Spring 2015

# Effects of online repetition practice on promoting mora awareness: Focusing on vowel length

Natsumi Suzuki
*Purdue University*

*Graduate School Form 30*
*Updated 1/15/2015*

**PURDUE UNIVERSITY**
**GRADUATE SCHOOL**
**Thesis/Dissertation Acceptance**

This is to certify that the thesis/dissertation prepared

By Natsumi Suzuki

Entitled

EFFECTS OF ONLINE REPETITION PRACTICE ON PROMOTING MORA AWARENESS: FOCUSING ON VOWEL LENGTH

For the degree of Master of Arts

Is approved by the final examining committee:

Atsushi Fukada
Chair

Jessica Sturm

Mariko Moroishi Wei

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): Atsushi Fukada

Approved by: Madeleine Henry 4/16/2015

Head of the Departmental Graduate Program Date

# EFFECTS OF ONLINE REPETITION PRACTICE ON PROMOTING MORA AWARENESS: FOCUSING ON VOWEL LENGTH

A Thesis

Submitted to the Faculty

of

Purdue University

by

Natsumi Suzuki

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Arts

May 2015

Purdue University

West Lafayette, Indiana

# ACKNOWLEDGEMENT

I would like to offer my deepest appreciation to the chair of my committee member and my advisor, Professor Atsushi Fukada for his continuous help, support, and guidance since I arrived at Purdue University. This research would not have been possible without his assistance. Thank you for always being there when I needed help, and for being a great advisor.

I would like to extend my appreciation to my committee members, Professor Mariko Moroishi Wei and Jessica Sturm for the helpful feedback and giving me inspiration on the topic I wish to pursue as a researcher.

I also thank the Japanese 102 instructors for letting me collect the data from their classes, especially the coordinator and my good friend, Mayu Miyamoto for giving me support throughout the process of completing my thesis. I also thank the students who participated in my research as they are the motivation of my research.

I would also like to thank Reiko Akiyama, Yukie Kawakami, Ryota Kondo, and Sayuri Ogiuchi for helping me rate the data, and Namiko Uchida for helping me with my experiment.

I wish to thank my proofreader, Elena Semos for taking her time to read countless papers throughout my career at Purdue University, providing me with thoughtful and helpful feedback.

I am grateful for all the staff at the School of Languages and Cultures for giving me an opportunity to continuously grow, both as a researcher and as a language instructor.

And last but not least I would like to thank my parents for always supporting me with every decision I make, and Lukas Brenner for his unconditional support and encouragement throughout every step in completing the Master's Program. Thank you for always being there for me.

TABLE OF CONTENTS

Page

Page

# LIST OF TABLES

Table Page

# LIST OF FIGURES

# ABSTRACT

Suzuki, Natsumi. M.A., Purdue University, May 2015. Effects of Online Repetition Practice on Promoting Mora Awareness: Focusing on Vowel Length. Major Professor: Atsushi Fukada.

Japanese learners often have a difficult time in acquiring both accurate perception and production of Japanese special moras. Since errors on Japanese special moras can change the meaning of words, it is important to promote mora awareness in an early stage of the learners' language learning process. However, pronunciation practice including that on special moras is often dismissed in the classroom due to time constraints and other reasons. That is why an online system, available outside the class time, could be beneficial in promoting students' awareness of special moras. This study investigates the effectiveness of online listen-and-repeat practice to promote learners' awareness of special moras, focusing on long vowel production. The effectiveness of online practice is compared to when the practice is done in a classroom setting. In this research, 20 words containing long vowels and five fillers were selected. The subjects were 27 first-year Japanese students and were randomly divided in three groups; 1) Online Audio-Only group, 2) Online Audio-Visual group, and 3) In-Class Choral Repetition group (called In-Class from this point onwards). All three groups received a pretest, listen-and-repeat practice, a posttest, and a delayed posttest.

While the two training groups were given listen-and-repeat online practice with and without visual cues, the In-Class group received a classroom-style presentation and practice of the same set of words. The visual cue in this research modeled a "karaoke" system, where each mora in the word appears at the same time as that mora is being pronounced, and this system was designed so that learners will notice that the length of a long vowel is the same as any other mora.

Results indicated significant improvement on vowel duration accuracy for both groups that received listen-and-repeat practice in an online self-study environment, while participants who practiced in a classroom setting did not. However, although the descriptive statistics showed the greatest improvement for those who received the Audio-Visual treatment, statistical analysis did not show significant difference in the participants' improvement among the treatment types. Therefore, future studies are needed to further investigate the effectiveness of different input modalities that could be used with listen-and-repeat practice.

# CHAPTER 1: INTRODUCTION

## Motivation for This Study

In the field of Second Language Acquisition (SLA), pronunciation is usually given the least amount of classroom instruction, and scholars have also studied words and grammar for a much longer time compared to pronunciation. This is said to be the same in Japanese as a foreign language education as well.

There are many reasons why pronunciation has often been neglected in the classroom, which includes the overall trend of SLA and the popularity with the methods being used to teach a second language (L2), the lack of teaching and testing materials on pronunciation, and the lack of teacher training.

L2 classroom instruction in recent years have been based on the communicative method, which puts emphasis on doing creative work. The advocates of the popular communicative method encourages instructors to avoid doing repetitious drills because they believe that language acquisition does not occur from forming habits (Iba, 2008). However, pronunciation practice frequently requires repetition practice. Thus, the avoidance of repetition practice may have resulted in a lack of pronunciation practice.

Oguma (2002) states that in Japanese as a foreign language education, the learner's speaking ability is normally not tested, and thus, this results in limited amount of pronunciation training, because instructors tend to focus teaching on factors that

learners will be tested on, such as grammar and vocabulary items. Because of this, learners who have achieved an advanced level in terms of grammar and vocabulary capacity often do not show a parallel ability in their pronunciation skills. In relation to Oguma's statement, Toda (2008) states that for grammar and vocabulary items, there is usually a set goal that determines how many grammatical items and words learners should know at a specific level, but pronunciation normally does not have this set goal. Toda (2008) claims that one of the reasons why pronunciation lacks this set goal is that materials to teach pronunciation are lacking. It is easier for instructors to follow a textbook that explains how much progress learners should make in each chapter, but textbooks or materials in general are lacking in the field of pronunciation. This is problematic, because even if learners reach advanced level in terms of grammatical and vocabulary knowledge, it will be useless if they cannot communicate due to the hindrance of poor pronunciation.

Significant communicative problems could be created from various pronunciation errors, including learners having difficulty with pronouncing the phoneme that does not exist in their first language (L1). Empirical studies show that learners of Japanese have difficulty in acquiring non-native phonemic contrasts, namely the "special moras." Japanese words can be divided into a small unit called a mora, which is different from many other languages in the world (Toda, 2003). Within the moras, the "special moras" are said to be especially difficult for learners to acquire because they are non-syllabic, which many learners are probably not used to. It is important for learners to acquire these special moras, since Japanese words change their meaning depending on whether the word contains a special mora or not. Special moras include the long vowel,

the moraic obstruent, and the moraic nasal. Studies show that long vowels are especially difficult for learners to acquire, and thus they were chosen for the focus of this study (Yokoi, 1998).

In order to acquire special moras, proper instruction on Japanese pronunciation is needed. However, many instructors are hesitant to include pronunciation training in class because they state that it is too difficult and they do not know exactly how to teach pronunciation (Matsuzaki, 2001).

When instructors try to incorporate pronunciation practice in class, it is most likely to be incorporated when they introduce new words. When introducing new words, instructors often have the class repeat after they say the word, and studies show that these repeat-after-the-model practices are beneficial for improving pronunciation (Kawai & Hirose, 2000; Yoshida & Fukada, 2014). However, since practice on producing special moras or pronunciation practice in general is not being done in class normally, it is questionable if this in-class repetition practice is enough for learners to become aware of the phonemic contrasts, and to improve their vowel duration.

In recent years, more and more Computer Assisted Language Learning (CALL) programs have been developed to overcome the problem that many instructors have with their lack of ability or confidence in teaching pronunciation. Recent studies use CALL programs to find out a useful practice method to create awareness of the notion of "mora", and one of the most commonly used visual aids is showing learners waveform displays (see Figure 1) (Ofuka, 1997; Okuno, 2013; Motohashi-Saigo & Hardison, 2009).

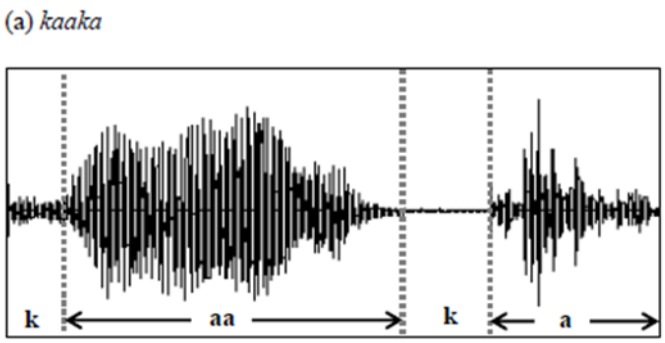


Figure 1: Example of Waveform Display
Source from Okuno (2013)

However, studies show that waveforms may not always be effective, because the length differences of the vowels are not always evident. Also, in experiments using waveforms, learners are often instructed to make their utterance as close as possible to the model waveform, but they do not receive any instruction on how to do so. Therefore, it is a guess-and-try process, and as such, it is difficult to assume that this could lead to acquisition.

The following paragraph is a summary of gaps in previous studies. The present research was designed to fill these gaps. First, previous studies show that repetition practice is beneficial for improving learners' pronunciation (Kawai & Hirose, 2000; Yoshida & Fukada, 2014), but not many studies focus on vowel duration in particular. Second, although studies show that repetition practice is beneficial, it does not compare the effectiveness of repetition practice when it is done in different environments. Third, most research uses waveform displays, and not enough studies use other input methods to

see its effectiveness. Fourth, not enough studies have conducted a delayed posttest to examine the retention after learners receive practice on vowel duration.

In order to fill these gaps, the design of the present research included: 1) focusing on learners' improvement in their vowel duration after doing repetition practice, 2) comparing the effectiveness of repetition practice when it is done in an in-class environment and in a self-study environment, 3) using a visual aid other than a waveform display, and 4) conducting a delayed posttest to examine learners' retention in their improvement in vowel duration. The purpose of this study is to investigate the effectiveness of online oral practice on creating awareness of the accurate vowel duration, particularly when repetition practice is conducted and also, out of the three input methods implemented in the study, which input method enhances acquisition the most in which environment. This research was designed to answer the following questions.

## Research Questions

**Research question 1.** Do learners of Japanese show significant improvement in their vowel length production after receiving repetition practice?

**Research question 2.** Among the three different treatment groups, which groups show the most improvement from the pretest?

**Research question 3.** Do different input methods influence learners' retention?

## Summary of Chapter 1

This chapter introduced the motivation for this study, as well as the areas where this study attempts to further investigate in promoting mora awareness among learners of Japanese. In the following chapter, literature reviews on the areas of history of pronunciation teaching, importance of practice in SLA, and recent attempts to create awareness of the special mora will be presented

# CHAPTER 2: LITERATURE REVIEW

The goal of this study is to investigate the effectiveness of repetition practice to improve Japanese learners' accurate vowel length production, as well as the effect that different practicing environments and practicing methods have on enhancing correct vowel length acquisition.

In this chapter, the general trend in today's pronunciation instruction and issues concerning pronunciation teaching will be reviewed and discussed. The chapter will first explain the basics of the Japanese phonological system since that is the main focus of the present study. After that it will briefly go over the history of pronunciation teaching and offer reasons why pronunciation has often been neglected in L2 instruction in recent years. Then it will be followed by reviews of previous studies on the importance of practice in SLA, effectiveness of repetition and drill practices, recent trends in pronunciation teaching, and attempts to create awareness of Japanese special moras using CALL programs.

## Japanese Phonology: Mora and Special Moras

Just like English, Japanese words can be divided into syllables, but they can also be divided into a small unit called a mora, which has some different characteristics compared to syllables. Traditionally, a syllable is said to be divided into three additional

sub-units, which are onset, nucleus, and coda. However, a mora is not divided further into sub-units, and each mora is supposed to have the same length of time; this is why a mora is called a timing unit (Tsujimura, 2013). For example, while the word *London* has two syllables, the same word is made up of four moras (/*lo.n.do.n*/). The Japanese writing system of *hiragana* and *katakana* (*katakana* used for words borrowed from other languages) is also based on moras. Taking *London* as an example again, the word is spelled as LO-N-DO-N. As explained earlier, this word consists of four moras and is also written with four *katakana.*

Most moras are also syllables (e.g. ka, sa, and ta), but there are moras that are non-syllabic: i.e. long vowel, the moraic obstruent, and the moraic nasal. Since these moras have a special status, they are called special moras. Toda (2003) claims that special moras are difficult to acquire regardless of the learner's L1 because Japanese time unit structure is different from many other languages in the world, such as English, German and Swedish that are stressed-time and syllable-timed languages such as French, Italian, and Spanish. Toda states that the field of phonetics is known to have the most influence of the L1 so that is why the acquisition of the special mora is difficult among learners.

### Contrastive Nature of Special Moras

The contrastive nature of special moras can change the meaning of words, depending on whether the word contains a special mora or not, and that is why special moras are considered one of the aspects that affects intelligibility in Japanese. This is exhibited in the following examples: 1) *obasan* (middle-aged woman) vs. *obaasan* (old lady), 2) *kite* (please come) vs. *kitte* (please cut), and 3) *tani* (valley) vs. *tani* (academic

credit). There is a long vowel in the first example, and this is something learners should be careful of since learners could offend a middle-aged woman by accidentally calling her an “old lady.” The second example involves the single vs. double consonant contrast, and the speaker could be giving two totally different requests based on how accurately they pronounce the word. Finally in the last example, the word for “academic credit” contains the moraic /N/ while the word for “valley” does not.

As the examples above demonstrate, one mora can change the entire meaning of a word. That is why it is important for learners to create awareness of the special moras and to practice producing it in an early stage of Japanese language learning.

### Acquisition of Long Vowels

Studies have shown difficulties that learners have in accurately pronouncing special moras, and among the special moras, long vowels are said to be the most difficult ones to pronounce (Yokoi, 1998). For learners to produce accurate vowel length, they first need to acquire the ability to distinguish whether the vowel they hear is a long one or a short one. Recently, more empirical studies have been done in the fields of acquisition of Japanese long vowels, both in terms of learners’ perceptive ability and productive ability to examine how learners process the new sound system.

Studies have shown mixed results regarding the correlation between learners’ language level and their acquisition of the long vowels. In terms of perception, Enomoto (1992), Toda (1998), and Oguma (2000a) found that learners’ accurate perception increases as their language level increases, but Uchida (1993) and Minagawa (1995) found otherwise. Although Toda (1998) concluded that the advanced-level learners’

ability to distinguish between long and short vowels is more accurate than beginning-level learners, the advanced level students were still not able to accurately distinguish between the two in a fast-paced free conversation.

In terms of production, Tsuchiya (1992) found that learners' level and their ability to accurately produce vowel duration did not have any correlation, but Nagai (1997) found that advanced-level learners' production was more accurate than the beginning-level learners'. However, Nagai's advanced learners did not reach a native level in producing long vowels. Oguma (2001a) found that there were no differences between beginning and intermediate-level learners but learners' ability to accurately produce long vowels at the word level seems to improve once learners reach the advanced-level. However, in another study of hers, she observed that learners who are in a transition stage between beginning and intermediate-level improved pronouncing words that contain long vowels in short sentences (Oguma, 2001b).

Although there are disagreements about the correlation between language level and the acquisition of the vowel length contrast, very few studies show a result where advanced-level learners have achieved a near-native level in pronouncing long vowels. Studies have found that advanced-level learners are able to produce and perceive long vowels more accurately compared to beginning-learners, but are still below the level of a native Japanese speaker (Nagai, 1997; Toda, 1998). Oguma's study in 2006 added another level, a superior-level learners as her participants for her study. In her studies, the learner's level was classified according to ACTFL's (American Council on the Teaching of Foreign Languages) OPI (Oral Proficiency Interview) scores. The aim of this study was to investigate the production of vowel duration in a free conversation, so the

participants' data from the OPI was rated by eight native Japanese speakers. Since most production tests are conducted by either reading out loud individual words or short sentences, in which it is usually easier for learners to control their utterances, this study aimed to investigate how accurately they can pronounce the vowel duration in their free speech. The result showed that although 42% of the superior-level learners showed no unnaturalness in their utterance, the lengthening of short vowels in word-initial position and shortening of long vowels in word-final position were observed in the rest of the superior-level learners. Another study shows that it is easiest for learners to acquire long vowels in the word-initial position, then the long vowels in the word-medial position, and lastly long vowels located at the word-final position (Oguma, 2001a). Oguma's result showed that it is challenging for even learners who have reached the superior-level to produce the long vowels that are generally more difficult to acquire. Based on these previous studies, it can be argued that the acquisition process of correct vowel duration takes time, because learners' interlanguage must be changed or adjusted to the Japanese phonological system (e.g. changing English L1 speakers' stress-timed and Spanish L1 speakers' syllable-timed system to the Japanese mora-timed system).

### History of Pronunciation Teaching

SLA has gone through various stages where different methods have been used to teach the pronunciation of the target L2. According to Celce-Murcia, Brinton, and Goodwin (2013), the reform movement, which emerged in the 1890s was the first linguistic contribution to pronunciation teaching. During this movement, the International Phonetic Alphabet (IPA) was developed, which made it possible to describe and analyze

the sound systems of various languages. It is believed that later, the historians from the reform movement played a role in developing the Audiolingual approach in the 40s and 50s. Teachers using the Audiolingual method had their students imitate or repeat after the model, so that they could achieve native-like pronunciation. Some other methods, such as the Cognitive approach, the Silent way, and the Community Language Learning emerged in the 60s and 70s, but the Audiolingual method remained the dominant approach until the Communicative approach took hold in the 1980s. This approach, which is still the dominant method in language teaching today, focuses on language being the tool to communicate with, and the popularity of the Communicative approach method shifted the goal of pronunciation teaching. The goal was no longer to teach students to attain native-like pronunciation, but rather to achieve intelligible pronunciation (Celce-Murcia, Brinton, Goodwin, 2013).

Iba (2008) states that the intelligibility principle establishes its goal to simply being understandable, and does not put emphasis on becoming native-like. People who support the intelligibility principle claim that communication can be successful even if the foreign accent is noticeable. After this new approach gained popularity, the Audiolingual method became less and less popular and is not widely used today, because it is seen that forming habits by repetitious drills to become “native-like” is less interesting to learners and that more creative work is necessary in practicing speech production, rather than memorizing and imitating structural patterns in dialogues (Iba, 2008).

### Importance of Practice

Although it is frequently assumed by language teachers that practice is a necessity for learning a second language, empirical research on the issue on "practice" has been receiving less attention in the post-Audiolingual period as the idea of practice can bring bad memories of Audiolingual mechanical drills. However, researchers have started to explore a number of questions concerning what kind of practice is most effective, in what contexts, and for what kinds of learners (DeKeyser, 2007). Bransford, Brown, and Cocking (1999) stated that "in deliberate practice, a student works under a tutor (human or computer based) to rehearse appropriate practices that enhance performance" (p.166) and pointed out that when a specific skill is practiced in an efficient manner, it could accelerate the process to transfer the declarative knowledge to a real world performance criterion. When concepts introduced by non-linguists such as Bransford et al. are applied to language teaching, these researchers are not necessarily arguing that mechanical drills should be brought back into the classroom, but there is a need to contemplate what kind of practice is effective for what purpose.

### Effectiveness of Repetition and Drill Practices in Teaching Pronunciation

Wong and VanPatten (2003) claim drills as meaningless activity since the purpose of drill exercise is to correctly produce a form or structure, which can be completed without understanding the meaning. They state that drills "are not necessary or beneficial for foreign language acquisition of the development of fluency and should be discarded from instructional practice" (p.403). They make this claim based on their belief that acquisition is input dependent, and this input has to be meaning based or communicative

in nature. For this reason, they argue that drills will not lead to acquisition of grammatical structures. On the other hand, they state that learners who are advanced speakers or near-native speakers, have usually lived or studied abroad, and in that environment, they had to read in the L2, watch films and TV shows in the L2 and so on, and these are all interactions that learners can use to obtain meaningful input. Studies have shown that learners advance significantly when studying abroad (Lapkin, Hart, & Swain, 1995; Meara, 1994; Teichler & Maiworm, 1997), and Wong and VanPatten claim that "learners are not engaged in mechanical drilling with native speakers while studying abroad" (p. 417).

However, there are counterarguments to their statement as well. Leaver, Rifkin and Shekhtman (2004) wrote a response to Wong and VanPatten's article, stating that "more research is required about the kinds of drills that are needed, at what stages they are needed, and for what languages they are most effective" (p. 130). Empirical evidence that Wong and VanPatten cited was mainly collected from research conducted on learners of Spanish and French. Therefore, it may be problematic to generalize that drill practices are ineffective across all languages without finding more empirical evidence.

Also, there are instances where improvement was not seen from students who have studied abroad, counter to what Wong and VanPatten were claiming as evidence to argue that drills are ineffective in language learning. Dekeyser (2007) proposes that the lack of efficacy in those students who did not show improvement after studying abroad can be attributed to various factors, such as student's aptitude, personality, and motivation. Although these factors are something that language instructors cannot control, Dekeyser argues that instructors can help students by better preparing their

experience overseas both in terms of language knowledge and language learning strategies. Therefore, effective practice before going overseas is required as well as having "good" practice while they are abroad, which could be instructed by the teachers beforehand. This is why the idea of practice becomes crucial when making progress in language skill (DeKeyser, 2007).

Coming back to the debate concerning the effectiveness of drills, what is most problematic is not Wong and VanPatten's statement, but the interpretation that the readers of their article are making. Although this strong argument that Wong and VanPatten made is specifically about grammar instruction in the classroom, it seems like their argument has influenced the field of pronunciation instruction as well. One of the major reasons why pronunciation practice has been receiving less attention in the classroom could be because many instructors associate the repetition and drill practices for pronunciation with the Audiolingual method, which is not very popular today. However, the effectiveness of repetition and drill practices for pronunciation acquisition is an entirely different study, and it is dangerous to assume that since drills are not good for acquiring grammatical forms and structures (Wong & VanPatten, 2003), drills will also not be effective for acquiring accurate pronunciation skills. The aim of pronunciation practice is not geared toward the mastery of specific structures or rules, and that is why the efficacy of drills cannot be said to be the same for pronunciation practice. As Paulston (1974) stated, teachers should not abandon a method simply because it is old, but rather need to consider which methods that are older could still be used for specific aspects of language teaching.

**Recent Trends in Pronunciation Teaching**

Lambacher (1999) states that when pronunciation practice is incorporated in the classroom, recent trends show that the focus has been on the suprasegmental features of rhythm, stress, and intonation, and segmental features have received significantly less attention. It has been believed that prosodic features are more important for the improvement of intelligibility and therefore, improving learner's prosodic features including rhythm, stress, and intonation have been the main focus of pronunciation training in recent years (Lambacher, 1999). It is however, important to address the issue of the difficulty associated with learners learning non-native phonemic contrasts. If learners are not pronouncing the different phonemic contrasts accurately, this can create significant communicative problems because, for instance, not being able to accurately pronounce one word could lead to difficulty in comprehending the overall sentence, as it was demonstrated earlier using Japanese special moras. Therefore, it is important to look at the smaller units that learners are having difficulty with, because they are often times linked to the larger suprasegmental features (Lambacher, 1999).

This is where the aid of repetition and drill practices becomes useful, as it could help internalize the phonemic differences between the target language and the learners' L1. This is not the same as stating that drill practice should be used throughout the language class. However, as Leaver et al. (2004) stated, different methods should be used based on the different stages that learners are at, and using repetition practice especially in the initial stage of pronunciation practice could be beneficial in internalizing the sound differences from the learners' L1. Research shows that simple repeat-after-the-model exercises are beneficial and can help learners (Kawai & Hirose, 2000; Yoshida &

Fukada, 2014), but it is also true that this area has not been widely investigated in the field of SLA. This may be due to the same reason why pronunciation instruction is frequently neglected; researchers may have been avoiding conducting repetition drill research because it is generally a method that is not in favor among many scholars and instructors. In the case of Japanese, many researchers have focused on investigating perception skills rather than production skills, and the effect that repetition has on enhancing production skills has not been explored as much.

### Attempts to Create Awareness of the Special Mora

Schmidt (1993) has claimed that language learning is possible only when learners consciously notice all aspects of language, including pronunciation. Thus, the role of instruction becomes important as it could make certain features of the target language more salient, which would facilitate the learners to notice. Following this idea of noticing, many scholars have investigated ways to create awareness of the special moras among Japanese learners by using visual aids, including lip movements and waveform displays using CALL programs.

Hirata and Kelly (2010) conducted research to find out what kinds of training would help native English speakers to perceive phonemic vowel contrasts in Japanese. Their study included an audio-only group, in addition to three other groups; audio-mouth, audio-hands, and audio-mouth-hands groups. These methods were based on the previous research finding that multimodal information such as lip movements and hand gestures influence many aspects of language processing. For their study 60 participants, who were all native speakers of English, were divided into four groups and all of them took a pre-

and posttest where they listened to the input in a quiet lab and decided whether the vowel in each word was short or long. The audio-only group simply listened to the audio-track with a still image of a native speaker on the screen. Participants in this group did not receive any information on the length of the target vowels. Participants in the audio-mouth group watched a video where the speaker's mouth was clearly synchronized with the sentence they heard, while the speakers' body remained still. In the audio-hands condition, participants watched a video where the face of the speaker was obscured by using a digital pixelization technique, but having a clear image of the speaker's hand making gestures that corresponded with the length of vowels of the spoken words. Speakers produced quick hand flicks for each short vowel in the word, and when words contained both short and long vowels, the speaker produced one quick hand flick and a prolonged hand sweep that extended horizontally. For the audio-mouth-hand condition group, the digital pixelization technique was taken off the speaker's face and the participants were able to see both the mouth and hand gestures clearly while listening to the target words. Lip movements and hand gestures were chosen as the independent variables, and their effectiveness on Japanese vowel length perception was investigated.

The results showed that the proportion of correct responses from pre- to posttest improved for all four groups, but the participants in the audio-mouth condition group improved the most. One interesting finding was that seeing lip movements helped learners to perceive difficult second language phonemic contrasts, while seeing hand gestures did not. It may be generally assumed that the more visual information the learners receive, the more helpful it is, but the findings revealed otherwise, as the audio-only condition improved more than the audio-mouth-hand condition. It could be the case

that participants were overloaded with visual input, and got distracted rather than benefited from it. Another possibility is that the participants' attention may have been drawn onto the hand gestures more than the lip movements in the audio-mouth-hand condition since the hand gestures were more salient, and that distracted them from looking at the useful lip-movement. McGurk and MacDonald's (1976) study illustrated how much lip movement influences human's perception of sound (which is now called the McGurk effect), and Hirata and Kelly's finding supports this idea on lip movement having a great impact on perception (McGurk & MacDonald, 1976).

Although the findings of the study are interesting, it has some limitations. The participants in this study were not learners of Japanese; therefore, it becomes difficult to generalize or compare Hirata and Kelly's findings with other studies where they recruit their participants only from those who are studying Japanese or have studied Japanese and are at least familiar with the language. Students who are taking Japanese are probably constantly hearing words that contain long vowels in their classroom, therefore, the difference in input could have an effect on the result between participants who are learners of Japanese and participants who are not. Also, students who are taking Japanese are probably somewhat interested in Japanese if not at all, and that desire to learn Japanese could have had an effect on their improvement. If participants are not interested in Japanese, they probably do not care as much to make improvements in their utterance and hence, could have had a different outcome.

The next three studies involve the use of waveforms, which is one of the most frequently used visual aids in recent years.

Ofuka (1997) examined the effect of using kinetic and computer assisted visualization to promote Japanese long and short vowel awareness. These two approaches were chosen because previous studies had shown that kinetic movement training can assist students with production while the visual reinforcement training can provide the learners with feedback. Seven learners of Japanese in an intermediate-level course volunteered to participate in this study. In the kinetic movement training, participants stretched their arms as they pronounced the long vowels in order to help them visualize and notice the different lengths of vowels using Total Physical Response (TPR). Once they had noticed the difference, it was thought that this could help participants become aware that they could indeed improve pronunciation with the aid of training. In the computer assisted visualization approach, learners watched the waveform and energy values that were produced by a native Japanese speaker and had to make their utterance as close as possible to the waveform they saw. The recognition score of the pretest and posttest was used as the measure to compare the gains between the two tests. Although both treatments focused on the production of long and short vowels, the result showed that both treatments had positive effects on distinguishing the two vowels not only in terms of production, but also for recognition. The treatment was most effective for errors with long vowels with a pitch rise, which suggests that there is a correlation between pitch and long/short vowels.

Although the findings from this study are helpful for future studies, it may be difficult to apply the treatment as it was used in reality. This research was conducted with a small sample size so it was possible to implement a tutorial style lesson. However, there is a need to propose an idea on how to incorporate this tutorial style training in a larger

classroom. Also, this study did not use any statistical analysis, therefore, it is questionable whether this finding could be generalized or not.

Motohashi-Saigo and Hardison (2009) conducted a study using the waveform display to facilitate the acquisition of Japanese double consonants. Using a pretest-posttest design, they compared its effect on both perception and production of 40 English-speaking beginning-level learners of Japanese, who were divided into three treatment groups: auditory-visual (AV) group, auditory only (A-only) group, and control group. The difference between the AV group and the A-only group was that while the AV group received information about the double consonants and also saw the waveform comparing single and double consonants, the A-only group only received the same information about double consonants but was not given a waveform display. Both experimental groups conducted an identification task. To evaluate the effectiveness of the training on double consonant identification and to compare training types, accuracy scores of the means of pretest and posttest from the two treatment groups were submitted to a mixed ANOVA. The result showed that both groups made significant improvement on double consonant identification as well as production, but greater improvement was seen from the AV group. Motohashi-Saigo and Hardison also surveyed the participants after the test, and found positive comments. They summarized their findings by stating that the training increased the awareness of the differences between single and double consonants among participants, and waveform displays helped the participants in the AV training group to understand where they should focus their attention when producing words including the double consonant. Also, participants said they liked the Web-based format of the training, since it gave them flexibility to practice during their spare time.

The limitation of this Web-based format training is that no feedback was given to the participants. It would be interesting to see the effect that corrective feedback would have on enhancing the learner's pronunciation skills. If learners can improve significantly without feedback, this type of outside-of-classroom practice can be beneficial in that teachers do not have to worry about time when they want to incorporate pronunciation teaching into their curriculum.

Similarly to Ofuka (1997) and Hirata and Kelly's (2010) study, Okuno (2013) also conducted a study to investigate the factors affecting learners' ability to correctly perceive and produce Japanese vowel duration. Just like Motohashi-Saigo and Hardison's study (2009), Okuno divided the 64 participants, who had American English as their L1 into three groups; Audio-only (A-only), Audio-visual (AV), and no training (control) groups. Participants were all learners of Japanese, but their levels ranged from first-year to fourth-year. Instead of only comparing the modality of training, Okuno also compared factors such as different preceding consonants, vowel types, talker's voice and pitch pattern to investigate the various effects each variable had. In the pretest, participants took both production and perception tests, where they read out loud the given pseudo-words, as well as responding to a four-alternative identification task. During the training in the A-only group, participants listened to the stimuli and were asked to choose the word they heard from the options. Immediate feedback was given after the participants made their choice, saying whether the answer they chose was correct or not. If they chose the incorrect answer, they were given another chance to listen to each stimulus again. As for the participants in the AV group, while listening to the stimulus, they were provided with a waveform to watch and then chose what they heard, just like the A-only group.

Participants in this group also received feedback, and the waveform was shown with the feedback so that they could attend to the visual information while paying closer attention to the form. Participant's accuracy score for both production and perception were used for statistical analysis to compare the improvement from pretest to posttest.

The results showed a significant improvement on identification accuracy from pretest to posttest for both groups. The result also indicated that although participants did not receive any production training, their production accuracy improved for both of the experimental groups. Motohashi-Saigo and Hardison (2009) and Ofuka (1997) attested that transferring occurs during perceptual and production training tasks, and likewise, Okuno's (2013) result also showed the positive effect that perceptual training has on producing correct L2 vowel duration. Abbs, Gupta, and Khetarpal's (2008) finding could help to explain Okuno's result. In their study, they found that learners tend to unconsciously engage in internal silent repetition when receiving perceptual training. Therefore, if it is assumed that the same thing happened to Okuno's participants, the reason why production improved while only engaging in perceptual practice could be explained.

Okuno's findings also showed that there were certain types of consonant-vowel (CV) combinations that improve more through this training than some. The three token types, (in which the syllables are divided by the period) CVV.CV, CV.CVV, and CV.CV significantly improved after the training while the token type of CVV.CVV did not. In addition, Okuno found that pitch pattern affects the efficacy of increasing learners' identification accuracy. It was easier for participants to correctly identify the vowel

duration when the word had a Low-High pitch in the first syllable and with the High-High pitch in the second syllable.

However, contrary to previous studies (Motohashi-Saigo & Hardison, 2009; Ofuka 1997), Okuno found that waveform was not helpful for the participants in noticing the vowel duration, since there was no significant difference found between the two training groups. Okuno's study also did not investigate whether retention would occur among participants, which would be important to know to decide whether this type of practice should be implemented into the curriculum.

As seen from the studies reviewed in this chapter, it has been common to use a CALL program in recent years to see the effect it has on assisting learners to notice the vowel duration. The present study will also examine the effectiveness of online oral practice on creating awareness of the accurate vowel duration, particularly when repetition practice is conducted. It will also compare the effectiveness between practicing online in a self-study environment versus practicing in the classroom environment without using any computer device. In order to investigate the effectiveness of repetition practice as well as the effect that different practicing environments and practicing methods have on enhancing correct vowel length acquisition, three different input modalities were constructed: 1) Audio-Only 2) Audio-Video, and 3) In-Class. However, this study will not use the frequently used waveform display as the visual aid input. As shown in Okuno's study, waveform display does not always help to enhance the learner's perception on the difference in vowel duration because 1) many people do not understand what the waveform is representing, 2) waveforms do not always show a distinct difference between a short and a long vowel, and 3) many people do not know how to

make their utterance as close as possible to the waveform they see because they do not receive any instruction on how to do so. Therefore, a new visual aid is used in this study, and since the difference in retention between groups was not observed in Okuno's study, this study will fill in the gaps by conducting a delayed posttest to see if there is any difference in retention when participants receive different treatments.

## Summary of Chapter 2

In this chapter, the possible reasons to why pronunciation practice has been neglected in recent years were further discussed, as well as reviewing previous studies that use CALL programs to promote awareness of Japanese special moras. The next chapter explains the methodology used for this present study.

# CHAPTER 3: METHODOLOGY

## Overview

This study examined different methods that could help increase Japanese learners' mora awareness. Three different methods were used in order to practice Japanese long vowel production, and total of 25 words were chosen as the stimulus words for the present study. All three practice methods were a simple listen-and-repeat practice, but they were conducted in a different environment with different input. The participants were randomly divided into three groups: 1) Audio-Only, 2) Audio-Video, and 3) In-Class. Audio-Only and Audio-Visual used a video-based oral training computer application called *Speak Everywhere* (Fukada, 2013) in order to practice long vowel production, while the treatment for In-Class was conducted in a classroom setting. All three groups received a pretest, practice, a posttest, and a delayed posttest. Figure 2 shows an overview of the design and the procedure of the experiment.

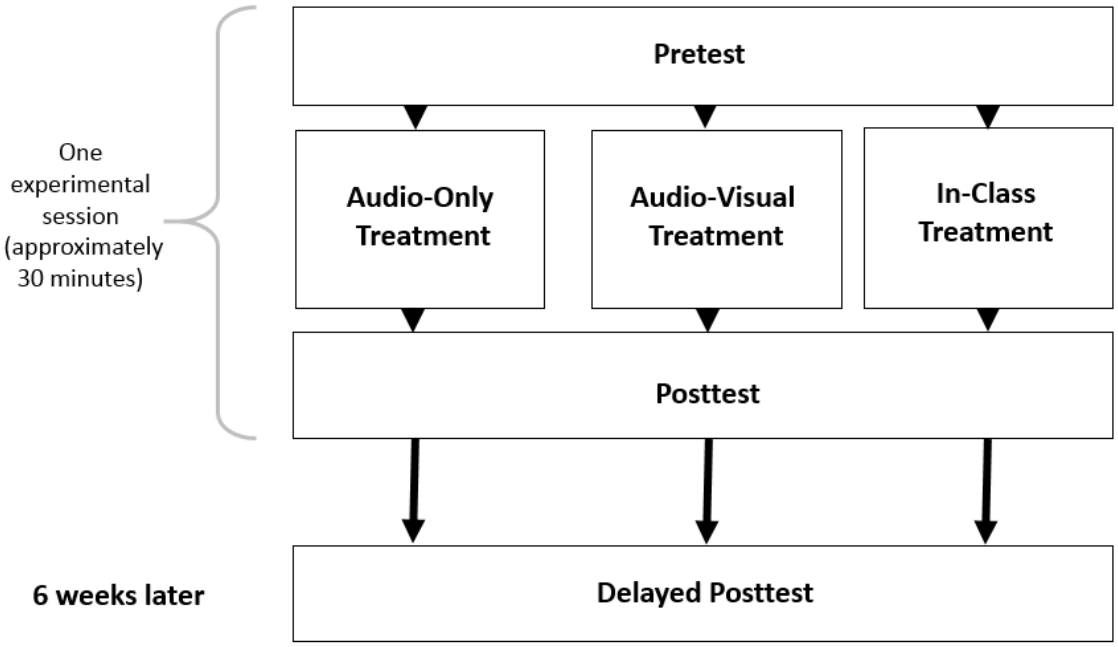


Figure 2: Overview of the Present Study

**Participants**

Participants in this study were those enrolled in Japanese 102 during the spring semester in 2015 at a Midwestern university. This is a course students can take after passing the Japanese 101 course (i.e. a total of approximately 60 hours of class instruction), which is an introductory course of Japanese offered at this university. Students can also take this course by getting placed in the course via placement test. In the elementary level courses (Japanese 101 and 102), students meet in class five times a week for 50 minutes each.

Table 1: Groups of the Present Study

| **Treatment groups** | **Participants**<br>**Total N=27** |
|---|---|
| Audio-Only<br>Audio-Visual<br>In-Class | N=10<br>N=11<br>N=6 |

Table 2: The Gender and L1Information for Each group

| **Group** | **Gender** | | **L1** |
|---|---|---|---|
| Audio-Only | M | 7 | English (4), English and American Sign Language (1), Chinese (5) |
| | F | 3 | |
| Audio-Visual | M | 9 | English (8), Chinese (3) |
| | F | 2 | |
| In-Class | M | 4 | English (5), Chinese (1) |
| | F | 2 | |

The participants have been using the software *Speak Everywhere* throughout the course. They have been using it to complete out-of-class assignments as well as conducting oral exams on it, thus, it was expected that the participants were familiar with the software.

In the textbook used at this university, (*Nakama 1*: written by Hatasa, Hatasa, & Makino) long vowels are introduced in the very beginning of the chapter where students

learn *hiragana*. Therefore, it was assumed that participants knew the basic concepts of Japanese long vowels.

This study was conducted on a voluntarily basis, and the participants were randomly divided into three experimental groups. 10 participants were assigned to be in the Audio-Only group, 11 participants were in the Audio-Visual group, and lastly, the In-Class group had 6 participants. In total, the participants consisted of 20 males and 7 females with their first language being either English or Chinese. The average age of the participants was 20.4 years old.

Over 80% of the participants had studied Japanese for one semester (meaning they had taken Japanese 101) prior to the semester in which this experiment was conducted. There were three participants who had studied Japanese in high school, ranging from two to four years. One of them was in the Audio-Visual group, and the other two were in the In-Class group. There were also two participants who had studied Japanese for only one to two months.

## Materials

**Stimulus words.** 25 words that the participants have not learned at the time this research was conducted were chosen (see Table 3). If already learned words, such as SE-N-SE-I (/sensee/) (“teacher”) were used, participants may have been able to say the long vowel correctly because they remember it as the part of the word, rather than knowing that it is a long vowel. The purpose of this study is to analyze whether learners are able to pronounce the words containing long vowel(s) solely from looking at the word and not from memorization, so that is why the unlearned words were purposely chosen. Within

the 25 words, 20 of them contained long vowels which had 23 long vowels in total since some words had more than one long vowel, and 5 words were chosen as fillers. All the filler words had corresponding words in the non-filler group that were identical to them except for the vowel length. In other words, each filler word and its corresponding word in the non-filler group formed a minimal pair (i.e. word with long vowel: YU-U-KI (/yuuki/) ("encouragement"); its minimal pair that was used as a filler: YU-KI (/yuki/) ("snow").

In *Nakama 1*, long vowels are introduced after all the *hiragana* have been presented. In the textbook, long vowels are explained as follows: "When the same vowel appears twice consecutively in a word, the two are pronounced as a continuous sound rather than as two separate vowels. This is called a long vowel" (Hatasa, Hatasa, & Makino, 2014, p.19).

As explained in the textbook, long vowels are created from simply extending the duration of the vowel sound. When the vowel sound of a character is /a/, it is extended by adding the *hiragana* for "A", which has the same vowel sound (i.e. for the word "mother", it is spelled as O-KA-A-SA-N and is pronounced as /okaasan/, and for the word "grandmother," it is spelled as O-BA-A-SA-N and is pronounced as /obaasan/). When the vowel has either an /i/ or /e/ sound, the vowel duration is extended by adding the *hiragana* for "I" (i.e. the word for "small" is spelled as CHI-I-SA-I and is pronounced as/chiisai/, and the word for "student" is spelled as GA-KU-SE-I and is pronounced as /gakusee/). Lastly, when the vowel sound is /u/ or /o/, the *hiragana* for "U" is used to extend the sound (i.e. "balloon" is spelled as HU-U-SE-N and is pronounced as /huusen/, and the word for "younger brother" is spelled as O-TO-U-TO and is pronounced as

/otooto/). The long vowel words that have the latter two cases are usually more difficult for learners of Japanese to grasp, because the corresponding *hiragana* for the sound /e/ is normally spelled with “E “, but instead of “E”, “I” has to be added instead. A similar situation is obtained with the /o/ sound; although the corresponding *hiragana* for the sound /o/ is spelled with “O”, “U” is added instead. There are exceptions where an /e/ vowel sound is extended by adding the *hiragana* for “E” or adding the *hiragana* for “O” to extend the /o/ vowel sound, but these cases are rare.

Out of the 20 words containing 23 long vowels, eight of the vowels were extended by adding “U” to the /u/ sound, nine of them were extended by adding ”I” when the vowel had the /e/ sound, and six of them had the vowel sound of /o/ and was extended by adding “U”. The latter two long vowels were assumed to be more difficult to produce, since the long vowel is not produced exactly how the *hiragana* is read, as explained earlier. If participants had enunciated each *hiragana* as they read the word, then it would not have become an accurate pronunciation of a long vowel.

Accent was not a variable in this study. The words used had these accentual properties: 13 out of 20 long vowel words have no accent, 5 of them have an accent on the first mora, and 2 of them have an accent on the second mora.

25 words were set in random order. However, the order was same for all pretest, treatment, posttest, and delayed posttest. The stimulus words are displayed in the following table. The pronunciation for each word is written in *romanization*, the long vowel is indicated by an underline, and the accent marker is placed above the mora that takes the accent.

Table 3: Stimulus Words

| ***Hiragana*** | **Spelling** | **Pronunciation** | **Definition** |
|---|---|---|---|
| ふつう | HU-TSU-U | hutsuu | Normal |
| すいてい | SU-I-TE-I | suitee | Estimate |
| ゆ・き | YU-KI | yuki | Snow |
| さいこう | SA-I-KO-U | saikoo | The highest |
| りゆう | RI-YU-U | riyuu | Reason |
| なんせい | NA-N-SE-I | nansee | South-west |
| つ・き | TSU-KI | tsuki | Moon/month |
| ゆ・うき | YU-U-KI | yuuki | Courage |
| おせいぼ | O-SE-I-BO | oseebo | Year-end gift |
| て・き | TE-KI | teki | enemy |
| ふう　とう | HU-U-TO-U | huu too | Envelope |
| く・うき | KU-U-KI | kuuki | Air |
| せつめい | SE-TSU-ME-I | setsumee | Explanation |
| つうき | TSU-U-KI | tsuuki | Ventilation |
| さ・いこ | SA-I-KO | saiko | The oldest |
| くつう | KU-TSU-U | kutsuu | Pain |
| ひ・こうき | HI-KO-U-KI | hikooki | Airplane |
| よてい | YO-TE-I | yotee | Plan |
| と・うき | TO-U-KI | tooki | Pottery |

Table 3: Stimulus Words (continued)

| ***Hiragana*** | **Spelling** | **Pronunciation** | **Definition** |
|---|---|---|---|
| くき | KU-KI | kuki | Stem |
| ていき | TE-I-KI | teeki | Regular/periodic |
| すいとう | SU-I-TO-U | suitoo | Water bottle |
| ゆう れい | YU-U-RE-I | yuuree | Ghost |
| ふくろう | HU-KU-RO-U | hukuroo | Owl |
| へい せい | HE-I-SE-I | hee see | Calmness |

**Online vocabulary practice.** The online oral practice software program called *Speak Everywhere* was used in the present experiment. The treatments for all three groups were a repeat-after-the-model practice, and this practice was conducted using *Speak Everywhere* for the Audio-Only and Audio-Visual group. All 25 words that these two groups listened to during their treatment were recorded by the present researcher. Although the In-Class group did not use this software to practice the words in their treatment, they used it to record their utterance for the pretest, treatment, posttest and delayed posttest, just like the other two groups. Thus, *Speak Everywhere* was used throughout the experiment for all three groups.

The Audio-Only and Audio-Visual used a headset with an attached microphone and the In-Class used a headset and a separate microphone. This was because it was more convenient for the In-Class treatment to have a headset and microphone that are separated, and the reason for this will be explained later in the procedure section.

Before the pretest, a sound check was conducted through *Speak Everywhere* to make sure that their audio was recorded properly and that their voice level was neither too soft nor loud. The volume of the headset was adjusted according to the participant's preference as well.

All groups conducted a pretest practice trial in the beginning of the session and this step was incorporated to familiarize the participants with the task. The format of the pretest practice trial was made exactly the same as the actual pretest and posttest; the difference is that the words used in the pretest practice trial were not used in the actual pretest, and the pretest practice consisted of only two words. The instruction for the pretest and each word with an English definition were displayed on the left-side of the screen, as shown in Figure 3. If a video was included in the exercise, it appeared on the upper right side of the screen. The format of the pretest, posttest, and delayed posttest was to say the word that appeared on the left-side of the screen and record their utterance.

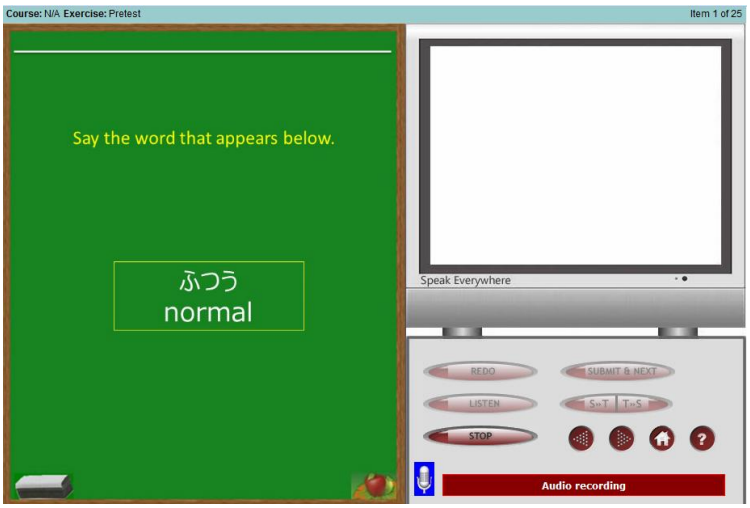


Figure 3: Pretest, Posttest, and Delayed Posttest Screen

Both Audio-Only and Audio-Visual groups also conducted a practice trial before starting their official treatment for the same reason. For both treatment groups, a video was set up on the upper right screen in addition to the instruction on the left side. For the Audio-Only group, a still image of the word appeared on their video screen, as shown in Figure 4.

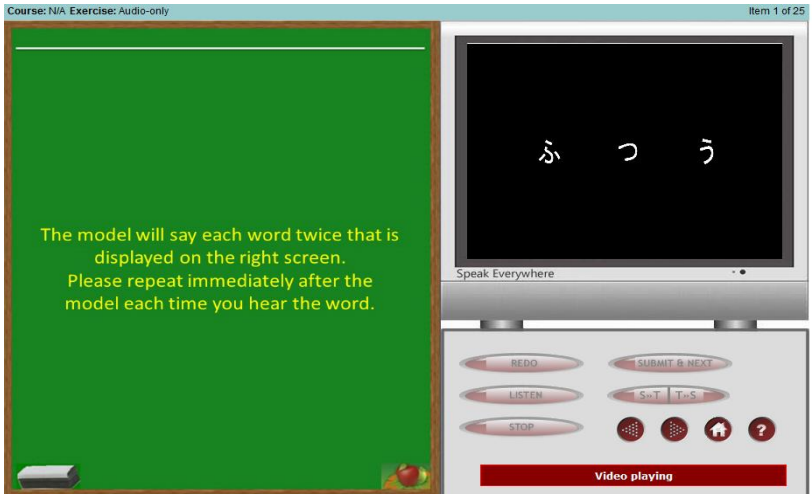


Figure 4: Audio-Only Treatment Group Screen

On the other hand, the Audio-Visual group had a video which modeled a "karaoke" system, where each mora in the word appears at the same time as that mora is being pronounced, and this system was designed to promote mora awareness among learners. An image is shown in Figure 5. Recent studies have focused on waveforms as visual cues in training learners to perceive the length difference, but found that this type of visual cue may not be effective since waveforms do not always show a clear distinction between a long and short vowel (Okuno, 2013). It is important for learners to

understand that the length of a long vowel is the same as any other mora, and thus, a visual aid that explicitly shows this concept was developed for this study. Participants in this treatment group were told to watch the video carefully when listening to the model saying the words.

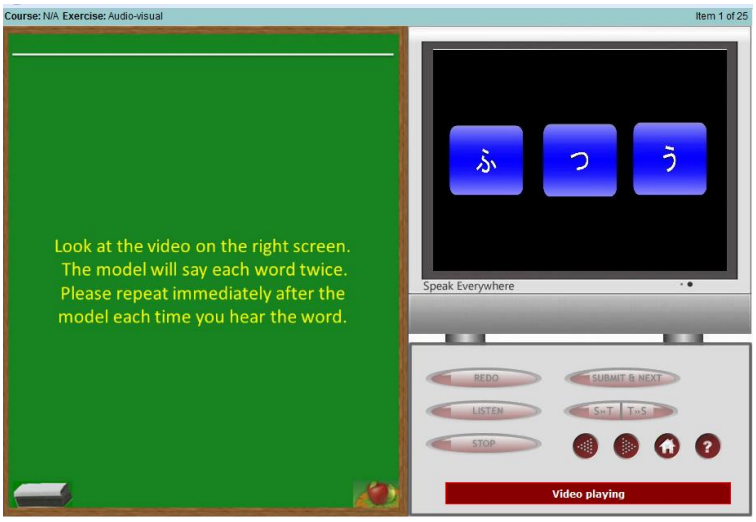


Figure 5: Audio-Visual Treatment Group Screen

Finally, the participants in the In-Class treatment also recorded their utterance during their treatment session.

**Procedure**

All three treatment groups took the pretest, treatment, the posttest, and six weeks delayed posttest. The pretest, treatment, and posttest were all completed within the same day. Six weeks after the pretest, practice, and posttest, all participants were asked to take

the delayed posttest to check their retention. The delayed posttest took the exact same format as the pretest and the posttest. For the delayed posttest, all participants took the test on their own, and were instructed to submit it within one week after the delayed posttest was assigned.

**Main experimental session.** Participants in the Audio-Only group and the Audio-Video group were assigned to meet with the present researcher individually, and the experiment was conducted in a small room. The experiment took place in a room where *Speak Everywhere* was already set up on the computer for them. All participants conducted this experiment using a school computer. The present researcher was the only person present in the room for these two groups.

The experiment for the In-Class treatment group took place in a computer lab where 6 participants came in at the same time. Their seats were assigned and were arranged in such a way that there would be at least one seat in between each participant, to ensure that they would not hear each other's voices as much and would not influence or distract each individual's recordings. *Speak Everywhere* was set up on the school computer for the participants in this group as well.

All participants first filled out the information sheet, where they were asked to provide their name, gender, age, major, first language (or languages they are most fluent in), and the length of time studying Japanese before the spring 2015 semester. Then, they were instructed to sign in to *Speak Everywhere* using an account that was set up by the present researcher.

Once logged in to *Speak Everywhere*, the first step all three treatment groups took was the sound test and after that all groups conducted the pretest practice trial, and

then they started their pretest. Although most participants did not choose to do so, they were given a choice to listen to their own recording and redo it if they wished to during the pretest and posttest. Most participants redid their recording when they noticed immediately that they said something different from what appeared on the screen. Since most of the participants studied Japanese for only one semester (16 weeks) prior to the present study, some of them still made mistakes when reading *hiragana*.

After the pretest, the Audio-Only group and the Audio-Visual group conducted the treatment practice trial and then the actual treatment. For both Audio-Only and Audio-Visual treatment groups, the model said each word appearing on the video screen twice. This was done to compare the effect with the In-Class treatment group. When an informal survey was taken among Japanese instructors at the university where this study took place, the average number of times the instructors had their students repeat the new vocabulary item in the classroom was twice. Therefore, the process went as follows; listen to the model and repeat, listen to the model again and repeat. For this present research, only the utterance that was made after the second repetition was recorded.

As for the In-Class group, they were instructed to take off their headsets, face the model and adjust the microphone position so that it could capture their voice during the treatment. The model for the In-Class group was a native speaker of Japanese who volunteered to do the pronunciation practice part of the experiment. The model is a 44 year-old female who is a Japanese instructor at the university where this study was conducted, but was not the instructor for any of the participants in the group. There, the model said the same 25 words that were in the pretest in the same order, and the participants repeated after the instructor. Since this treatment was conducted as a

simulation of the in-class environment, the model said every word twice. The participants' utterances during this repeat-after-the-instructor session were recorded. The participants were instructed to press the "RECORD" button when they were told to and then began the listen and repeat practice, and then pressed "STOP" once the repeating practice was over. Therefore, the difference between the online treatment groups and this In-Class group was that while only the second utterance was recorded for the two online groups, in the In-Class group, the whole practice session was recorded.

After the treatment, all participants completed the posttest. The entire process did not take more than 20 minutes for each group.

**Delayed posttest.** Participants received an email from the present researcher six weeks after the main experimental session was conducted. They were instructed to complete the delayed posttest on their own in a quiet room within one week from the day the email was sent. There was a link attached to the email where they could access *Speak Everywhere.* Participants logged in using the same account that they used from the main experimental session, and conducted the delayed posttest, which had the exact same procedure as the pre- and posttests. The delayed posttest did not take more than 5 minutes to complete.

### Measurement

Native speaker raters judged all the recordings collected from the pretest, treatment, posttest, and delayed posttest, and judged whether each vowel was appropriate in length or not. Two female native speakers of Japanese (Rater 1 the age of 35 and rater 2 being 37 years old) rated the pretest, posttest, and the delayed posttest data. Both raters

were born in Japan but are living in the United States currently, teaching Japanese at the university where this study was conducted. However, neither one of them were instructors of the participants.

Another two native speakers rated the practice round for all three treatment groups. Rater 1 is a female age 25 and rater 2 is a 25 year-old male, and both of them were born in Japan. Both of the raters for the practice round had also been teaching Japanese at the university where this study took place, but neither one of them were the instructors of the participants.

The raters were given multiple choices for each word, and as they listened to each word, they chose what they heard. The choices of the words were all written in *hiragana*, but for the long vowel words that are not pronounced exactly as it is written, the *romanization* version of the words were given to make the distinction clear.

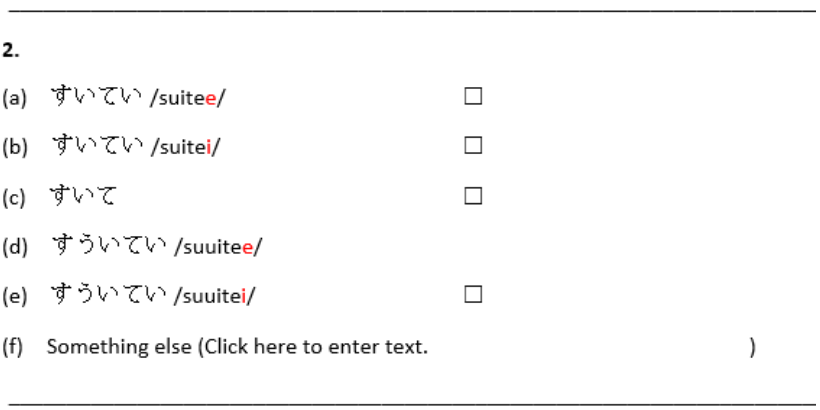


Figure 6: Sample Task for the Raters

The choices of the words were given based on what the present researcher often heard while she was sitting through the experiment with the Audio-Only and Audio-Visual groups. A choice of "something else" was also given to the raters, where they were asked to write what they thought they heard if it was not on the list of choices. All four raters rated the production individually, and for raters who rated the pretest, posttest, and delayed posttest, both of them first rated all the pretests, and then all the posttests for all 27 participants. Once all data for the delayed posttests were gathered, the same two raters conducted the same task.

For the practice round, the recording of the In-Class group was different from the other two because the whole practice session was recorded, unlike the online groups where only their second repeating utterances were recorded. Therefore, the two raters were told to rate the second repeating utterances while listening to the recordings taken from the In-Class participants.

The raters simply had to choose or write what they thought they heard, and later, the present researcher assigned points according to the following criterion. Each word was scored based on whether the participants said the length of each syllable in the word correctly. If they said it correctly, they got one point; if they did not, they got zero points. For example, the word HI-KO-U-KI (/hikooki/) ("air plane") has three syllables (hi/koo/ki). If they pronounced it as (hi/koo/ki), which is correct, they got three points. If they said (hii/koo/ki), they got two points because the last two syllables, (koo/ki) were pronounced correctly but the participant said the first syllable as a long vowel where it is really a short vowel and hence, got no points. If they said (hi/kou/ki), enunciating the /u/

sound in the second syllable, they also got only two points because that is not the correct way to pronounce a long vowel word.

The average of two raters' scores given to each word became the final score for that word. Then, those 25 vocabulary scores were added up and its sum was used as the score for each test (including the practice). Therefore, each participant had four scores in total, (scores for the pretest, practice, posttest, and delayed posttest) each being a measure of how accurately they controlled the length of vowels in the stimuli. These scores were submitted to various statistical analyses.

**Data Analysis**

Repeated-measures ANOVA was used to test Hypotheses 1 to 3 with the significance level set to 0.1. The sphericity assumption was tested with the Mauchly Sphericity Test. The statistical software package SPSS version 22 was used to perform this analysis.

Repeated-measures ANOVA was also used to test Hypotheses 4 to 7 with the significance level set to 0.1. The homogeneity of variance was tested with Levene's Test, and the Tukey pairwise comparison procedure was used as the post-hoc. The statistical software packages SPSS version 22 and MiniTab version 17 were used to perform this analysis.

**Summary of Chapter 3**

In this chapter, the methodology used for this present study was explained. In the next chapter, the following research questions will be answered by testing the hypotheses associated with each question.

**Research question 1.** Do learners of Japanese show significant improvement in their vowel length production after receiving repetition practice?

In order to answer research question 1, the following hypotheses were formed:

***Hypothesis 1***: There will be significant gain in the score within each group from pretest to practice.

***Hypothesis 2***: There will be significant gain in the score within each group from pretest to posttest.

***Hypothesis 3***: There will be significant gain in the score within each group from pretest to delayed posttest.

**Research question 2.** Among the three different treatment groups, which groups show the most improvement from the pretest?

In order to answer research question 2, the following hypotheses were formed:

***Hypothesis 4***: There will be significant gain difference among the three groups from pretest to practice. The score will be in the following descending order; Audio-Visual, Audio-Only, and In-Class group.

***Hypothesis 5***: There will be significant gain difference among the three groups from pretest to posttest. The score will be in the following descending order; Audio-Visual, Audio-Only, and In-Class group.

***Hypothesis 6***: There will be significant gain difference among the three groups from pretest to delayed posttest. The score will be in the following descending order; Audio-Visual, Audio-Only, and In-Class group.

**Research question 3.** Do different input methods influence learners' retention?

In order to answer research question 3, the following hypothesis was formed:

***Hypothesis 7***: There will be greatest retention from posttest to delayed posttest in the following descending order: Audio-Visual, Audio-Only, and In-Class group.

# CHAPTER 4: RESULTS AND DISCUSSIONS

In this chapter, the data collected from pretest, practice, posttest, and delayed posttest are analyzed and discussed.

## Descriptive Statistics

Prior to data analysis, one participant's record was removed from the In-Class group because she started the pretest before she was instructed to do so. Therefore, a data set of 26 records was used in the following data analyses.

Table 4 displays a summary of means, standard deviation (SD), minimum and maximum scores observed in each treatment group, and Table 5 shows the mean and SD of gained score from pretest to practice, posttest, and delayed posttest in each treatment group. Figure 7 shows the box plot of the overall scores for all treatment groups from the pretest, practice, posttest, and delayed posttest.

Table 4: Summary of the Measurements Used for Statistical Analysis

| | Audio-Only | | | | Audio-Visual | | | | In-Class | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Min | Max | Mean | SD | Min | Max | Mean | SD | Min | Max |
| Pretest | 43.7 | 6.825 | 32.5 | 53.5 | 45.4 | 7.406 | 29.0 | 54.0 | 54.3 | 2.332 | 51.5 | 57.3 |
| Practice | 54.8 | 3.302 | 47.0 | 58.0 | 55.4 | 1.241 | 53.5 | 57.0 | 56.7 | 0.837 | 55.5 | 57.5 |
| Posttest | 48.4 | 6.894 | 36.3 | 57.3 | 50.4 | 5.176 | 36.5 | 57.3 | 56.4 | 1.387 | 54.5 | 58.0 |
| Delayed | 46.1 | 7.316 | 35.8 | 57.5 | 48.8 | 6.329 | 32.3 | 54.5 | 54.9 | 1.557 | 53.3 | 57.3 |

Table 5: Summary of Means and Gain Score

| | Audio-Only | | Audio-Visual | | In-Class | |
|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD |
| Pretest | 43.7 | 6.825 | 45.4 | 7.406 | 54.3 | 2.332 |
| Practice | 54.8 | 3.302 | 55.4 | 1.241 | 56.7 | 0.837 |
| Posttest | 48.4 | 6.894 | 50.4 | 5.176 | 56.4 | 1.387 |
| Delayed | 46.1 | 7.316 | 48.8 | 6.329 | 54.9 | 1.557 |
| Gain 1 (Pretest-Practice) | 11.1 | 5.916 | 10.0 | 7.755 | 2.5 | 1.841 |
| Gain 2 (Pretest-Posttest) | 4.7 | 2.458 | 5.1 | 4.640 | 2.1 | 2.529 |
| Gain 3 (Pretest-Delayed) | 2.4 | 2.723 | 3.5 | 3.649 | 0.7 | 2.485 |

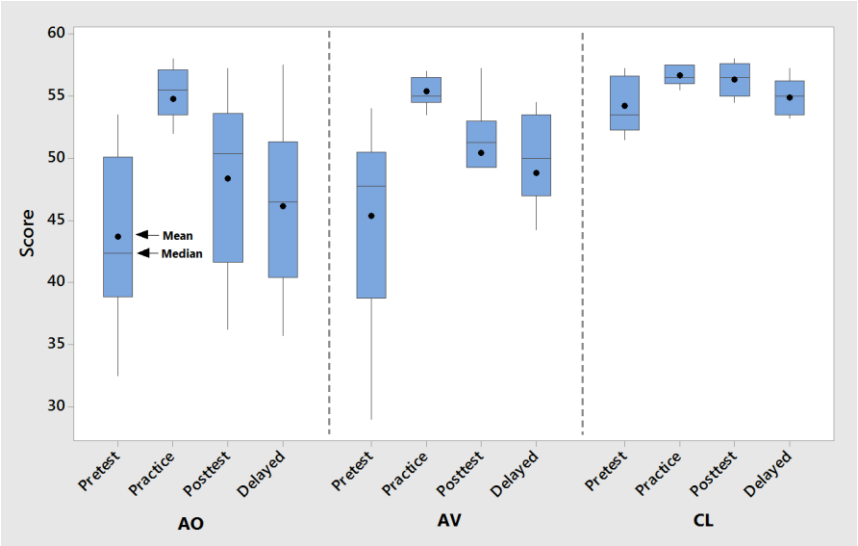


Figure 7: Box Plot of the Overall Scores
(AO = Audio-Only, AV = Audio-Visual, CL = In-Class)

One-way ANOVA found a significant difference among the three groups in the pretest scores. However, since this study compares the gain scores from pretest to practice, posttest, and delayed posttest, we do not consider the initial lack of homogeneity to be crucial. (see the limitation section in Chapter 5).

## Research Questions and Hypotheses

**Research question 1.** In order to answer research question 1, "Do learners of Japanese show significant improvement in their vowel length production after receiving repetition practice?", hypothesis 1, 2, and 3 were examined.

Prior to a Repeated-measures ANOVA, Mauchly's Sphericity Test was conducted. The results are shown in Table 6. Where there was a violation of the assumption of sphericity, the Greenhouse-Geisser Correction was applied. The *p*-Value of the Repeated-measures ANOVA were significant for all three treatment groups (see Table 7, Table 9, and Table 11). Therefore, pairwise comparison was conducted for further analysis.

Table 6: Results of Mauchly's Test of Sphericity

| | Mauchly's Test of Sphericity (p-Value) | Interpretation |
|---|---|---|
| Audio-Only | 0.102 | No violation of sphericity |
| Audio-Visual | 0.039 | Violation of sphericity; use Greenhouse-Geisser correction |
| In-Class | 0.633 | No violation of sphericity |

Table 7: Result of Repeated-measures ANOVA for Audio-Only Group

**Tests of Within-Subjects Effects**

Measure: MEASURE_1

| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| factor1 | Sphericity Assumed | 681.362 | 3 | 227.121 | 21.971 | .000 |
| | Greenhouse-Geisser | 681.362 | 1.699 | 401.135 | 21.971 | .000 |
| | Huynh-Feldt | 681.362 | 2.052 | 331.974 | 21.971 | .000 |
| | Lower-bound | 681.362 | 1.000 | 681.362 | 21.971 | .001 |

Table 8: Pairwise Comparison for Audio-Only Group

| | Pretest | Practice | Posttest | Delayed |
|---|---|---|---|---|
| Pretest | | **0.000** | **0.000** | **0.026** |
| Practice | | | 0.005 | 0.001 |
| Posttest | | | | 0.049 |
| Delayed | | | | |

Table 9: Result of Repeated-measures ANOVA for Audio-Visual Group

**Tests of Within-Subjects Effects**

Measure: MEASURE_1

| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| factor1 | Sphericity Assumed | 575.119 | 3 | 191.706 | 13.411 | .000 |
| | Greenhouse-Geisser | 575.119 | 1.618 | 355.467 | 13.411 | .001 |
| | Huynh-Feldt | 575.119 | 1.885 | 305.161 | 13.411 | .000 |
| | Lower-bound | 575.119 | 1.000 | 575.119 | 13.411 | .004 |

Table 10: Pairwise Comparison for Audio-Visual Group

| | Pretest | Practice | Posttest | Delayed |
|---|---|---|---|---|
| Pretest | | **0.002** | **0.005** | **0.010** |
| Practice | | | 0.008 | 0.007 |
| Posttest | | | | 0.131 |
| Delayed | | | | |

Table 11: Result of Repeated-measures ANOVA for In-Class Group

**Tests of Within-Subjects Effects**

Measure: MEASURE_1

| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| factor1 | Sphericity Assumed | 20.375 | 3 | 6.792 | 3.736 | .042 |
| | Greenhouse-Geisser | 20.375 | 1.675 | 12.163 | 3.736 | .086 |
| | Huynh-Feldt | 20.375 | 2.742 | 7.430 | 3.736 | .048 |
| | Lower-bound | 20.375 | 1.000 | 20.375 | 3.736 | .125 |

Table 12: Pairwise Comparison for In-Class Group

| | Pretest | Practice | Posttest | Delayed |
|---|---|---|---|---|
| Pretest | | 0.041 | 0.137 | 0.590 |
| Practice | | | 0.657 | 0.058 |
| Posttest | | | | 0.025 |
| Delayed | | | | |

***Hypothesis 1.*** There will be significant gain in the score within each group from pretest to practice.

Using pairwise comparison, the improvement between the pretest to practice was compared among all three treatment groups (see Table 8, Table 10, and Table 12).

Table 13: Result of Pairwise Comparison (Pretest-Practice) – H1

| | Pretest Mean | Practice Mean | Gain | Significance Level (p-Value) |
|---|---|---|---|---|
| Audio-Only | 43.7 | 54.8 | 11.1 | 0.000 |
| Audio-Visual | 45.4 | 55.4 | 10.0 | 0.002 |
| In-Class | 54.3 | 56.7 | 2.5 | 0.041 |

Table 13 summarizes the result of the pairwise comparison for all three treatment groups from pretest to practice. Significant improvement in the mean scores were seen in all three treatment groups. Therefore, this result confirms hypothesis 1, and indicates that learner's production in vowel length improves immediately after they hear the model saying the word, regardless of the practice type.

***Hypothesis 2.*** There will be significant gain in the score within each group from pretest to posttest.

The same analytical procedure as hypothesis 1 was undertaken to examine hypothesis 2 to see the improvement between the pretest to posttest within groups (see Table 8, Table 10, and Table 12).

Table 14: Result of Pairwise Comparison (Pretest-Posttest) – H2

| | Pretest Mean | Posttest Mean | Gain | Significance Level (p-Value) |
|---|---|---|---|---|
| Audio-Only | 43.7 | 48.4 | 4.7 | 0.000 |
| Audio-Visual | 45.4 | 50.4 | 5.1 | 0.005 |
| In-Class | 54.3 | 56.4 | 2.1 | 0.137 |

Table 14 summarizes the result of the pairwise comparison for all three treatment groups from pretest to posttest. Significant improvements in the mean scores from pretest to posttest were found in Audio-Only and Audio-Visual group, while the improvement was not found significant for the In-Class group ($p$=0.137). Therefore, the result only partially confirmed hypothesis 2; no short term improvement was observed in the In-Class group.

***Hypothesis 3.*** There will be significant gain in the score within each group from pretest to delayed posttest.

The same analytical procedure as the previous two was also undertaken to examine hypothesis 3 to see the improvement between the pretest to delayed posttest within groups (see Table 8, Table 10, and Table 12).

Table 15: Result of Pairwise Comparison (Pretest-Delayed) – H3

| | Pretest Mean | Delayed Mean | Gain | Significance Level (p-Value) |
|---|---|---|---|---|
| Audio-Only | 43.7 | 46.1 | 2.4 | 0.026 |
| Audio-Visual | 45.4 | 48.8 | 3.5 | 0.010 |
| In-Class | 54.3 | 54.9 | 0.6 | 0.590 |

Table 15 summarizes the result of the pairwise comparison for all three treatment groups from pretest to delayed posttest. Significant improvements in the mean scores from pretest to delayed posttest in Audio-Only and Audio-Visual were found, while the improvement was not found significant for the In-Class group ($p$=0.590).

Therefore, the result only partially confirmed hypothesis 3; no long term gain was seen in the In-Class group.

**Research question 2.** In order to answer research question 2, "Among the three different treatment groups, which groups show the most improvement from the pretest?", hypotheses 4, 5, and 6 were examined. Figure 8 shows the box plot of the overall gained scores from pretest to practice, posttest, and delayed posttest for all three treatment groups.

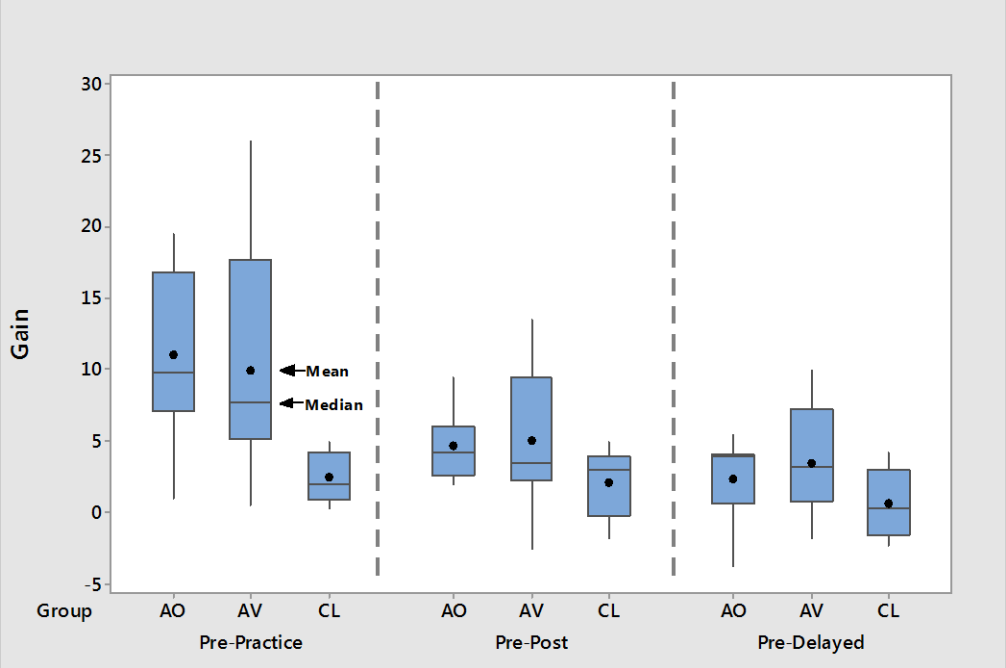


Figure 8: Box Plot of the Gain (= Gained Score)
(AO = Audio-Only, AV = Audio-Visual, CL = In-Class)

Before using the Repeated-measures ANOVA to analyze the data, Levene's test was first used to assess the equality of variance. The results are shown in Table 16. It was

found that the three treatment groups have equal variance (*p*-Value greater than the set significance level 0.1), thus, the data was further analyzed using Repeated-measures ANOVA. When the *p*-Value in the Repeated-measures ANOVA was found significant, the Tukey pairwise comparison procedure was used as the post-hoc to further compare the effectiveness between the different treatments.

Table 16: Results of Levene's Test for Equality of Variance

| | Levene's Test for Equality of Variance (p-Value) | Interpretation |
|---|---|---|
| Pretest-Practice | 0.265 | Equal variances assumed |
| Pretest-Posttest | 0.260 | Equal variances assumed |
| Pretest-Delayed | 0.692 | Equal variances assumed |

***Hypothesis 4.*** There will be a significant gain difference among the three groups from pretest to practice. The score will be in the following descending order; Audio-Visual, Audio-Only, and In-Class group.

Using Repeated-measures ANOVA, significant difference in the gained score from pretest to practice was found between groups (see Table 17). Therefore, the Tukey pairwise comparison was used as the post-hoc to further compare the gain scores between groups (see Table 18).

Table 17: Result of Repeated-measures ANOVA (Pretest-Practice) – H4

| Source | Gain | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| Gain | Linear | 712.116 | 1 | 712.116 | 35.226 | .000 |
| Gain * Group | Linear | 135.316 | 2 | 67.658 | 3.347 | .053 |
| Error(Gain) | Linear | 464.964 | 23 | 20.216 | | |

Table 18: Tukey Pairwise Comparison (Pretest-Practice) – H4

| | Audio-Only | Audio-Visual | In-Class |
|---|---|---|---|
| Audio-Only | | 0.768 | 0.048 |
| Audio-Visual | | | 0.015 |
| In-Class | | | |

As Table 18 shows, no significant difference in the gained score was found between Audio-Only and Audio-Visual groups ($p$=0.768), but there was a significant difference between Audio-Only and In-Class groups ($p$=0.048), as well as Audio-Visual and In-Class group ($p$=0.015). Thus, hypothesis 4 was partially confirmed; although it was not confirmed that the Audio-Visual group will have the most gained score among the three groups, the two groups that conducted the treatment online had a greater improvement from pretest to practice compared to the In-Class group.

***Hypothesis 5.*** There will be a significant gain difference among the three groups from pretest to posttest. The score will be in the following descending order; Audio-Visual, Audio-Only, and In-Class group.

The same analytical procedure as hypothesis 4 was undertaken to analyze hypothesis 5, and the difference in gained scores from pretest to posttest were compared between three groups. The result of the Repeated-measures is shown in Table 19.

Table 19: Result of Repeated-measures ANOVA (Pretest-Posttest) – H5

| Source | Factor | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| Factor | Linear | 179.404 | 1 | 179.404 | 27.954 | .000 |
| Factor * Group | Linear | 16.021 | 2 | 8.011 | 1.248 | .306 |
| Error(Factor) | Linear | 147.609 | 23 | 6.418 | | |

The result of the Repeated-measures ANOVA showed that there is no significant gain difference among the three groups from pretest to posttest ($p$=0.306), therefore, a post hoc was not conducted for further analysis. Although the Audio-Visual group did show the most gain when solely looking at the raw scores as shown in Figure 8, the significant gain difference was not proven statistically. Therefore, this result did not support hypothesis 5.

***Hypothesis 6.*** There will be a significant gain difference among the three groups from pretest to delayed posttest. The score will be in the following descending order; Audio-Visual, Audio-Only, and In-Class group.

The same analytical procedure as hypotheses 4 and 5 was undertaken to analyze hypothesis 6, and the difference in gained scores from pretest to delayed posttest were

compared between three groups. The result of the Repeated-measures ANOVA is shown in Table 20.

Table 20: Result of Repeated-measures ANOVA (Pretest-Delayed) – H6

| Source | factor1 | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| factor1 | Linear | 54.913 | 1 | 54.913 | 10.887 | .003 |
| factor1 * Group | Linear | 13.826 | 2 | 6.913 | 1.371 | .274 |
| Error(factor1) | Linear | 116.007 | 23 | 5.044 | | |

The result of the Repeated-measures ANOVA showed that there is no significant gain difference among the three groups from pretest to delayed posttest ($p$=0.274), therefore, a post hoc was not conducted for further analysis. Similarly to hypothesis 5, Figure 8 shows that both groups that practiced online had a greater gain from pretest to delayed posttest test compared to the group that practiced in class. However, the gained difference between the treatment types were not proven statistically.

**Research question 3.** In order to answer research question 3, "Do different input methods influence learners' retention?", hypothesis 7 was examined. The same procedure was used as research question 2 to analyze the data. Before using the Repeated-measures ANOVA, Levene's test was first used to assess the equality of variance. The result is shown in Table 21. It was found that the three treatment groups have equal variance ($p$-Value greater than the set significance level 0.1), thus, the data was further analyzed using Repeated-measures ANOVA.

Table 21: Result of Levene's Test for Equality of Variance

| | Levene's Test for Equality of Variance (p-Value) | Interpretation |
|---|---|---|
| Posttest-Delayed | 0.145 | Equal variances assumed |

***Hypothesis 7.*** There will be greatest retention from posttest to delayed posttest in the following descending order: Audio-Visual, Audio-Only, and In-Class group.

The same analytical procedure as hypotheses 4, 5 and 6 was undertaken to analyze hypothesis 7, and differences in scores (=negative gained scores, as it was assumed that the score would decline somewhat from posttest to delayed posttest) were compared between three groups. The result of the Repeated-measures ANOVA is shown in Table 22.

Table 22: Result of Repeated-measures ANOVA (Posttest-Delayed) – H7

| Source | factor1 | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| factor1 | Linear | 35.806 | 1 | 35.806 | 8.465 | .008 |
| factor1 * Group | Linear | 1.555 | 2 | .778 | .184 | .833 |
| Error(factor1) | Linear | 97.292 | 23 | 4.230 | | |

The result of the Repeated-measures ANOVA showed that there is no significant retention difference among the three groups from posttest to delayed posttest ($p$=0.833), therefore, a post hoc was not conducted for further analysis. The result did not confirm

hypothesis 7, and research question 3 can be answered; it was not statistically proven that different treatment types influence learners' retention in this study.

### Summary of the Findings

**Research question 1.** Hypotheses 1, 2, and 3 were examined in order to answer research question 1, "Do learners of Japanese show significant improvement in their vowel length production after receiving repetition practice?"

In hypothesis 1, it was postulated that there will be significant gain in the score within each group from pretest to practice. The result confirmed the hypothesis, showing that all three groups significantly improved their pronunciation when engaged in a listen-and-repeat exercise, and that they can make their vowel duration sound closer to the model immediately after listening to it. Therefore, hypothesis 1 supports the previous studies' claim that simple repeat-after-the-model exercises are beneficial for improving pronunciation, (Kawai & Hirose, 2000; Yoshida & Fukada, 2014) and shows that it is helpful in introducing the accurate vowel length to the learners' sound system.

It was hypothesized in hypothesis 2 that there will be significant gain in the score within each group from pretest to posttest and it was partially confirmed, showing that there was a short term improvement for learners who did the listen-and-repeat practice online in a self-study environment, while a short term improvement was not seen among learners who engaged in the listen-and-repeat practice in an in-class environment ($p$=0.137). This means that participants who practiced online were able to retain the accurate vowel length they heard from the model during the listen-and-repeat exercise immediately after the practice, while participants who practiced in class could not.

In hypothesis 3, it was hypothesized that there will be significant gain in the score within each group from pretest to delayed posttest, and this was also partially confirmed, showing that there was a long term improvement for learners who did the listen-and-repeat practice online in a self-study environment, while a long term improvement was not seen among learners who received the listen-and-repeat practice in an in-class environment ($p$=0.590). Similarly to the findings in hypothesis 2, this indicated that participants who practiced in a self-study environment was able to retain the accurate vowel length they heard from the model during the listen-and-repeat exercise even six weeks after the practice, while participants who practiced in class could not.

From these three hypotheses, research question 1 can be answered. Learners of Japanese show significant improvement in their vowel length production regardless of how they receive repetition practice when they repeat immediately after the model. However, different results were obtained when the learners were tested afterwards. Learners who did the listen-and-repeat practice online showed both short and long term improvement in their vowel length production from pretest, even without listening to the model. However, learners who did the listen-and-repeat practice in an in-class environment showed neither short nor long term improvement in their vowel length production. These results show that listen-and-repeat practice is beneficial for both short term and long term improvement in their vowel length production when the practice is done in an online self-study environment.

**Research question 2.** Hypotheses 1 to 3 were examined to investigate whether learners of Japanese show significant improvement in their vowel length production after receiving repetition practice. The overall effectiveness of repetition practice was found

when participants received practice in an online self-study environment, so the next step was to investigate how effective different repetition practice methods are on enhancing the acquisition process of accurate vowel length. This was research question 2, and it was examined through hypotheses 4 to 6.

Hypothesis 1 attested that the scores from pretest to practice improved for all three treatment groups, and that regardless of treatment types, learners are able to mimic the mora length accurately immediately after listening to the model. In Hypothesis 4, a comparison among the groups was conducted to investigate which treatment group(s) shows the most improvement from pretest to practice immediately after listening to the model. The result showed that there were significant differences between the Audio-Only and In-Class groups, as well as between the Audio-Visual and In-Class groups, but a significant difference was not found between the two groups that received their treatment online ($p$=0.768). Therefore, it can be concluded that learners are able to model the word better immediately after listening to it when they are practicing in an online self-study environment than when they practice in a classroom setting.

One possible explanation for the difference might be that the online practice groups' participants were able to listen to their own utterance, and compare that with the model. As explained earlier, the online practice groups also listened and repeated after the model twice, just as in the In-Class group. It could be thought that from the first time they repeated after the model and the second time they listened to the model, they were able to hear the differences between their own utterance and the model, and that is why they were able to produce it better the second time they repeated. On the other hand, it may be the case for the In-Class group that the participants were not able to compare their

utterance well enough with the model because other participants were repeating the words at the same time, which may have distracted them or made it difficult for them to hear their own utterance clearly. This could explain why the In-Class group did not make as much improvement in their vowel length production as the two online practice groups.

Hypotheses 2 and 3 showed that there were significant improvements in their vowel length production from pretest to posttest and from pretest to delayed posttest for the two online practice groups (Audio-Only and Audio-Visual), while no significant improvement was found for the In-Class group. Hypotheses 5 and 6 were tested to further analyze whether there would be a significant difference in the gained score both in terms of short and long term improvement among the three different treatment groups. It was hypothesized that the Audio-Visual group would show the greatest short and long term improvement since it received a visual aid in addition to the model audio that the other two groups also did.

However, the result from hypotheses 5 and 6 showed otherwise, and a significant difference was not detected in the gain scores from both pretest to posttest ($p$=0.306) and from pretest to delayed posttest ($p$=0.274) among the three groups. At the level of descriptive statistics, the differences in gains among the groups came in the hypothesized order (In-Class<Audio-Only<Audio-Visual) both from pretest to posttest and from pretest to delayed posttest (see Table 23). However, the differences were not enough to reach statistical significance.

Table 23: Gain Scores from Pretest to Posttest, and from Pretest to Delayed

| | Audio-Only | | | | Audio-Visual | | | | In-Class | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Gain | Mean | SD | Min | Max | Mean | SD | Min | Max | Mean | SD | Min | Max |
| Pretest-Posttest (H5) | 4.7 | 2.458 | 2.0 | 9.5 | 5.1 | 4.640 | −2.5 | 13.5 | 2.1 | 2.529 | −1.8 | 5.0 |
| Pretest-Delayed (H6) | 2.4 | 2.723 | −3.8 | 5.5 | 3.5 | 3.649 | −1.8 | 10.0 | 0.7 | 2.485 | −2.3 | 4.3 |

From the findings from hypotheses 4 to 6, research question 2 can be answered. From pretest to practice, the groups that showed the most improvement in controlling vowel length were the two groups that received the repetition practice online, but it was not determined which online group did better than the other.

**Research question 3.** The final hypothesis was constructed in order to answer research question 3; "Do different input methods influence learners' retention?" In the same way as hypotheses 4 to 6, it was hypothesized that participants in the Audio-Visual group will show the greatest retention since they received visual aid in addition to the model audio, which was thought to help internalize the vowel duration. It was hypothesized that the Audio-Only group would show the next greatest retention, and the In-Class group would show the least retention among the three groups.

The result showed that there were no significant differences in retention among the three treatment groups ($p$=0.833), and therefore, did not confirm the final hypothesis.

**Discussions**

Although the first four hypotheses were all partially confirmed, the last three hypotheses were not confirmed, contrary to expectations. Possible reasons for these results are discussed below.

**Visual aid.** First, the "karaoke-model" visual aid that was used for the Audio-Visual group may not have been salient enough to affect the learner's perception of the mora length. The result may have been different if other visual aids that had been shown to be effective in other studies had been combined with the "karaoke-model." For example, showing a video where the model is saying a word next to the "karaoke-model" screen could have facilitated the acquisition process more since studies show that lip movement influences humans' perception of sound (McGurk & MacDonald, 1976). Also, the "karaoke-model" visual aid may not have been enough to create awareness of the concept of "mora." If the participants in the Audio-Visual group had first received a short lecture about the concept of mora, and understood that Japanese is mora-timed and is different from their L1 (e.g. stressed-time languages, syllable-timed languages), they may have been able to control the vowel length better after watching the "karaoke-model" visual aid.

**Length of training.** Another reason why different treatment types did not show differences in their effectiveness may be that the acquisition process to internalize accurate mora length in their sound system simply takes time. Unlike Japanese pitch-accent, acquiring accurate mora length is not about memorization. To acquire correct pitch-accent, learners need to memorize whether the word has an accent or not and if it does, where the accent falls on each word. However, once they know this, they can predict the pitch pattern of the word. Therefore, pitch-accent acquisition is strongly correlated with memorization, and a practicing method that can facilitate the learners' memorization process is needed. On the other hand, once learners acquire the system of

mora-timing, they should be able to pronounce many words accurately in terms of mora durations in Japanese.

There were some participants in the Audio-Only and Audio-Visual group who made great improvement in their score from pretest to posttest and pretest to delayed posttest. In the Audio-Only group, there were two participants who improved more than 8 points from pretest to posttest, and three participants from the Audio-Visual group improved as much. From pretest to the six-week delayed posttest, one participant in the Audio-Only group gained 5.5 points, and there were three participants from the Audio-Visual group who gained at least 5.5 points. One participant in the Audio-Visual group even gained 10 points from pretest to delayed posttest. Seeing these few participants' results where they made a significant improvement from the pretest, with the most participants being from the Audio-Visual group, it may be plausible that these participants acquired short-term, if not long term mora length awareness after practice, and because they did, they were able to improve their scores to a large degree from the pretest. They may not have been able to pronounce many words accurately in the pretest because they did not have this mora length awareness in their sound system, but it could be that once they got it after the practice, they made a significant improvement in their overall scores.

Looking at these participants' scores, it could be said that practice, especially the Audio-Visual group's practice, may have helped to improve their mora length awareness. However, acquiring this "sense" could take time and a large amount of practice may be needed to facilitate the acquisition process.

The idea that acquiring the system of mora-timing takes time has been discussed in previous studies as well. Previous studies show that even learners who have reached the advanced or superior-level often times have difficulty in pronouncing long vowels (Nagai, 1997; Oguma, 1999; Toda, 1998). Given that, it seems unreasonable to expect the participants in this study to acquire correct vowel duration after practicing only once. If this is the case, it could be that the Audio-Visual treatment was more effective than the other treatments (as it could be seen at the level of descriptive statistics), but one-time practice using this method was not enough to show a difference in the improvement between the other groups.

**Location of the long vowel.** Lastly, the result may have been affected by the words that were chosen for this study. Out of the 25 words that were used in this study, there were 23 long vowels in total, of which 7 were located in the word-initial position, 2 in the word-middle position, and 14 in the word-final position. Oguma's study (2006) showed that it is easiest for learners to acquire long vowels when they are at the word-initial position and is most difficult when it is located at the word-final position, and more than two-thirds of the words used in this study had long vowels located at the most difficult, word-final position. If the majority of the words used in this study had long vowels either at the word-initial or word-medial position, there may have been greater improvement on the overall score and a difference in improvement among the treatment groups may have been detected. However, in order to find that out, a longitudinal study would be required to see how long it takes for learners to acquire the long vowels at different positions, and if different methods facilitate the acquisition process more than the others.

## Summary of Chapter 4

This chapter presented research findings and discussed the possible reasons for the obtained result. In the final chapter, pedagogical implications, limitations of the study, and lastly, directions for future studies are provided.

# CHAPTER 5: CONCLUSIONS

## Introduction

This study has investigated the effectiveness of repetition practice for learners' acquisition of special moras, focusing on long vowel production, and compared the different practicing environments as well as different practicing methods in order to find which environment and method enhances the acquisition process the most. This final chapter will summarize the research findings and also discuss pedagogical implications, limitations of the study, and directions for future studies.

## Research Findings

First, statistical analysis showed significant improvement in participants' vowel duration from pretest to practice in all three groups that received listen-and-repeat practice. This means that regardless of input modality, learners can make their vowel duration sound closer to the model immediately after listening to it, and that listen-and-repeat practice is effective in introducing the accurate vowel length to the learners' sound system.

Also, significant improvement was observed between pretest and posttest and pretest and delayed posttest for both groups that received their treatment online, but not

in the In-Class group. This means that participants who practiced online were able to retain the accurate vowel length immediately after and six weeks after they practiced, while participants who practiced in class could not. This study showed that just having one time listen-and-repeat practice on vowel duration has significant short and long term improvement in their vowel length production skills when the practice is done in an online self-study environment, but practicing in an in-class environment is not enough to ensure learners' improvement in their vowel duration.

Second, among the three practice methods, Audio-Only and Audio-Visual groups showed significantly greater gains than the In-Class group between the pretest and practice. This might have been due to the fact that the participants in the online practice groups were able to listen to their own utterance and compare that with the model better than the In-Class group, since they were practicing in a self-study environment where they could hear the differences between their own utterance and the model more clearly compared to when there are other learners repeating at the same time. However, a significant difference in the gain score was not found between the Audio-Only group and the Audio-Visual group.

From pretest to posttest and pretest to delayed posttest, statistical analysis did not show significant difference among groups, although the descriptive statistics came in the hypothesized order (In-Class<Audio-Only<Audio-Visual). It was expected that the Audio-Visual group would show the most improvement in controlling vowel length because it received additional visual aid in addition to the model audio, which was thought to help internalize the vowel duration and would show a clear correspondence between the sound and the Japanese writing system. The reason for the unexpected

outcome might have been caused by the lack of repeated training and lack of salience of the “karaoke-model” visual aid.

Third, it was examined whether different input methods influence learners’ retention. It was also expected that the Audio-Visual group would show the greatest retention from the same reason, but the statistical analysis showed no significant difference in retention from posttest to delayed posttest among the three groups. The reason for the unexpected result may be the same as the above, and therefore, further investigation after reconstructing a visual aid with repeated training may be needed to see the differences among treatment types.

**Pedagogical Implication**

One of the aims for the present research was to compare the effectiveness of repetition practice done in an online self-study environment and in-class environment, and a significant short and long term improvement was only found in learners who practiced in an online self-study environment. When instructors have their students engage in repetition when they introduce new words in class, they may feel satisfied with incorporating a small pronunciation practice in class, as a way of justifying their action as an instructor. However, since it was found that practicing in class may not be very helpful in improving learners’ vowel duration, instructors should assign this type of pronunciation practice as a homework assignments, which would be beneficial for both learners and instructors. This is beneficial for learners because when they do these repetition practices at home, it enhances their vowel length production skills both in short and long term more than doing it in class. This is also beneficial for the instructors

because by applying online practice outside of class, the limited class time can be used for doing other communicative activities that can be done only in the classroom. Also, by making it an individualized online practice, pronunciation practice can be incorporated into the curriculum even if instructors do not know how to teach pronunciation.

### Limitations of the Present Study and Future Possibilities

The present research was designed to investigate whether learners' long vowel production ability improves after receiving listen-and-repeat practice, and whether learners are able to pronounce the words containing long vowel(s) from only looking at the word, and not from memorization. For the listen-and-repeat practice, different input modalities were used and were conducted in different environments to compare the effectiveness of the repetition practice. That is why participants were randomly divided into three different treatment groups, and that only unlearned words were chosen for this study. However, there are five limitations to this study that need to be disclosed:

**Significant difference in pretest means.** As mentioned in Chapter 4, a significant difference in pretest means were found when One-way ANOVA was conducted, in spite of the randomized group assignment. However, since all the participants for this study were volunteers, it was not possible to re-divide the participants into three groups and have them do the pretest again. Therefore, a desired research design would be where the participants could be re-divided into different groups (if needed) after obtaining the result from the pretest.

**Number of participants.** Due to the restricted environment of the research, there were only 27 learners who participated in this study, and the results of this research could have been largely affected by the number of participants. Although significant improvement differences among treatment groups were not attested in this study, the result of the descriptive statistics showed that the Audio-Visual group improved the most, as hypothesized. Therefore, this experiment should be tested again with a larger sample size.

**Length of training.** Participants in this study received training on producing accurate vowel length only once, and this one-time practice many not have been enough to promote mora awareness, as studies show that acquiring accurate vowel length production takes time, as mentioned in Chapter 4. Therefore, longitudinal study with repeated training is needed for future research.

**Other visual input modalities.** Also discussed in Chapter 4, improvement differences among the treatment groups may have been observed if the Audio-Visual group received visual input in addition to the "karaoke-model" visual aid. Future research should compare the effectiveness of the different treatment types again after constructing another visual aid in addition to the one that was used in this research.

**Target word selection.** The location of the long vowels of the 20 target words used in this study was not equally distributed, and this may have affected the overall score of the improvement. Since studies show that the acquisition process of long vowels differs according to the location of the long vowel, as mentioned in Chapter 4, it would be interesting to compare the time it takes to acquire long vowels at certain locations with the aid of visual input.

## Future Studies

For future studies, the following three factors should be included:

**Feedback.** Although participants in this study did not receive any feedback during their training, the two groups that received training online showed significant improvement in their vowel duration from pretest to practice, pretest to posttest, and pretest to delayed posttest. Since they were able to make significant improvement even without any feedback, it would be interesting to see how much more improvement participants practicing in an online self-study environment would be able to make if they received feedback.

**Perception training and test.** The present study only focused on the production of long vowels and not on the perception of long vowels. However, as mentioned in Chapter 2, previous studies discuss that perceptive ability is strongly correlated to productive ability. In future studies, it would be interesting to examine whether the instruction used in this present study might also help enhance learners' perceptive ability.

**Various proficiency levels.** Since it is important for learners to create awareness of special moras and to practice producing it at an early stage, the participants for this present study were limited to beginning-level learners. However, future studies should compare the effectiveness of online self-study repetition practice on different levels, since studies show that many learners have difficultly acquiring long vowels regardless of their proficiency level. It would also be beneficial to determine which level of learners can benefit the most from online self-study repetition practice.

## Final Conclusion

The results of this study suggest that vowel duration training using listen-and-repeat practice is most effective when it is conducted in an online self-study environment. This study showed that repetition practice is beneficial for producing accurate vowel duration, although this type of practice is often times avoided among language instructors. However, instructors who are hesitant in incorporating repetition practice in class do not necessary have to, because this type of practice could be easily done using CALL programs such as *Speak Everywhere*. In fact, making this an out-of-class activity would be more beneficial for the learners as well.

While much research must be done on the acquisition of vowel duration, and that future research should be expanded on the aspects listed in the limitation section, it is hoped that the results obtained from this study will encourage instructors to incorporate pronunciation practice into language curricula.

# LIST OF REFERENCES

# LIST OF REFERENCES

Abbs, B., Prahlad, G., & Khetarpal, N. (2008). Is overt repetition critical to expressive word learning? The role of overt repetition in word learning with and without semantics. *Applied Psycholinguistics, 29*(4), 627-667.

Bransford, J., Brown, A., & Cocking, R. (1999). *How people learn: Brain, mind, experience, and school*. National Academy Press.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2013). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge University Press.

DeKeyser, R. (Ed.). (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge University Press.

Enomoto, K. (1992) Interlanguage phonology: The perceptual development of durational contrasts English-speaking learners of Japanese, *Edinburgh Working Papers in Linguistics, 3,* 25-35.

Fukada, A. (2013). An online oral practice/assessment platform: Speak Everywhere. *The IALLT Journal, 43*(1), 64-77

Hirata, Y., & Kelly, S. D. (2010). Effects of lips and hands on auditory learning of second-language speech sounds. *Journal of Speech, Language, and Hearing Research*, *53*(2), 298-310.

Iba, M. (2008). The Influence of Model Sounds on the Speech Production of Japanese Learners of English. *Gengo To Bunka*, *12*, 45-66.

Kawai, G., & Hirose, K. (2000). Teaching the pronunciation of Japanese double-mora phonemes using speech recognition technology. *Speech Communication*, *30*(2), 131-143.

Lambcher, S. (1999). A CALL tool for improving second language acquisition of English consonants by Japanese learners. *Computer Assisted Language Learning*, *12*(2), 137-156

Lapkin, S., Hart, D., & Swain, M. (1995). A Canadian interprovincial exchange: Evaluating the linguistic impact of a three-month stay in Quebec. *Second Language Acquisition in a Study Abroad Context,* 67-94.

Leaver, B. L., Rifkin, B., & Shekhtman, B. (2004). Apples and Oranges are both fruit, but they don't taste the same: A response to Wynne Wong and Bill VanPatten. *Foreign Language Annals, 37*(1), 125-132.

Matsuzaki, H. (2001). Nihongo no Onsei Kyouiku. *Computer onseigaku,* 207-258.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264,* 746-748.

Meara, P. (1994). The year abroad and its effects. *Language Learning Journal, 10.1,* 32-38.

Minagawa, Y. (1995). Kankokujin gakushuusha ni okeru akusento, tokushuhaku choushunouryoku to gakushuu nensuu ni yoru sono henka. *Dai Juunikai Daini Gengo Shuutoku Kenkyuukai Haihu Shiryou*.

Motohashi-Saigo, M., & Hardison, D. M. (2009). Acquisition of L2 Japanese geminates: Training with waveform displays. *Language Learning & Technology, 13*(2), 29-47.

Nagai, K. (1997). Eikokujin gakushuusha ni yoru nihongo hatsuon no tokuchou ni tsuite. *Nihon Onsei Gakkai Zenkoku Taikai Yokoushuu,* 53-58.

Ofuka, E. (1997). *Teaching pronunciation of Japanese long and short vowels: Kinesthetic and computer-assisted visualization approaches.* (Unpublished master's thesis). University of Toronto, Toronto, Canada.

Oguma, R. (2000). Eigo bogowasha ni yoru chouon to tanon no chikaku. *Sekai No Nihongo Kyouiku, 10,* 43-55.

Oguma, R. (2001a). Nihongo gakushuusha no chouon sanshutsu ni kansuru shuutoku kenkyuu - chouon ichi ni yoru nanido to shuutoku junjyo. *Nihongo Kyouiku, 109,* 110-117.

Oguma, R. (2001b). Nihongo gakushuusha ni yoru chouon to tanon no sanshutsu – hatsuon ni taisuru chuuido ga oyobosu eikyou -. *Takushoku Daigaku Nihongo Kiyou, 11,* 79-97.

Oguma, R. (2002). Nihongo no chouon to tanon ni kansuru chuukan gengo kenkyuu no gaikan. *Gengobunka To Nihongokyouiku Tokushuugou,* 189-200.

Oguma, R. (2006). Shizen hatsuwa ni mirareru nihongo gakushuusha no chouon to tanon no shuutoku katei. *Sophia Linguistica, 54,* 193-205.

Okuno, T. (2013). *Acquisition of L2 vowel duration in Japanese by native English speakers* (Unpublished doctoral dissertation). Michigan State University, East Lansing, Michigan.

Paulston, C. B. (1974). Linguistic and communicative competence. *TESOL quarterly*, *8*(4), 347-362.

Schmidt, R. (1993). Awareness and second language acquisition. *Annual review of applied linguistics*, *13*, 206-226.

Teichler, U., & Maiworm, F. (1997). *The ERASMUS experience: Major findings of the ERASMUS evaluation research project.* Office for Official Publications of the European Communities.

Toda, T. (1998). Nihongo gakushuusha ni yoru sokuon, chouon, hatsuon no chikaku hanchuuka. *Bungei Gengo Kenkyuu Gengo Hen, 33*, 65-82.

Toda, T. (2003). Acquisition of special morae in Japanese as a second language. *Onsei kenkyuu,7(2),* 70-83.

Toda, T. (2008). *Nihongo kyouiku to onsei*. Kuroshio Shuppan.

Tsuchiya, J. (1992). Gaikokujin gakushuusha no nihongohatsuwa ni okeru gomatsu no boin no nagasa no hikaku taishou. *Monbushou Jyuuten Ryouiki Kenkyuu Heisei Yonendo Kenkyuuseika Houkokusho,* 131-142.

Tsujimura, N. (2013). *An introduction to Japanese linguistics*. John Wiley & Sons.

Uchida, T. (1993). Chuugokujin nihongo gakushuusha ni okeru chouon to sokuon no choukakuteki ninchi no tokuchou. *Kyouiku Shinrigaku Kenkyuu*, *41*(4), 414-423.

Wong, W., & Patten, B. (2003). The evidence is IN: Drills are OUT. *Foreign Language Annals*, *36*(3), 403-423.

Yokoi, K. (1998). Gakushuusha no hatsuonjou no mondaiten shiteki to sono kouka - daikibo kurasu wo taishou ni shita hatsuon chekku to fiidobakku -.*Heisei Juunendo Nihongokyouikugakkai Shunkitaikai Yokoushuu-*, 154-164.

Yoshida, K., & Fukada, A. (2014). Effects of Oral Repetition on Learners' Japanese Word Accentuation. *The IALLT Journal for Language Learning Technologies*, 44(1),17-37.