing submissions each week. Furthermore, it had to be completed in a couple of days in order to return the feedback to students during a weekly cycle.

Therefore, if thoroughly accurate data is not necessary, there are alternative methods to be considered for marking the accuracy; for example, Tamai's (2005) check-point method, which involves marking only every other five words like in a cloze test (as cited in Kadota, 2007, p. 238). This check-point method is designed for marking English sentences, but it can be arranged for marking Japanese sentences by breaking the sentence into *bunsetsu*, or minimum meaning-making blocks in a sentence. If the shadowing performance does not reflect the grades, it may be best to have students mark

4.2. Implications for Implementing Shadowing Practice in Japanese Classes

their own performance as this would encourage more active noticing of their own mistakes rather than passively receiving the feedback marked by the instructor. Besides accuracy, Kadota (2007) also suggested marking prosodic features of shadowing performance, such as rhythm and intonation. However, this may involve the marker's subjective perception of 'good pronunciation' (Ishihara et al., 2011); it may be considered fair feedback as long as consistency is maintained during the study period.

4.3. Implications for further study

The primary limitation of this study is the lack of the experimental group to compare the effect of shadowing. Although between-groups comparison was made based on the different amounts of shadowing practice frequency, the implication of the results is not statistically valid, particularly since the size of the 'control group' is significantly small. This was due to the constraints of the shadowing as part of the unit assessment which did not allow the shadowing assessment as an option. Therefore, if the shadowing is not compulsory for the unit assessment, it may be best to recruit volunteer participants who are willing to conduct shadowing as an extracurricular activity for the experimental group while the rest of the unit students participate in the control group. If a joint project with other universities is possible (Kuramoto et al, 2006), a large number of participants can be expected.

In terms of the target language proficiency levels, this study conducted shadowing during the first semester in intermediate and advanced level Japanese language units, whose end-of-the-academic-year achievement levels are aimed at JLPT N3 and N2, respectively. Therefore, at the time of the study period, the participants' current Japanese proficiency levels were likely between beginner and intermediate levels. The commonality between these levels was that the manageable shadowing speed was between 320 m/m and 350 m/m, with more advanced level participants preferring 380 m/m. Therefore, it is assumed reasonable that 400 m/m (native-like speed) or even faster speed can be used for expert learners of Japanese whose proficiency level is at JLPT N1 level. On the other hand, the fact that the intermediate level learners in this study have managed a speed this close to the native-like

4.3. Implications for Further Study

speed may well suggest a possible study targeting pure beginner level as there is almost no shadowing study conducted at this level. The main reason for this can be assumed that the learners' vocabulary and grammar knowledge is quite limited and no sufficient shadowing audio materials are available. However, the flexible audio production method presented in this study would allow production of various shadowing model audio at suitable speed progression rates specific to a pure beginner level with a focus on sound recognition if their language knowledge is too limited.

References

- Baddeley, A. (1992). Working Memory. *Science*, 255(5044), 556-559.
- Baddeley, A. (2000). The Episodic Buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417-423.
- Baddeley, A. (2002). Is Working Memory Still Working? *European Psychologist*, 7(2), 85-97.
- Bloch, B. (1950). Studies in Colloquial Japanese IV: Phonemics. *Journal of Linguistic Society of America*, 26(1), 86-125.
- Broadbent, D. E. (1952). Speaking and Listening Simultaneously. *Journal of Experimental Psychology*, 43(4), 267-273.
- Chang, J. (2008). The Result and Prospection in Applying Shadowing Training Into Japanese Conversation Class (2) -Using Wenzao Ursuline College of Languages, Department of Japanese as Model-. *Taiwan Application Japanese Association*, 5, 415-430.
- Darro, V., & Fabbro, F. (1994). Verbal Memory During Simultaneous Interpretation: Effects of Phonological Interference. *Applied Linguistics*, 15(4), 365-381.
- Dewey, D., Brown, J., & Eggett, D. (2012). Japanese Language Proficiency, Social Networking, and Language Use during Study Abroad: Learners' Perspectives. *Canadian Modern Language Review*, 68(2), 111-137.
- Edwin, L., Kyaryll, S., Lise, S., & Gary, L. (1981). Goal Setting and Task Performance: 1969-1980. *Psychological Bulletin*, 90(1), 125-152.
- Gregersen, T., & Horwitz, E. (2002). Language Learning and Perfectionism: Anxious and Non-Anxious Language Learner's Reactions to Their Own Oral Performance. *The Modern Language Journal*, 86(4), 562-570.
- Hamada, Y. (2012). An Effective Way to Improve Listening Skills through Shadowing. *The Language Teacher*, 36(1), 3-10.
- Hamada, Y. (2014). The Effectiveness of Pre- and Post-Shadowing in Improving Listening Comprehension Skills. *The Language Teacher*, 38(1), 3-10.

- Horwitz, E., Horwitz, M., & Cope, J. (1986). Foreign Language Classroom Anxiety. *The Modern Language Journal*, 70(2), 125-132.
- In'nami, Y. (2009). A meta-analysis of test format effects on reading and listening test performance: Focus on multiple-choice and open-ended questions *Language Testing*, 26(2), 219-244.
- Ishihara, S., Tsurutani, C., & Tsukada, K. (2011). What Constitutes "Good Pronunciation" from L2 Japanese Learners' and Native Speakers' Perspectives? A Perception Study. *Electronic Journal of Foreign Language Teaching*, 8(1), 277-290.
- Iwashita, M. (2008). The effectiveness of Shadowing Training on Japanese Leareners - Based on monthe-long longitudinal study - *Bulletin of the Graduate School of Education, Hiroshima University*, 25(7), 219-228.
- JapanFoundation. (2009). *Onsei wo oshieru (Teachin Pronunciation)*. Tokyo: Hitsuji Books Ltd.
- JapanFoundation. (2012). Japanese-Language Proficiency Test. from <http://www.jlpt.jp/e/about/points.html>
- Kadota, S. (2007). *Shadowing to Ondoku no Kagaku (Shadowing and the Science of Oral Reading)*. Tokyo: CosmoPier Publishing Company.
- Kadota, S., & Tamai, K. (2004). *Ketteiban, Eigo Shadowing (The Definitive Edition English Shadowing)*. Tokyo: CosmoPier Publishing Company.
- Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of Working Memory in Children with ADHD. *Journal of Clinical and Experimental Neuropsychology*, 24(6), 781-791.
- Kobayashi, M. (2002). Cloze Test Revisited: Exploring Item Characteristics with Special Attention to Scoring Methods. *The Modern Language Journal*, 86(4), 571-586.
- Kondo, R. (2012). Effects of Providing scripts in Shadowing Practice on Learner's Japanese-Language Acquisition. *Jaapanese Studies Journal*, 28(2), 43-57.

- Krashen, S. (1985). *The Input Hypothesis: Issues and implications*. London: Longman.
- Kuramoto, A., Shiki, O., Nishida, H., & Ito, K. (2006). *Seeking Effective Instructions for Reading: The impact of Shadowing, Parallel reading, and Reading Aloud Tasks*. Paper presented at the A Symposium Held at an International Joint Forum on 1st- and 2nd- Language Sciences Honolulu: University of Hawaii at Manoa.
- Kurata, K. (2007). A Fundamental Study on the Cognitive Mechanism of Shadowing in Japanese - From the view point of oral reproduction, memory span and sentence structure -. *Bulletin of the Graduate School of Education, Hiroshima University*, 2(56), 259-265.
- Kyo, H. (2012). Nihongo Gakushu ni Mirareru Shadowing no Kouka. *Meikai Nihongo*, 17, 103-111.
- Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.
- Lambert, S. (1992). Shadowing. *Meta*, 37(2), 263-273.
- Maeda, A. (1998). Nichijo Kaiwa ni Okeru Joshi no Shoryaku (Particle Omission in Daily Conversation). *Bulletin of the International Student Center, Nagasaki University*, 6, 43-70.
- Markus, H., & Nurius, P. (1986). Possible Selves. *American Psychologist*, 41(9), 954-969.
- Mochizuki, M. (2006). On the Application of Shadowing Method to Japanese Language Education : With Special Reference to Relevance of Learners' Japanese Competence to Their Self-Assessment of the Effects of Shadowing. *Kansai University Audio-Visual Education*, 29, 37-53.
- Nakanishi, T., & Ueda, A. (2011). Extensive reading and the effect of shadowing. *Reading in Foreign Language*, 23(1), 1-16.
- Nakayama, T. (2011). Weak Forms in Shadowing: How can Japanese EFL Learners Perform Better in Shadowing Tasks? *The Society of English Studies*, 41, 17-31.

- Saito, Y., Nagasawa, Y., & Ishikawa, S. (2011). Effective Instruction of Shadowing Using a Movie. *JALT Conference Proceedings, JALT 2010*, 139-148.
- Shiki, O., Mori, Y., Kadota, S., & Yoshida, S. (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*, 21, 81-90.
- Spiller, E., Daro, V., BFabbro, F., & Bosatra, A. (1990). Audio-Phonological and Neuropsychological Aspects of Simultaneous Interpretation: Role of Auditory Shadowing. *Scand Audiol*, 19, 81-87.
- Takahashi, E., Fukuda, N., Iwashita, M., & Sakoda, K. (2010). A Study of Shadowing for Advanced Language Learners - Learners' noticing and instructors' support - *Bulletin of the Graduate School of Education, Hiroshima University*, 2(59), 299-308.
- Tamai, K. (2002). On The Effects of Shadowing on Listening Comprehension - Keynote Lecture at the 3rd Annual Conference of JAIS. *Interpretation Studies*(2), 178-192.
- Tanaka, H. (2014). Motivational Intervention and Satisfying Learners' Need for Competence. *JALT Journal*, 36(1), 91-124.
- Vygotsky, L. S. (1978). Interaction Between Learning and Development. In M. Gauvain & M. Cole (Eds.), *Reading on the Development of Children* (pp. 29-36). New York: W.H. Freeman and Company.
- Watanabe, A., Oba, R., Shimizu, T., Sugiyama, M., Nohara, Y., & Sakuta, N. (2013). *Japanese Language Proficiency Test N2 - Complete Mock Exams*. Tokyo: J Research Publisher.
- Yang, J. (2008). Motivational Orientations and Selected Learner Variables of East Asian Language Learners in the United States. *Foreign Language Annals*, 36(1), 44-56.

Appendices

Appendix A

Anonymous Survey

Q1. Class ☐ JPS201 ☐ JPS301

Q2. Gender ☐ male ☐ female

Q3. I am ☐ domestic student ☐ international student

Native language:

Other language(s):

Q4. Period of studying Japanese (formal education)

Q5. Already passed JLPT (Japanese Language Proficiency Test)

☐ N5 ☐ N4 ☐ N3 ☐ N2 ☐ N1

☐ N/A

Q6. Do you have any previous experience of shadowing before this unit?

☐ Yes ☐ No

If Yes, please briefly explain your experience of shadowing (e.g., class, institution, etc.)

Q7. Do you think shadowing improves your Japanese listening ability?

☐ Strongly agree ☐ Agree ☐ Not sure ☐ Disagree

☐ Strongly Disagree

Q8. Do you think shadowing improves your Japanese speaking ability?

☐ Strongly agree ☐ Agree ☐ Not sure ☐ Disagree

☐ Strongly Disagree

Q9. Do you think shadowing motivates you more in studying Japanese?

☐ Strongly agree ☐ Agree ☐ Not sure ☐ Disagree

☐ Strongly Disagree

Q10. How much do you think you have improved your shadowing skill?

☐ Very little ☐ Little ☐ Somewhat ☐ Much

☐ Significantly

Q11. How many times did you practice shadowing before submission (including re-recording)?

Minimum: ☐ 0 ☐ 1-2 ☐ 3-5 ☐ 6-10 ☐ more than 10

Maximum: ☐ 1-2 ☐ 3-5 ☐ 6-10 ☐ more than 10

Usually: ☐ 1-2 ☐ 3-5 ☐ 6-10 ☐ more than 10

Q12. Do you think the shadowing materials were

☐ Very easy ☐ Easy ☐ Moderate ☐ Difficult

☐ Very difficult

Q13. Which do you think the most effective for the model audio speed for your

improvement?

- ☐ W05: 240/m ☐ W06: 270/m ☐ W07: 300/m ☐ W08: 320/m
☐ W09: 350/m ☐ W10: 380/m

Q14. Do you think the script should be available before the shadowing for more effective outcome?

- ☐ Yes ☐ No ☐ It does not matter

Q15. What do you think positive aspect(s) of shadowing practice?

Q16. What do you think negative/difficult aspect(s) of shadowing practice?

Q17. How do you think shadowing practice can be improved?

Q18. Any comments/suggestions?

- ☐ You may quote my response above in your study.
☐ You may NOT quote my response above in your study.

Appendix B
Ethics Approval letter

Ethics Application Ref: (5201400169) - Final Approval

Dear Dr Svetanant,

Re: ('The effect of Shadowing Practice: The case study of Japanese language learners in Australian universities')

Thank you for your recent correspondence. Your response has addressed the issues raised by the Faculty of Arts Human Research Ethics Committee. Approval of the above application has been granted, effective (6/03/2014). This email constitutes ethical approval only.

If you intend to conduct research out of Australia you may require extra insurance and/or local ethics approval. Please contact Maggie Feng, Tax and Insurance Officer from OFS Business Services, on x1683 to advise further.

This research meets the requirements of the National Statement on Ethical Conduct in Human Research (2007). The National Statement is available at the following web site:

<http://www.nhmrc.gov.au/files/nhmrc/publications/attachments/e72.pdf>.

The following personnel are authorised to conduct this research:

Dr Chavalin Svetanant
Hideki Sumiyoshi

NB. STUDENTS: IT IS YOUR RESPONSIBILITY TO KEEP A COPY OF THIS APPROVAL EMAIL TO SUBMIT WITH YOUR THESIS.

Please note the following standard requirements of approval:

1. The approval of this project is conditional upon your continuing compliance with the National Statement on Ethical Conduct in Human

Research (2007).

2. Approval will be for a period of five (5) years subject to the provision of annual reports.

Progress Report 1 Due: 6th March 2015

Progress Report 2 Due: 6th March 2016

Progress Report 3 Due: 6th March 2017

Progress Report 4 Due: 6th March 2018

Final Report Due: 6th March 2019

NB: If you complete the work earlier than you had planned you must submit a Final Report as soon as the work is completed. If the project has been discontinued or not commenced for any reason, you are also required to submit a Final Report for the project.

Progress reports and Final Reports are available at the following website:

http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/human_research_ethics/forms

3. If the project has run for more than five (5) years you cannot renew approval for the project. You will need to complete and submit a Final Report and submit a new application for the project. (The five year limit on renewal of approvals allows the Committee to fully re-review research in an environment where legislation, guidelines and requirements are continually changing, for example, new child protection and privacy laws).

4. All amendments to the project must be reviewed and approved by the Committee before implementation. Please complete and submit a Request for Amendment Form available at the following website:

http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/human_research_ethics/forms

5. Please notify the Committee immediately in the event of any adverse effects on participants or of any unforeseen events that affect the continued ethical acceptability of the project.

6. At all times you are responsible for the ethical conduct of your research in accordance with the guidelines established by the University. This information is available at the following websites:

<http://www.mq.edu.au/policy/>

[http://www.research.mq.edu.au/for/researchers/how to obtain ethics approval/human research ethics/policy](http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/human_research_ethics/policy)

If you will be applying for or have applied for internal or external funding for the above project it is your responsibility to provide the Macquarie University's Research Grants Management Assistant with a copy of this email as soon as possible. Internal and External funding agencies will not be informed that you have approval for your project and funds will not be released until the Research Grants Management Assistant has received a copy of this email.

If you need to provide a hard copy letter of approval to an external organisation as evidence that you have approval, please do not hesitate to contact the Faculty of Arts Research Office at ArtsRO@mq.edu.au

Please retain a copy of this email as this is your official notification of ethics approval.

Yours sincerely

Dr Mianna Lotz
Chair, Faculty of Arts Human Research Ethics Committee
Level 7, W6A Building
Macquarie University
Balaclava Rd
NSW 2109 Australia
Mianna.Lotz@mq.edu.au