THE EFFECTIVENESS OF AUDIO-VISUAL TRAINING ON NON-NATIVE ENGLISH SPEECH PERCEPTION AND PRODUCTION

by

CHIA-NI SHEN

A THESIS

Presented to the Department of Linguistics, Communication Disorders & Sciences
and the Robert D. Clark Honors College
in partial fulfillment of the requirements for the degree of
Bachelor of Arts

June, 2019

An Abstract of the Thesis of

Chia-Ni Shen for the degree of Bachelor of Arts
in the Department of Psychology and Communication Disorders & Sciences to be taken
June, 2019

Title: The Effectiveness of Audio-Visual Training on Non-Native Speech Perception
and Production

Approved: _____

Dr. Melissa Baese-Berk

This Project examines the effectiveness of audio-visual training on non-native

English speech perception. Previous research utilizing audio-visual training has been

employed in the field of speech pathology, showing positive outcomes on improving

speech among dyslexic children. However, few studies to date have examined its use in

second language learning, specifically labial and labiodental consonants (i.e., /b, p, m, f,

v/), which are known to be challenging for many second language learners. The aim of

this project is to explore audio-visual training across 3 native language groups,

Mandarin, Japanese, and Arabic, who are all English language learners. Participants

undergo a training regimen designed to examine the effects of audio-visual and audio-

only training. We hypothesized that perception performance will improve after two days

of training and will improve more after audio-visual training and 2) perception

improvement will rely heavily on a participants' language background and known

difficulties with specific labial and labiodental sounds. Results from this study showed

audio-visual training having no effect on Mandarin and Arabic speakers but an effect on

Japanese speakers. Additional results suggest that the effectiveness of audio-visual

training may be language-specific, which may have pedagogical implications on its use

in second language teaching and learning.

Acknowledgements

I would like to thank Professor Melissa Baese-Berk for trusting and welcoming me into her lab and for providing me with tremendous guidance and support throughout my research journey; Professor Louise Bishop for her continuous encouragement and thoughtfulness, for believing in me always and imparting me with daily wisdom; and Dr. Karen McLaughlin for her infectious smile and enthusiasm during office hour visits and reminding me that you're never too old to play with play-doh even in a college neuroanatomy class.

Thank you to the Undergraduate Research Opportunities Program for the financial support in making this research a reality. Also a huge thank you to the American English Institute, Karen, Deborah, and many teachers on their generous support and collaboration, for opening up their classrooms and allowing students to participate in this research.

I would like to express thanks to Farah Abu-Khater, Ayami Iijima, and Menju Lin for their help in translating the research documents into their respective languages; Jennifer Bennett for modeling in the training video, and research assistants Tellena Trebon and Aubrey Whitty for running subjects.

Lastly, I would like to thank the many friends and family who showered me with endless love and moral support, especially to my mom and dad for always having faith in me and cheering me on even from halfway across the world.

Table of Contents

List of Figures

List of Tables

Non-Native Speech Sounds: A Problem

Learning a new language is not an easy task. Human perception of speech, or our ability to process and understand language through listening and/or seeing, is highly influenced by linguistic experience. Among adults, the native language (L1) carries substantial weight on how one processes non-native speech sounds. Identification of non-native speech sounds, or phonemes (a unit of sound that distinguishes one word from another in a specific language; e.g., the words pat versus bat in English are distinguishable by phonemes /p/ and /b/), may become difficult when the target second language (L2) phoneme is mistakenly confused with one in L1. For example, a typical native English speaker may perceive words "pat" and "bat" as two different words, but an English as a second language learner may perceive the two words as "bat" and "bat". Instead of perceiving the two phonemes, /p/ and /b/, as being qualitatively different, the second language learner perceived both as being qualitatively the same where /p/ was ascribed as being the same as /b/. Another common example is Japanese English learners' perception of the English /l/ and /r/. Second language learning therefore, becomes a difficult task when the learner is unable to make appropriate phoneme discriminations. Inability to make appropriate phoneme discriminations may have negative implications on second language learner's abilities to comprehend and communicate effectively with native speakers of the L2. Rise in communication difficulties, in turn, may have influence on non-native speakers' abilities to engage and participate in L2's culture. Second language acquisition research does show, however, that non-native perceptual difficulties may be reversible by extensive perceptual training.

Research has shown that non-native perceptual difficulties may be improved through intensive conversational training (MacKain, Best, and Strange, 1981), high variability laboratory training (Logan, Lively, Pisoni, 1991, 1993; Lively et al., 1993, 1994; Bradlow, Pisoni, Yamada, Tohkura, 1997; Shinohara & Iverson 2018), and more recently, audio-visual training (Kartushina, Hervais-Adelman, Frauenfelder, Golestani, 2015; Peng, Chen, Wang L., Wang H., 2018; Lim & Holt, 2011; Wang, Spence, Jongman, Sereno, 1999; Wang, Behne, Jiang, 2008, 2009; Hazan, Sennema, Faulkner, Ortega-Llebaria, Iba, Chung, 2005, 2006; Sekiyama, 1997), where audio-visual training was more effective on improving non-native speech perception than auditory alone training (Hazan et al., 2005, 2006, 2010).

Perception of speech is both an auditory and visual task. Incorporation of audio-visual training is useful since it combines the use of audio and visual cues and creates speech representations similar to what learners may encounter in face to face communication. Contrary to auditory alone training, visual cues may carry important linguistic information and add to learners' overall understanding and perception of an L2. Research have shown that lip-reading can provide valuable information when speech is unclear. This is demonstrated in speech-in-noise identification tasks where audio-visual training was superior in speech intelligibility than auditory alone (Lidestam, Moradi, Pettersson, Ricklefs, 2014; Sumby & Pollack, 1954). Evidence of the visual stream containing important speech information in speech processing is also shown in the McGurk effect. The McGurk effect is a perceptual phenomenon where visual cues for one sound (/ka/) is dubbed with audio cues for another (/pa/), and results

in the perception of a third sound (/ta/) (Mcgurk & Macdonald, 1976). In sum, visual cues can have powerful influence on one's perception of speech. This may be critical in second language perception during situations where auditory information is not reliable in distinguishing speech sounds. Visual information, therefore, may be supplemented to aid learner's understanding of unfamiliar speech sounds.

Previous investigation of audio-visual training have been employed in the field of speech pathology and second language acquisition of the English /r/ and /l/ among Japanese speakers. However, little research has employed its use on more visually salient phonemes, labials and labiodentals. The current research will be used to bridge this gap.

Background

Categorical Perception

How a listener perceives and identify sounds is important in the understanding of second language acquisition. One of the key characteristics of how listeners organize speech sounds is through categorical perception. Categorical perception is the perception that different sensory phenomena are qualitatively, or categorically different. In Liberman, Harris, Hoffman, and Griffith (1957), researchers used a speech synthesizer and created a series of syllables panning categories /b/, /d/, /g/, followed by /a/. The researchers varied the speech-like sounds in small steps along an acoustic continuum and asked English listeners to label them as b, d, or g. Results showed that listeners tend to divide the continuum into three sharply defined phonemic categories. In a second experiment, researchers measured listeners' abilities to discriminate between sounds. Results showed that listeners were better at discriminating sounds between categories than within categories. In other words, sounds that have different phonemic labels are typically easier to distinguish than those in the same phoneme class. Reports on other languages revealed that phonemic categories were formed at different boundaries along the b-d-g continuum. In sum, while speech sounds could be many and various, listeners tend to reduce the many sounds they hear into distinct categories already available in their native phonemic categories. Discrimination of these sounds are much easier when the sounds are between categories, identified as being qualitatively different, than within categories, identified as being qualitatively the same. However, because the phonemic boundaries formed along the continuum are different

with each language, a problem may arise when listeners try to discriminate sounds in a second language.

Perceptual Assimilation Model (PAM)

One theoretical model that underlies perception of non-native speech is Best's (1995) Perceptual Assimilation Model (see figure 1). According to PAM, perception of L2 sounds are based on their gestural (articulatory organs, place of articulation, manner of articulation) similarity to an existing native phoneme and can either be assimilated, uncategorized, or non-assimilable. The degree to which L2 sounds are correctly discriminated after assimilation vary from poor to excellent depending on the goodness of fit and assimilation type. Best (1995) posits that there are 5 patterns of assimilation that occurs during assimilation of contrasting non-native phones: 1) single-category (SC) assimilation, where both L2 sounds are assimilated into one category and are perceived as either a good, acceptable, or deviant exemplar of the L1 category, 2) two-category (TC) assimilation, where two L2 sounds are assimilated to two different L1 categories, 3) Category-Goodness (CG) assimilation, where both L2 sounds are assimilated to the same L1 category, but with varying goodness of fit (one L2 sound may be a better exemplar of the L1 than the other), 4) uncategorized-categorized, where one L2 sound is assimilated to an L1 category, but the other is not, 5) uncategorized-uncategorized assimilation, where neither L2 sound is assimilated into any of the L1 category.

Research interested in PAM have primarily focused on three assimilation types: SC, TC, and CG. A common example of SC assimilation is of Japanese speakers

assimilating the English /l/ and /r/ contrasts to the Japanese /r/ (Bradlow et al., 1997). Typically, the discrimination scores for this assimilation type have been very low since both sounds are perceived as being phonetically equal. An example of TC assimilation is of French speakers' assimilation of the English /ð/ and /θ/ contrasts as /d/ and /t/ (Jamieson and Morosan, 1986). Discrimination scores for TC assimilation have usually been high since the two sounds being compared are labeled as categorically different. Example of CG assimilation is of French speakers' assimilation of the English /w/ and /r/ contrasts and assimilating the English /r/ as being /w/-like (Hallé, Best, and Levitt, 1999). Discriminations scores for CG can be low, but usually not as low as SC because there is a goodness difference (one sound is a good fit, the other a poor fit); whereas, there is no goodness difference in SC (both sounds are equally a good fit, or a deviant fit).

Lastly, non-assimilable occurs when listeners perceive the sounds as non-speech sounds and do not assimilate the sounds into the native phonological space. An example would be English listener's perception of Zulu clicks as being non-speech sounds (Best & Avery, 1999). Results show that English listeners discriminate the Zulu clicks quite well.

Overall, the PAM model illustrates the native phonological system of having strong influence on a listeners' perception of an L2. While discrimination of L2 sounds may be good when the goodness of fit of the assimilation is excellent, poor goodness of fit may result in increased discrimination problems and perceptual difficulties. Therefore, assimilation of an L2 sound into an L1 category may not always be the best

6

representation of what the L2 sound is perceptually like to a native speaker and second language learners may have to explore other ways to make the distinctions clearer.

Figure 1. Best's Perceptual Assimilation Model (PAM)
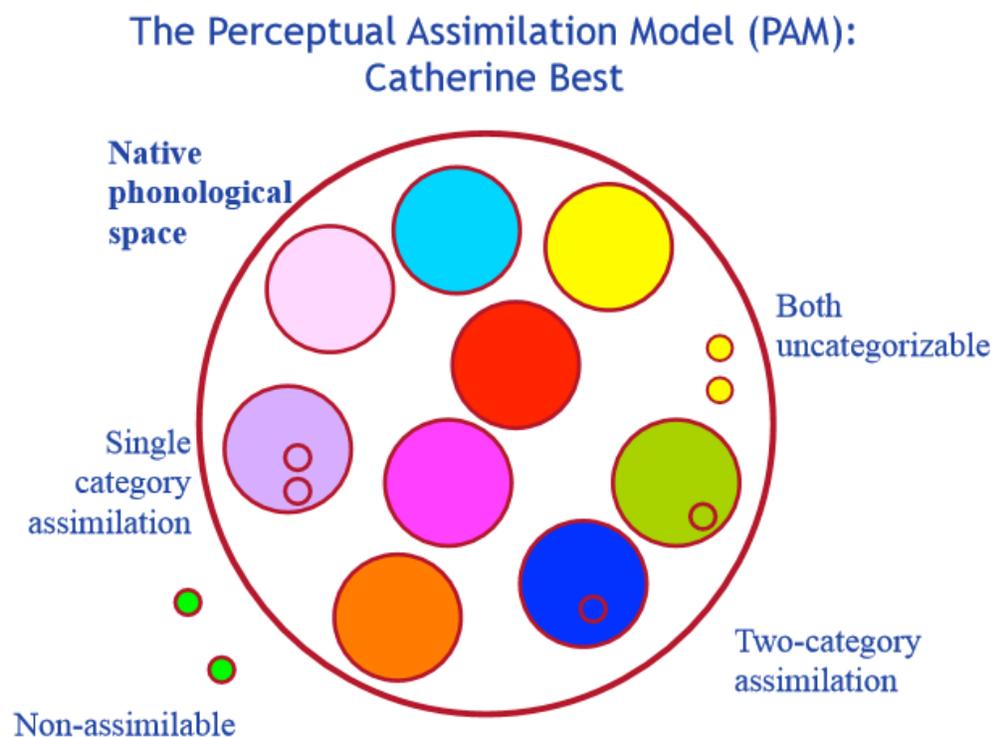
Labials/Labiodentals

The place of articulation plays an important role in how speakers perceive sounds. Place of articulation is where articulation of a consonant occurs. Consonants that are produced by the lips (labials, labiodentals) are easier to distinguish than consonants produced further in the vocal tract (see Figure 2). Labials are consonants made with two lips such as the English /p/, /b/, /m/ in words pie, buy, and my; whereas, labiodentals are consonants made with the lower lip and upper front teeth such as the English /f/ and /v/ in words fie and vie. Overall, increased awareness of how the

7

articulators move during voice productions may aid second language learner's perception and production, especially when specific phonemes are absent in the learner's L1. Labial and labiodental consonants are a good start because of their high visual saliency.

Figure 2. Illustration of labial and labiodental place of articulation in the vocal tract.

Labial place

Bilabial:
     p, b, m
Labiodental:
     f, v
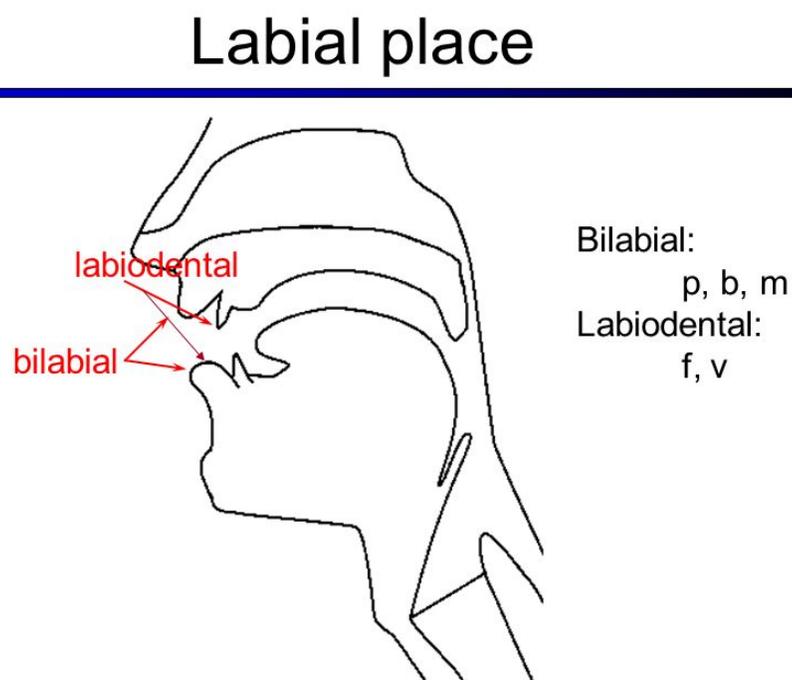
Figure thanks to Jennifer Venditti

The distribution of labial and labiodental consonant phonemes differ across phonemic inventories of each language. These differences are illustrated in Table 1 across English, Arabic, Mandarin, and Japanese.

	English		Arabic		Mandarin		Japanese	
	Labial	Labiodental	Labial	Labiodental	Labial	Labiodental	Labial	Labiodental
Plosive	p b		b		p pʰ		p b	
Nasal	m		m		m		m	
Fricative		f v		f		f	ɸ	

Table 1. Phonemic Inventory of labial/labiodental contrasts across English, Arabic, Mandarin, and Japanese. This table was adapted from phonemic inventories available on American Speech-Language-Hearing Association's website)

In the English phonemic inventory, labial consonants /p/, /b/, /m/ and labiodental consonants /f/ and /v/ are present. However, not all of these consonants are present in other languages. In the Arabic phonemic inventory, there are no English consonants /p/ and /v/. In the Mandarin phonemic inventory, there are no English consonants /b/ and /v/. Lastly, in the Japanese phonemic inventory, there are no English consonants /f/ and /v/.

Previous second language acquisition research have shown Arabic speakers having difficulties discriminating between /b/-/p/ and /v/ contrasts (Binturki, 2001) and Japanese speakers between /b/-/v/ contrasts (Yoshida & Hirasaka, 1983) and Mandarin speakers with /v/ (Zhang & Hayashi, 2015); however, the amount of research dedicated to investigating these consonants is meager compared to the extensive research

9

conducted in other areas such as Japanese speaker's discrimination of /l/ and /r/ (Logan et al., 1991, 1993; Lively et al., 1993, 1994; Bradlow et al., 1997; Shinohara et al., 2018) and Mandarin tones and pronunciation (Wang, Spence, Jongman & Sereno, 1999; Wang, 2013; Peng, Chen, Wang L., & Wang H., 2018). Further investigation of perceptual difficulties experienced with labial/labiodental consonants would add to the field.

The phonemic inventory illustrates the range of sounds that listeners are typically exposed to in their native language, but also the possible assimilations that may arise of not having an L2 sound. Although cross examination of phonemic inventories may not always accurately predict learners' perceptual difficulties, it does give researchers and educators a better idea of learner's base linguistic knowledge.

Use of Audio-visual Training

Computer-based training and audio-visual software are now steadily used across various linguistics and speech pathology research. Audio-visual training is pairing audio material with visual material. One of the basic advantages of using audiovisual training is the ability to develop better phonemic awareness through incorporation of sounds and digitized speech alongside words and high-quality graphics (Magnan et al., 2004). The development of this resource provides researchers opportunities to make research designs more creative and interactive. This is especially helpful in developmental research, since it allows researchers to develop game-like presentations to maintain child interest while still administering immediate feedback. The primary use of audio-visual training typically involves a listening and reading task. Application of this resource has widely been explored in speech pathology, specifically dyslexia. Dyslexia is a disorder that involves difficulties learning to read or interpret words. One hypothesis suggest that dyslexia is the result of having phonological deficits, the inability to discriminate acoustic cues important in distinguishing phonemes apart. This was explored in various dyslexia research. In Magnan et al. (2004), the study administered daily audiovisual voicing exercises to children 30 minutes a day, 4 days a week, over 5 weeks. Participants were assessed on reading task before and after the training. The ending results showed that training in ortho-phonological representations significantly improved the participant's reading ability. This is consistent with Veuillet et al. (2007), who assessed auditory deficits, altered voicing sensitivity, in children with dyslexia through a categorical perception test using bilabials /ba/ and /pa/. Results from this study showed that auditory deficits are reversible by training and that audiovisual

training was effective in improving reading comprehension among dyslexic children. The summation of the two studies demonstrates the effectiveness of audiovisual training in improving speech perception, such as categorical perception, among dyslexic children. Implications from these studies have surprising connections to second language learning, for one of the main issues second language learners face when learning a new language is also issues with categorical perception. Application of audiovisual training in second language acquisition would be helpful in evaluating whether use of audiovisual training have implications beyond speech pathology. Further, it would be interesting to see if the training also offers learning effects in production.

There are robust studies on second language acquisition and non-native speech, many of the studies present focus on exploring perception and production of English /r/ and /l/ among native Japanese speakers. Many of these explore the use of direct visual feedback to enhance English perception and production. A study by Tateishi (2013) investigates whether production training using ultrasound as visual feedback leads to improved production and perception on /r/ and /l/ contrast. Results showed potential production improvement in /l/ but no perception improvements in either of the phonemes. Similarly, Hattori (2010) also explores this relationship but with real-time spectrograms and hand mirrors to help guide productions. But like Tateishi (2013), Hattori (2010) also only saw improvements in productions and not perception. These results assume the possibility that perception and production do not develop in tandem, and echoes results expressed in Baese-Berk & Samuel (2016), that productions and perceptions may have separate mental representations.

One study that specifically looks at perception and production through use of audio-visual training is Hazan et al. (2005). The study assessed perception and production of consonants: /b, v, l, r/ by Japanese learners of English before and after an audiovisual training. For the purpose of my research, I will evaluate the results of /b/ and /v/. Participants were divided into one of three trainings: audio training (only hearing), audiovisual training (hearing and seeing), and visual (only seeing) training. The recording used for each training featured a British English speaker saying contrasting minimal pairs aloud. Participants were asked to make judgement about the recording based on their training (did you hear, hear/see, or see "van" or "ban" or "pan"?). Participants continued to do this for 10 sessions throughout a 4-week period, with each session lasting about 40 minutes. Perception changes were measured by a pre/post-test featuring nonsense words with the two consonants. Results found that audio-visual training improved the perception of the labial/labiodental contrasts more than the auditory alone training. These results show that audiovisual training is effective in improving perception of English consonants that are marked visibly on the lips. However, there were study limitations. The study only focused on distinguishing between /b/ and /v/ due to voicing confusions between /b/ and /p/. This suggests consonant perception probably improved as a result of visual and not audio input. Further, the study also did not assess for participant productions of these consonants. Furthermore, the study also only utilized one population sample. Therefore, I plan to address these limitations in my research by using audiovisual training to:

1. Evaluate for all English labial/labiodental consonants /b, p, m, f, v/

2. Evaluate for when consonants are in the initial position versus final position

3. Evaluate participant perception and production before and after training

4. Incorporate a broader and more diverse population sample

Purpose

The purpose of this study is to assess non-native speech production and perception before and after audiovisual training, specifically labial and labiodental consonants (e.g., /b, m, p, f, v/), as these have previously been demonstrated to be challenging for many second language learners. Further, these consonants are visible on the lips of participants, though the differences between them are typically NOT visible. We will train and test participants in two language learning paradigms: audio-only and audio-visual. Participants will be tested to examine the effectiveness of each paradigm in improving second language learning, specifically on whether audiovisual training is helpful in second language perception and production.

Methods

The following research was approved by the International Review Board from the University of Oregon.

Participants

A total of 21 international students (ages 18 to 40, 11 female) with an L1 in Arabic (N= 4), Japanese (N=8), and Mandarin (N=9) recruited from the American English Institute (AEI) and the broader University of Oregon community participated in the study (see appendix D for recruitment script). All were non-native speakers of English and had only been in the states for an average of 7 months. Participants all had self-reported normal to corrected vision and hearing. Each were reimbursed 20 dollars for their participation in the study. One subject was later removed from the analysis due to inconsistent scores. Participants were randomly assigned into 1 of 2 training paradigms (audio-only vs. audio-visual) as illustrated in Table 2.

Table 1. Distribution of participants across the two training paradigms.

	Audio-Only (A)	Audiovisual (AV)
L1 Mandarin	4 (2 F, 2 M)	5 (3 F, 2 M)
L1 Japanese	4 (2 F, 2 M)	4 (2 F, 2 M)
L1 Arabic	2 (2 M)	2 (1 F, 1 M)
Total	10	11

The following study followed a similar methodology as the one used by Hazan et al., 2005, 2006; however, instead of assessing production and perception of liquids,

this study only focused on the production and perception of labials (/b/, /p/, /m/) and labiodentals (/f/, /v/) as few studies to date have focused comparing and contrasting the production and perception of these consonants, which are usually visually salient to native English speakers but not second language learners.

Questionnaire:

On the first day of the study, before beginning the experiment, participants filled out a survey asking about their basic demographic information and first and second language experiences (see appendix B). This included years spent learning English, and any language proficiency test scores, if they know or wish to provide them. Questionnaires were filled out on paper and matched to participation code only. These provided researchers better understanding of participants' background and experiences with other languages.

Consent and Debrief:

All participants were presented with a consent form (see appendix A) before the experiment and a debrief form (see appendix C) after the experiment. Consent and debrief forms were available in each of the participant's respective native language.

Pre- and Post-test materials

Real and nonsense words (the same set used in the training, see Appendix E) and a list of 12 sentences were used to evaluate consonant identification before and after the training (see Appendix F). The sentences featured the consonants /b/, /p/, /m/, /f/, and /v/ throughout various syllable structure (CV, VCV, and VC) and were used to

gather participants' base language perception and production ability. Each item was recorded once, yielding a total of 132 tokens. No instruction or feedback was given on the production of these sounds. Participants were asked to try their best to reproduce the tokens on their own without any help or input.

Training materials

24 sets of English minimal pairs (real and nonsense words) were recorded, with each featuring a labial/labiodental (/b/,/p/, /m/, /f/, /v/) consonant in either the initial or final position of a word (e.g., ban vs. cub, see Appendix E). There were 12 sets of initial position words and 12 sets of final position words. All minimal pairs within the same string followed the same syllable structure (ban vs. pan, not ban vs. back). Equal number of minimal pairs across all labial/labiodental consonants were created, for a total of 120 tokens.

Stimuli

A female native English speaker (age 21) enrolled in the University of Oregon recorded the pre/post-test and training items. Video recordings were recorded in a soundproof room located at the Speech Perception & Production Lab located in the Spoken Language Research Laboratories in Eugene. The video was taken on a clear backdrop, and proper lighting was adjusted to ensure good video quality. Because the study was interested in observing identification of labial/labiodentals, the video camera was zoomed in and aligned to only feature the speaker's mouth. No other facial features

were included in the video clips. Sound recordings were recorded on a lapel mic that was attached to both the talker and recording camera. Complete recordings were digitally transferred to computers in the lab, and later edited so that the start and end frames did not have prolonged pauses. Audio recordings were created after by converting video recordings to a .wav file with an audio sample rate of 44.1Khz

Pre- and Post-Test

Production:

Participants were presented with a list of experimental tokens (see Appendix E) on a sheet of paper and were asked to say each aloud without any input or help. Recordings were recorded using a blue yeti microphone connected to the computer. Complete recordings were saved individually on an audio analysis program called Praat (SOFTONIC INTERNATIONAL S.A. 2019). After completing the list of words, they were asked to read aloud the list of 12 sentences, which featured consonants measured in the training. The productions were recorded and saved again on Pratt for later acoustic analysis.

Perception:

This test was administered auditorily on the lab computers. Participants were engaged in a 2 Alternative Forced Choice (2AFC) identification task. An audio recording of a token was played alongside two words displayed on the screen, one on the left and one on the right. Participants were asked to select which one of two options

they heard by pressing keys "z" (word on the left) and "m" (word on the right). An example would be did you hear "pin" or "bin". Feedback was not provided. Response time and accuracy was measured.

Perception Training

The audio and audiovisual training paradigms were designed using PsychoPy (Peirce, Gray, Simpson, MacAskill, Höchenberger , Sogo, Kastman, Lindeløv, 2019) and later uploaded and run in sound proof rooms located in the Speech Perception & Production Lab. The recording stimuli were presented to participants through a headset while participant recordings were recorded through a usb connected yeti microphone. Presented stimuli were set at a good hearing rate and were adjustable according to participants' liking. After completing both parts of the pre-test, participants were assigned to one of two paradigms: audio-only (A) and audiovisual (AV), The training task was an adaptation of the High Variability Phonetic Training (HVPT) used across many identification studies (Logan, Lively, Pisoni, 1991; Bradlow, Pisoni, Akahane-Yamada, Tohkura, 1997; Hazan et al., 2005; Hazan, Sennema, Faulkner, Ortega-Llebaria, Iba, Chung, 2006). This was the same 2AFC perception task described above but with immediate feedback after each response (either "correct!" or "incorrect!"). There was no time limit to respond, but each stimuli was only presented once. A response was required before participants could move on to the next token. Participants were given practice trials with the experimenter present before starting on the assigned (A or AV) training paradigm. In each training paradigm, participants either heard (A) or heard and saw (AV) the stimuli and responded by pressing keys "z" (word on the left)

or "m" (word on the right). The display of the AV training was similar to the A training but with a video of the native English speaker producing the sound (see figure 3). The entire training was two sessions long, held over the period of two days. Each training session was approximately 35 to 40 minutes, for a combined total of 80 minutes of training. The 120 tokens were randomized across 4 blocks, to make up a total of 480 tokens per day and 960 tokens across two days.

The procedure for day 2 of the experiment were similar to day 1. Participants were asked to continue the 2AFC training task followed by a post-test, which was the same as the pre-test described above, where participant's production and perception were measured.

Figure 3. Display of audio-only training (A) versus audio-visual training (AV)
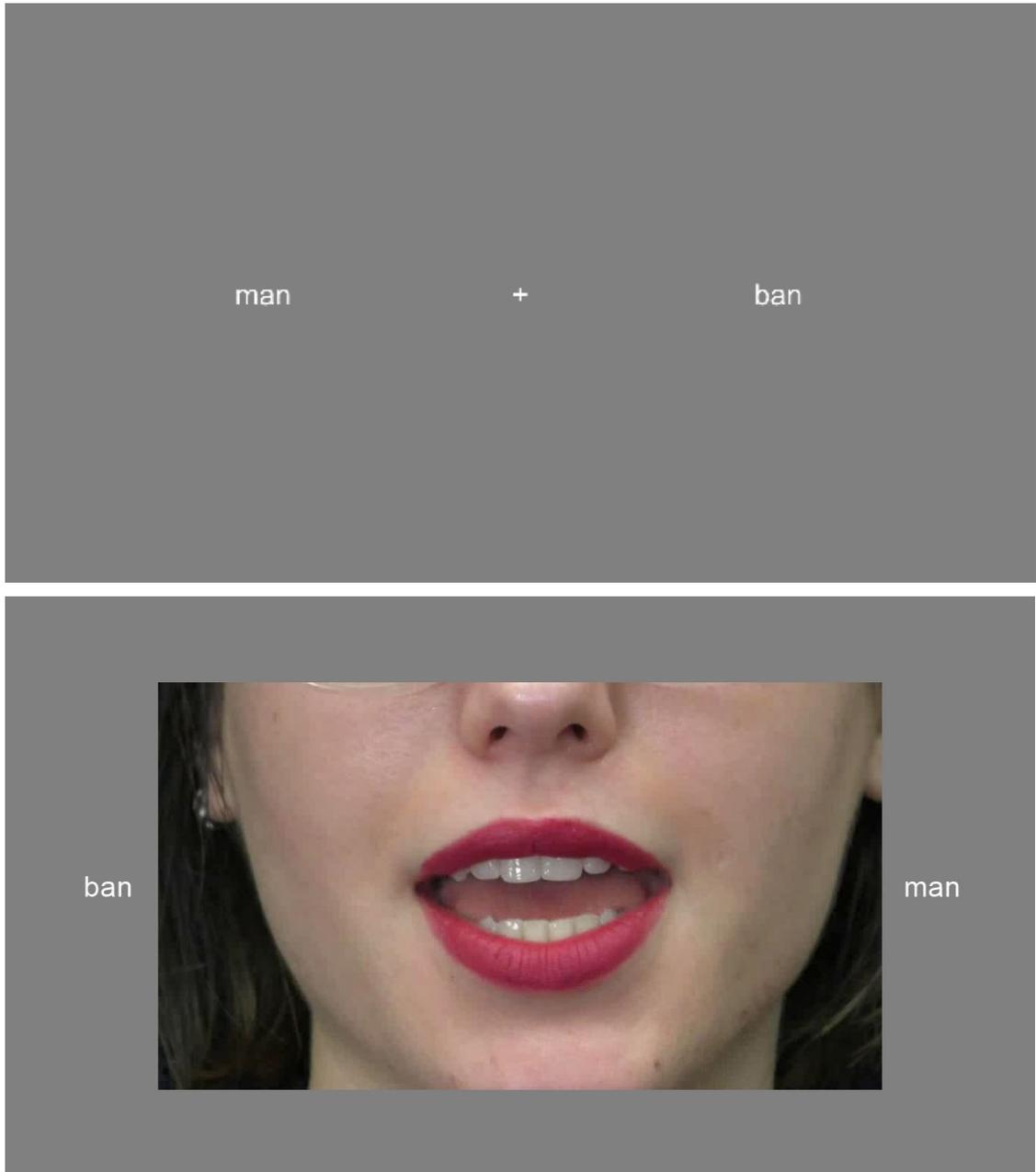
Analysis:

      The data gathered from the 2AFC task were analyzed quantitatively using mixed analysis of variance (ANOVA) on SPSS to assess any training effects on improvement of labial/labiodental consonants (IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp).

Results

Overall performance

Before conducting any analysis, the data was evaluated for normality through SPSS through its explore feature and was checked for a skewness and kurtosis value between -2 and +2 and an even distribution on a histogram (George & Mallery, 2010). One outlier was identified and was removed from the data analysis. Although the pre/post-test were assigned, later evaluation of the data discovered that there was no assessment of consonant final position. This is important as consonant position is one of the independent variables of interest in this study. Therefore, the pre/post test data were not used, and data from day 1 and day 2 were used instead to evaluate participants' overall performance across time. A repeated measures of variance (ANOVA) was used to calculate difference in condition (A, AV) on time (day 1, day 2). This interaction is illustrated in Figure 4.

Figure 4. Interaction of time x training on identification performance
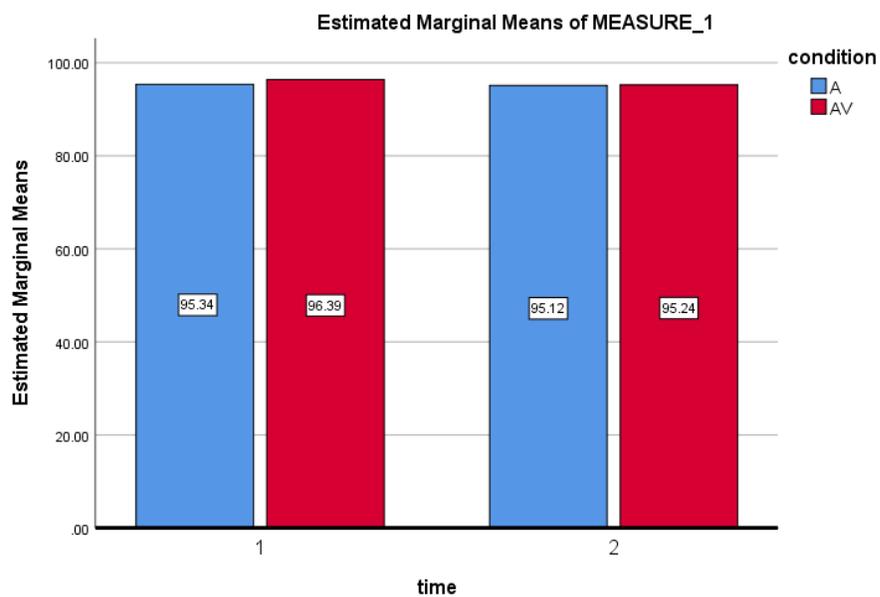
In the A condition, there was a difference of day 1 and day 2 score of -0.22%
while in the AV condition, there was a difference of day 1 and day 2 score of -1.15%.
Overall, neither training had a significant effect on participants' performance on the
identification task and scores between the 2 training days did not significantly change.

The overall performance between the 3 L1 group in each condition were
calculated to observe their overall accuracy on the identification task. Mean percent
correct scores across A and AV for each L1 group were illustrated in Figure 5.

Figure 5. Mean % correct scores of the 3 languages (Arabic, Japanese, and Mandarin)
across A and AV training
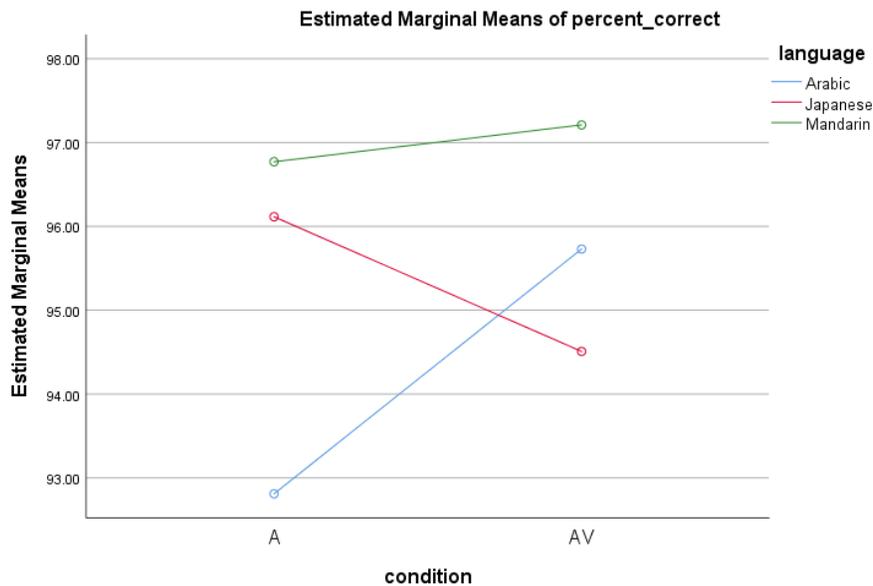
For Arabic L1 speakers, the mean percentages were 92.81% (no s.d. because
there was only 1 subject) in the A condition, 95.73% (s.d. 2.06) in the AV condition and
a combined average of 94.27%. For the Japanese L1 speakers, the mean percentages
were 96.12% (s.d. 1.36) in the A condition, 94.51% (s.d. 3.29) in the AV condition,

with a combined average score of 95.32%. Lastly, for the Mandarin L1 speakers, mean

percentages were 96.77% (s.d. .2.86) in the A condition and 97.01 (s.d. .711) in the AV

condition, with a combined average of 96.89%. The averaged mean total in A across all

L1 was 96.03% (s.d. 2.31) and 95.95% (s.d. 2.34) in AV, with a combined average of

96.89%. The mean percentages of all these scores showed that Mandarin L1 speakers

had the highest performance accuracy, followed by Japanese L1 speakers, and lastly

Arabic L1 speakers. However, the differences between them were very small, with only

about 1% difference between each group. These results suggest that performance scores

between each L1 were not significantly different from each other.

Training effects on specific paired contrasts

Though all participants scored near ceiling results, observations were made of

lower accuracy scores on specific paired labial/labiodental contrasts in each condition

(see Figure 6). For Arabic speakers, the mean score of the /b/-/p/ initial paired contrast

was 57.29% in A and 82.29% in AV. The /f/-/v/ final paired contrast was 68.75% in A

and 93.75% in AV. For Japanese speakers, the mean score of the /b/-/v/ initial paired

contrast was 78.95% in A and 86.22% in AV. Scores for Mandarin speakers were

consistent in both conditions. Lower bound scores, however, were found on the /f/-/v/

final paired contrast with 78.65% in A and 77.5% in AV. To see whether the differences

above were a result of the training assigned, a multivariate analysis of variance

(MANOVA) was used for Arabic with scores from /b/-/p/ initial, /b/-/v/ initial, final, /f/-

/v/ initial, final as dependent variables and training as the independent variable. Results

from the MANOVA showed that there was not a statistically significant difference in

scores between the 2 trainings. For Japanese speakers, a one-way ANOVA using scores

from /b/-/v/ initial as dependent variables and training as an independent variable,

results showed training had no significant effect on score changes. Overall, training had

no effect on improving known challenging labial/labiodental contrasts for specific L1

groups. Performance on the identification task, however, do seem to be consonant

specific as some pose to be more challenging than others.

Figure 6. Mean % correct scores of paired contrasts (/b/-/p/, /b/-/v/, /f/-/v/) in initial and final position across the 3 languages (Arabic, Japanese, Mandarin) in A and AV training

Audio-Only

	bp initial	bp final	bv initial	bv final	fv initial	fv final
Arabic	57.29	70	80.21	75	84.38	68.75
Japanese	97.89	90.16	78.95	72.07	100	97.85
Mandarin	97.92	89	97.92	95.65	97.4	78.65

Audio-visual

	bp initial	bp final	bv initial	bv final	fv initial	fv final
Arabic	82.29	67	96.88	96.74	94.79	93.75
Japanese	93.91	91	86.22	71.2	97.45	96.35
Mandarin	97.92	90.4	100	98.25	95.83	77.5
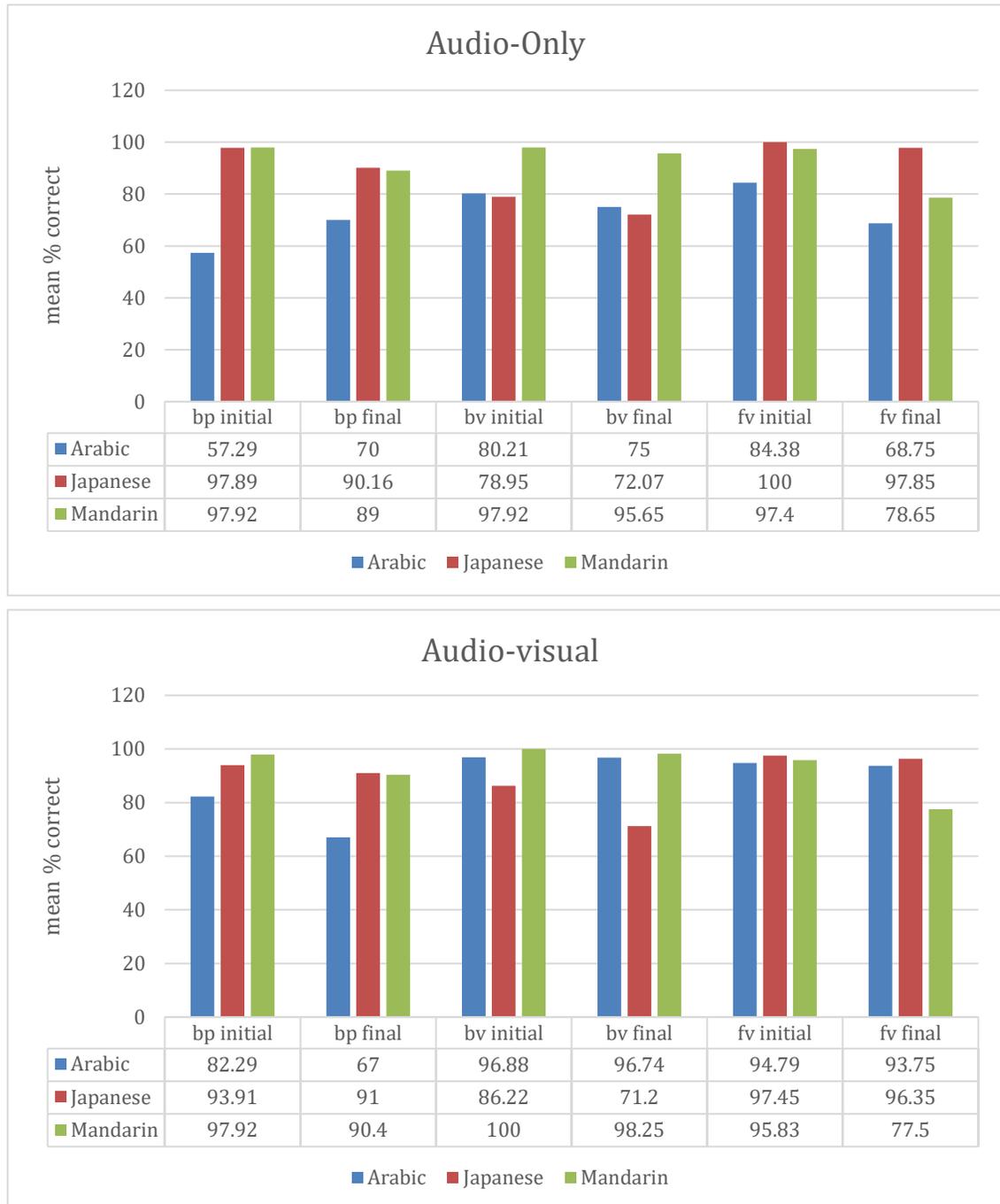
Condition effects on consonant position across L1 groups

Changes in scores were observed between position (initial, final) of paired contrasts across both conditions. (see Figure 6). For Arabic speakers, the score difference between initial and final position on the /f/-/v/ contrast was 12.71% in A and 15.29% in AV while the score difference between initial and final position on the /b/-/p/ contrast was 12.71% in A and 15.29% in AV. For Japanese speakers, the difference between initial and final position on the /b/-/v/ contrast was 6.88% in A and 15.02% in AV and the difference between the /b/-/p/ paired contrast was 7.73% in A and 2.91%. Lastly, for Mandarin speakers, the difference between initial and final position of the /b/-/p/ paired contrast was 8.92% in A and 7.52% in AV while the initial and final /f/-/v/ paired contrast was 18.75% in A and 18.33% in AV. The data suggest consonant position played an important role on participants' performance on the identification task.

In order to see if these within-subjects differences were statistically significant and whether training played a factor, a repeated measures ANOVA was used by assigning paired contrasts (ones specifically known to be difficult for non-native speakers in each group) and position (initial, final).

Arabic

In a repeated measures ANOVA using paired contrasts (/b/-/p/, /f/-/v/) and position (initial, final) as within group factors, and training (A, AV) as between subjects factors, results showed that there was not a significant effect of paired contrasts, position, or condition on participants' performance on the identification task.

Japanese

In the Japanese sample, the position between /b/-/v/ and /b/-/p/ contrasts were evaluated. A repeated measures ANOVA was performed using paired contrasts (/b/-/v/, /b/-/p/) and position (initial, final) as within-subjects factors and condition (A, AV) as between-subjects factors. Results from the test showed that there was a main effect for position [F $(1,6)$ = 6.824 p = .040] and a main effect for contrast [F$(1,6)$ = 8.981 p= .024] , and a statistically significant position*contrast*condition interaction [F$(1,6)$ = 5.23, p = .014] (see figure 7). Overall, results from the test show that consonants were easier to identify in the initial position than in the final position. Additionally, it was easier to identify between /b/-/p/ than /b/-/v/. Lastly, identification of these paired contrasts in specific positions changed as a result of training. Japanese speakers were able to perform better on identifying /b/-/p/ in the final position in the AV training than in the A training. Additionally, they were also able to identify /b/-/v/ in the initial position more accurately in the AV training. Results here suggest AV training of having a positive effect on improving Japanese speakers' perception of some challenging labial/labiodental sounds.

Figure 7. Mean % correct of position x contrast x condition for Japanese speakers

Estimated Marginal Means of MEASURE_1
at condition = A

Estimated Marginal Means of MEASURE_1
at condition = AV
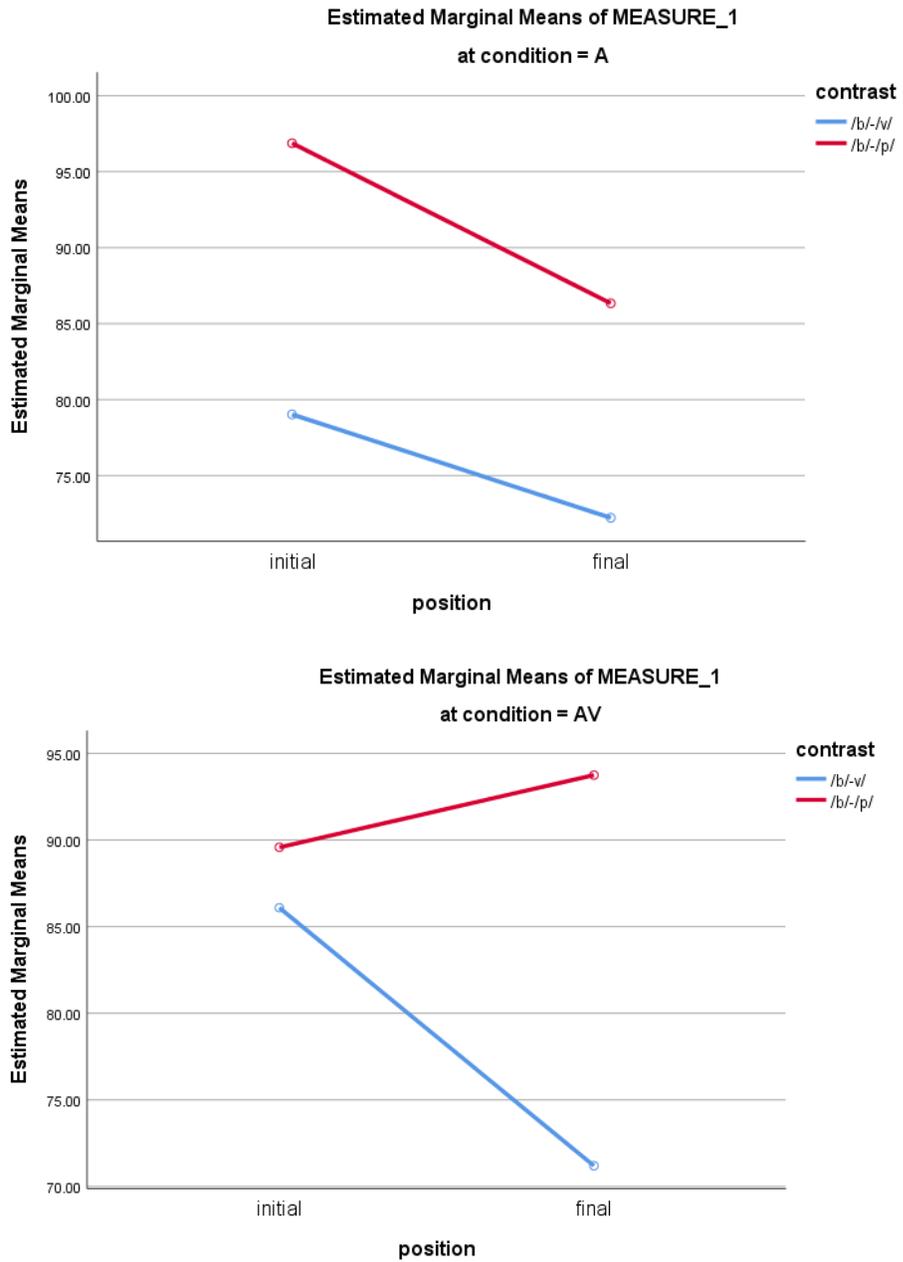
Mandarin

In the Mandarin sample, the within-subjects effect of the /b/-/p/ and /f/-/v/ paired

contrast in the initial and final position were evaluated. A repeated measures ANOVA

31

was used with paired contrasts (/b/-/p/, /f/-/v/) and position (initial, final) as within-subjects factors and condition (A, AV) as between subjects factors. Results from the test showed that there was a main effect for position $[F(1,7) = 27.291, p = .001]$ and a main effect for contrast $[F(1,7) = 9.086, p = .020]$ and a contrast*position interaction $[F(1,7) = 6.049, p = .044]$ (see figure 8). These results indicate that position played a significant role in participants' identification of these paired contrasts. Identification were also made easier when consonants were in the initial than final position. Additionally, correct identification was influenced by the type of paired contrasts asked to be identified. Participants were able to identify /f/-/v/ initial words (e.g., fan and van) with high accuracy while /f/-/v/ final words (e.g., cuf and cuv) were difficult.

Figure 8. Mean % correct of position x contrast for Mandarin speakers
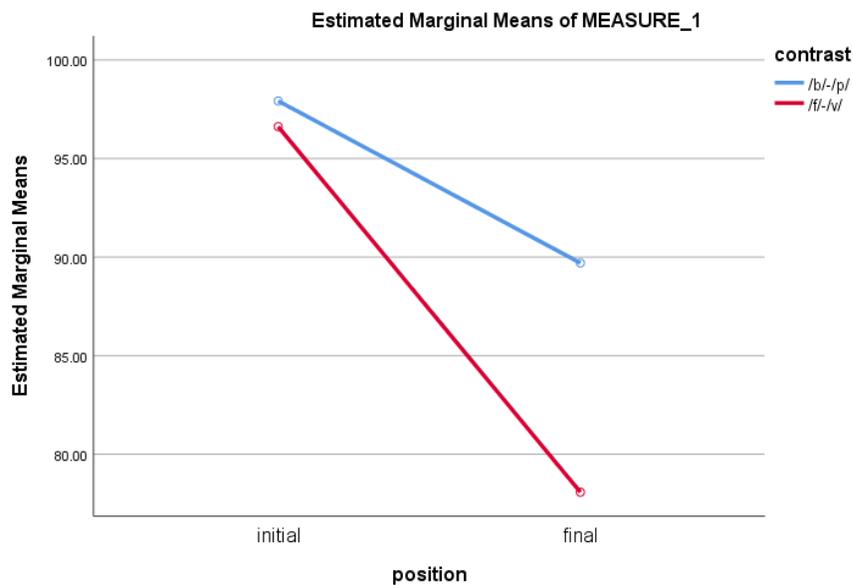
Discussion

Contrary to the initial hypothesis, there was no significant difference between A or AV training on improving participants' identification of labial/labiodental sounds. Scores between the two days also did not significantly change. Overall, the averaged mean percent correct scores across both training paradigms were very high for each L1 group (Arabic = 94.27%, Japanese = 95.32%, Mandarin = 96.77%), which left little room for improvement. I suspect that one of the main reasons behind these ceiling scores were participants' prior English experience. According to the Language Experience Questionnaire administered before the training, many participants reported previous English learning/training in their native country for an average of at least 8 years before enrolling at the AEI. Additionally, most participants were level 5 or 6 students at the AEI, which are two of the highest-level classes at the AEI. Therefore, many of the participants were, in fact, not beginning English learners, but semi-experienced English learners. However, performance scores were not consistent across all labial/labiodental paired contrasts as lower accuracy scores were observed across each group. For Arabic speakers, lower correct scores were observed following identification of labial/labiodentals /p/, /b/, /v/, /f/ sounds, which was consistent with other studies focusing on Saudi ESL learners (Flege, 1980; Binturki 2001). In an analysis of pronunciation errors of Saudi ESL learners, Binturki (2001) also found that Arabic learners tend to have lower accuracy means when trying to discriminate between /b/-/p/ and /v/ contrasts. This was probably true due to a result of voicing confusions between /b/ and /p/ as /p/ is absent in the Arabic phonemic category. In fact, many of the lower accuracy paired contrast observed were all relative to missing phonemes in

each L1 group's native phonemic category, with Japanese speakers having difficulty identifying between /b/, /p/, /v/ sounds and Mandarin speakers with /f/ and /v/ sounds. These trends show how categorical perception may have had a strong influence on participants' performance on the identification task; thereby supporting the PAM (Best, 1995) model of second language perception. Accuracy scores, overall, may be a result of having category specific confusions between L1 and L2 and identification scores may have been low as a function of speakers integrating L2 sounds into already existing L1 categories (Best, 1995). These observations posit that participant scores were, in a great degree, influenced by knowledge or experience of their L1, which supports our second hypothesis that participant performance on the identification task will be heavily influenced by their language background and known difficulties of labial/labiodental sounds.

Results from the assessment on consonant position showed great difference between how labial/labiodental consonants were identified when placed in the initial versus the final position of words. Both Japanese and Mandarin speakers were better at identifying labial/labiodental consonants when placed in the initial position of a word than the final position of a word. This was consistent with findings from Zhang and Hayashi (2015) in a study on perception of English syllable-final consonants by Mandarin and Japanese speakers, where speakers were better at identifying syllable initial-consonant (82.6%) than syllable-final consonant (67.8%). This may be because in Mandarin and Japanese, consonants rarely appear in the final position of words. In Mandarin, words typically end in a vowel sound. The only two consonants that ever do appear in the final position are nasals /n/ and / ŋ/ (Duanmu, 2000). Similarly, Japanese

34

also favors final vowel sounds, and the only consonant that ever does appear in the final is the placeless nasal /N/ (Nasukawa, 2010). The noticeable differences between the high scores of when consonants were in the initial position versus when consonants were in the final position show the importance position has in determining how participants performed on the identification task. Moving forward, it would be helpful to narrow the scope and target only final position consonants, as those tend to be more challenging for Japanese and Mandarin speakers.

Although there was no A or AV training effect on participants' identification of labial/labiodental consonants, there was however, specific effects on Japanese participants' ability to identify between paired /b/-/p/ and /b/-/v/ consonants in the initial and final position.

Results from the data showed that the AV training was more effective on improving Japanese participants' perception of the /b/-/p/ final contrasts than the A training. Additionally, the AV training was also effective on improving Japanese participants' perception of the /b/-/v/ initial contrasts than the A training. Results from here suggest that the effect AV training has on improving identification of labial/labiodental contrasts may be position and contrast specific. The addition of gestural cues in the AV training was only useful when trying to distinguish the /b/-/v/ initial contrasts and not the /b/-/v/ final contrasts. Overall, there seems to be different weight in terms of how audio and visual cues are used to benefit Japanese participants' perception of certain labial/labiodental contrasts. Therefore, it's not really the case that one training is superior than the other on improving second language perception, but more that the effects may be language-specific.

Implications

Although various audio-visual training studies have been implemented in second language learning and acquisition research, few have explored its effects on improving labial/labiodental speech contrasts among various language groups. This research was created to bridge that gap. While general results did not show a change in performance between audio-only training and audio-visual training on Arabic, Japanese, and Mandarin participants' perception of labial/labiodental speech contrasts there was, however, an effect for specific paired contrast in specific positions (Japanese participants' perception of the /b/-/p/ final contrast and /b/-/v/ initial contrast). This research identified perceptual difficulties that were language specific and discovered paired contrast and position as important factors to participants' perceived difficulties of labial/labiodental sounds. Information about this research would be useful for teachers and all communication partners. Better understanding of challenging paired contrasts can help teachers with effective lesson planning and build curriculums that cater specifically to students' needs. Likewise, increased awareness of challenging speech contrasts can help us all be more patient and empathetic communication partners and create social environments that prompt active communication and engagement.

Limitations & Future Directions

One of the biggest limitations was length of the training task. Unlike many other audiovisual training studies where training was longitudinal (Bradlow et al., 1997; Logan et al., 1991, Hazan et al, 2005,2006), our training was only 2 days. The change in effects that could be observed from such a short time frame is usually marginally small (as seen from the data), therefore, it would be helpful to lengthen the time for future designs in order have a better measure of training effects. The pre-test and post-test were not used for the data analysis due to design flaw of not incorporating final consonant position contrasts in the test. Analysis of training data may not have been the most accurate measure of performance scores due to immediate feedback, which might have boosted participant's performance scores. Performance scores may have been different if there was no feedback. In addition, the study could also benefit from having a bigger sample size. Data from the Arabic sample were not reliable due to how small the sample size was (1A to 2AV) and therefore, much of the data collected from that sample were not used in the analysis because the sample size was uneven and too small. Another important limitation is the variability of labial/labiodental consonants used across each L1. The near ceiling performance scores were likely a result of high variability of labial/labiodental consonants used in the design. Results from the data show that the labial/labiodental challenges were group specific. Therefore, over variability of other already known easy labial/labiodental contrasts may have resulted in participants' overall higher scores on the identification task. Therefore, future directions would be to remove the known easy contrasts and only measure for the known challenging ones in order to get an accurate measure of the effect A and AV training

may have on perception of select labial/labiodental consonants. The position of where labial/labiodental consonants were placed within a word also plays an important role on how well it is identified, with consonant-final being more difficult. Future research could narrow the scope and focus on the effect A and AV training have on improving consonant-final sounds. Lastly, while data of speech productions were collected, they were not analyzed. It would be good to analyze the data from speech productions to see if A or AV training had any effects on labial/labiodental speech productions. The steps for production analysis will be as follows:

1. Analysis of non-native production data will be rated and scored by native English speakers at the University of Oregon.

2. Participants will listen to a series of productions made by the second language learners and be asked to rate the quality of the productions on a scale of 1-5 (e.g., 1: very difficult to understand, 5: very easy to understand).

Conclusion

Although neither the audio-only or audio-visual training had an overall effect on non-native speakers' perception of labial/labiodental consonants, there was a small effect on Japanese participants' perception of the /b/-/p/ final contrast and the /b/-/v/ initial contrast. Therefore, effectiveness of the training may be contrast and position specific. Results from this finding suggest that, perhaps with more training, more improvements will occur across each language group. Application of audio-visual training in second language acquisition research is valuable as it combines the use of both audio and visual cues to create multi-modal speech representations that may better guide non-native speech perception. While previous research has mainly focused on other non-native consonant identifications, results from this research report that there are evident perceptual difficulties with labial/labiodental consonants across different language group that are left to be explored. Therefore, further investigation on this topic would be good to help narrow down language specific perceptual difficulties so that more targeted audio-visual training regimens may be created to better improve non-native speaker's perception of non-native sounds.

APPENDICES

Appendix A: Informed consent form (English)

You are invited to participate in an honors thesis research conducted by Chia-Ni (Jennie) Shen and Melissa M. Baese-Berk, Ph.D. You are asked to participate in this research because you are a non-native English speaker at least 18 years of age enrolled in the American English Institute. This research is interested in investigating the influence of audiovisual learning on speech and language development. We ask that you read this form and ask any questions that you may have before agreeing to be in the study.

The Study:

      The experiments for this research will examine how audiovisual input influence an individual's speech perception and production. We will show you a series of words on a computer screen and ask you to listen and/or observe them. We may ask you to repeat the words aloud and record your speech for later acoustic analysis. You may also be asked make judgments on a word by pressing buttons or clicking on the screen. All of your responses will be recorded. We will also ask you to complete questionnaires related to your language experience and proficiency.

      Your participation today will last no longer than 1 hour today and 1 hour tomorrow. You will receive 2-3% of extra credit for participating. The exact percentage is up to your instructor. Note that this is one option for extra credit offered by your instructor. Please talk to your instructor if you wish to do an alternative extra credit assignment.

Your Rights:

      Your participation is voluntary. Know that your decision on whether or not to participate will not affect your relationship with the University of Oregon, the UO Linguistics Department, and/or the American English Institute (AEI). If you decide to participate, you are free to withdraw participating in the experiment at any time without penalty. If you choose not to participate, an alternative yet comparable assignment has been created by your instructor at the AEI. Please communicate with your instructor if you wish to complete the alternative assignment instead.

      Your identity will be protected at all times and any data published from this research will not be linked back to your name. You will be assigned a code and will only be identified by this code throughout both the experiment and data analysis. All personal data collected will be saved in a password-protected computer, separate from other data gathered for the experiment. Your name will only be kept until the end of the term to notify instructors of your participation to assign the appropriate extra credit. All files, linking back to your name, will be deleted at the end of the term.

Risks/benefits: There are no foreseeable risks nor benefits from participating in this study.

Results from this study could, however, yield further knowledge in second language acquisition and non-native speech and language learning. Likewise, this research could also provide further insight on speech pathology, and development of better language learning tools and programs throughout different language learning institutions.

Contact Information:

Dr./Ms./Mr. _____has explained this study to you and answered your questions. If you have other questions or research-related problems, you may contact Chia-Ni (Jennie) Shen, jshen@uoergon.edu (971) 506-2149 or Dr. Melissa Baese-Berk, mbaesebe@uoregon.edu (541) 346-3899. If you have questions regarding your rights as a research subject, please contact Research Compliance Services, 5237 University of Oregon, Eugene, OR 97403, (541) 346-2510, or email ResearchCompliance@uoregon.edu.

Recording Permission

Your signature indicates that you have read and understood the information provided above;

- that you willingly agree to have your speech recorded.
- that these recordings will be coded so that your name will not be associated with these recordings.

Participant's Signature _____   Date
_____

Data Sharing Permission

Your signature indicates that you have read and understand the information provided above;

- that you willingly agree to have your recordings made available to other scientists in a public repository.
- that these recordings will be coded so that your name will not be associated with these recordings.
- that excerpts of the audio recordings may be used in conference or classroom presentations for the investigator's honors thesis.

Participant's Signature _____   Date
_____

Your Signature:

Your signature indicates that you have read and understand the information provided above;

- that you willingly agree to participate;
- that you may withdraw your consent at any time and discontinue participation without penalty;
- that you will receive a copy of this form if so desired; and
- that you are not waiving any legal claims, rights, or remedies.

Participant's Signature _____ Date
_____

Appendix A: Informed consent form (Simplified Chinese)

<div align="center">

同意书

</div>

您被邀请参与Chia-Ni (Jennie) Shen 和 Melissa M. Baese-Berk, Ph.D 的毕业论文研究问卷。此问卷调查对象为非英文母语、满18岁以上，并且目前就读于American English Institute之学生。研究目的为调查『视听学习对口语表达能力和语言提升的影响』。请您在问卷开始前阅读并完成此表格，有任何问题请在此时提出。

您好：

这向研究将探讨视听训练如何影响个人的口说语感和语言表现，我们将在计算机屏幕上显示一系列词语，并要求您聆听或观察它们，同时可能会要求您重复朗读这些词语并记录下您的语音，以便日后进行声学学分析。您也可能会被要求透过按下按钮或点击屏幕来对单词进行判别，所有的选择都将会被记录下来，最后请您填写关于您个人语言经验和熟练程度的问卷调查。

此问卷作答时间：一天将不会超过1小时，但需包含当日1小时以及次日1小时，当您完成这项问卷后，您将获得＄10/小时，总共20美元。如果您在问卷中途停止参与，以每30分钟计次，您将获得5美元的奖励，超过30分钟将取整计次。例如：如果您完成1-30分钟，您将收到＄5美元，如果您完成31-60分钟，您将收到＄10美元，依此类推。

您的权力：

您的参与是自愿的，您决定是否参与此次问卷并不会影响您与 University of Oregon、UO Linguistics Department 或 American English Institute (AEI)的关系。如果您决定参加，您可以在任何时候退出测验而不受到处罚。

您的身份将始终受到保密，本研究未来发布的任何数据都不会提及您的姓名，受测者将个别分配一个代码，并且只有在研究和数据的分析过程中才会通过此代码进行识别；收集到的所有个人数据将被保存在有监管系统的计算机中，并与测验中收集到的其他数据分开，所有与您有关的数据分析或文件将在学期结束后予以删除。

风险/益处：参与本研究没有可预见的个人风险或益处。

然而，这项研究的结果可以提供研究人员对于第二外语、非母语的语言学习方面有更进一步的了解，并且在未来于不同的语言学习机构中开发出更好的语言学习工具和程序。

联络信息：

Dr./Ms./Mr. _____已经向我解释了这项研究并回答了我的问题。

如果您还有其他问题或其他与研究有关的问题，请联系Chia-Ni (Jennie) Shen, jshen@uoergon.edu (971) 506-2149 或者 Dr. Melissa Baese-Berk, mbaesebe@uoregon.edu (541) 346-3899。如果您对于作为研究受测者的权力有任何疑问，请联系Research Compliance Services, 5237 University of Oregon, Eugene, OR 97403, (541) 346-2510, 或者 email ResearchCompliance@uoregon.edu.

录音权限：

您的签名表示您已阅读并理解上面提供的信息

- 您同意您的语音被记录下来。

- 这些录音将被编码，以便您的名字不会与这些录音档相关联。

受测者签名：_____ 日期_____

数据分享权限：

您的签名表示您已阅读并理解上面提供的信息

- 您同意把您的语音文件提供给公共数据储存库中的其他研究人员。

- 大学教授可能有权查看这些数据。
- 这些录音将被编码，以便您的名字不会与这些录音档相关联。
- 录音档的摘录可用于研究者的毕业论文、会议或课堂范例中。

受测者签名：＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿ 日期＿＿＿＿＿＿＿

您的签名：
您的签名表示您已阅读并理解上面提供的信息
- 您同意参与测验
- 您可以在任何时候撤回您的同意，退出测验而不受到任何处罚。
- 您可以索取此份表格的副本。
- 您不放弃任何法律索赔、权力或补救措施。

受测者签名：＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿ 日期＿＿＿＿＿＿＿

Appendix A : Informed consent form (Traditional Chinese)

同意書

您被邀請參與Chia-Ni (Jennie) Shen 和 Melissa M. Baese-Berk, Ph.D 的畢業論文研究問卷。此問卷調查對象為非英文母語、滿18歲以上，並且目前就讀於American English Institute之學生。研究目的為調查『視聽學習對口語表達能力和語言提升的影響』。請您在問卷開始前閱讀並完成此表格，有任何問題請在此時提出。

您好：

這向研究將探討視聽訓練如何影響個人的口說語感和語言表現，我們將在電腦屏幕上顯示一系列詞語，並要求您聆聽或觀察它們，同時可能會要求您重複朗讀這些詞語並記錄下您的語音，以便日後進行聲學學分析。您也可能會被要求透過按下按鈕或點擊屏幕來對單詞進行判別，所有的選擇都將會被記錄下來，最後請您填寫關於您個人語言經驗和熟練程度的問卷調查。

此問卷作答時間：一天將不會超過1小時，但需包含當日1小時以及次日1小時，當您完成這項問卷後，您將獲得＄10/小時，總共20美元。如果您在問卷中途停止參與，以每30分鐘計次，您將獲得5美元的獎勵，超過30分鐘將取整計次。例如：如果您完成1-30分鐘，您將收到＄5美元，如果您完成31-60分鐘，您將收到＄10美元，依此類推。

您的權力：

您的參與是自願的，您決定是否參與此次問卷並不會影響您與

University of Oregon、UO Linguistics Department 或 American English

Institute (AEI)的關係。如果您決定參加，您可以在任何時候退出測驗而不受到處罰。

您的身份將始終受到保密，本研究未來發佈的任何數據都不會提及您的姓名，受測者將個別分配一個代碼，並且只有在研究和數據的分析過程中才會通過此代碼進行識別；收集到的所有個人數據將被保存在有監管系統的電腦中，並與測驗中收集到的其他數據分開，所有與您有關的數據分析或文件將在學期結束後予以刪除。

風險/益處：參與本研究沒有可預見的個人風險或益處。

然而，這項研究的結果可以提供研究人員對於第二外語、非母語的語言學習方面有更進一步的了解，並且在未來於不同的語言學習機構中開發出更好的語言學習工具和程序。

聯絡資訊：

Dr./Ms./Mr. _____已經向我解釋了這項研究並回答了我的問題。

如果您還有其他問題或其他與研究有關的問題，請聯繫Chia-Ni (Jennie) Shen, jshen@uoergon.edu (971) 506-2149 或者 Dr. Melissa Baese-Berk, mbaesebe@uoregon.edu (541) 346-3899。如果您對於作為研究受測者的權力有任何疑問，請聯繫Research Compliance Services, 5237 University of Oregon, Eugene, OR 97403, (541) 346-2510, 或者 email ResearchCompliance@uoregon.edu.

錄音權限：

您的簽名表示您已閱讀並理解上面提供的信息

- 您同意您的語音被記錄下來。

- 這些錄音將被編碼，以便您的名字不會與這些錄音檔相關聯。

受測者簽名：＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿ 日期 ＿ ＿ ＿ ＿ ＿ ＿ ＿

數據分享權限：

您的簽名表示您已閱讀並理解上面提供的信息

- 您同意把您的語音檔提供給公共數據儲存庫中的其他研究人員。

- 大學教授可能有權查看這些數據。

- 這些錄音將被編碼，以便您的名字不會與這些錄音檔相關聯。

- 錄音檔的摘錄可用於研究者的畢業論文、會議或課堂範例中。

受測者簽名：＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿ 日期 ＿ ＿ ＿ ＿ ＿ ＿ ＿

您的簽名：

您的簽名表示您已閱讀並理解上面提供的信息

-　您同意參與測驗

-　您可以在任何時候撤回您的同意，退出測驗而不受到任何處罰。

-　您可以索取此份表格的副本。

-　您不放棄任何法律索賠、權力或補救措施。

受測者簽名：＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿ 日期 ＿ ＿ ＿ ＿ ＿ ＿ ＿

Appendix A: Informed consent form (Japanese)

母語が英語でない話者に対する視聴覚訓練の効果
研究への参加についての同意書 AEI

これは、Chia-Ni（Jennie）ShenとMelissa M. Baese-Berk博士（Ph.D.）が行う卒業論文研究の参加についての補足資料です。言語と言語の発達に及ぼす視聴覚学習の影響についての研究のため、オレゴン大学に入学した18歳以上の英語が母国語ではない話者が対象になります。同意の署名をしていただく前に、この研究についてご理解いただき、この研究についてわからないことがあった場合には、遠慮なくお問い合わせください。

研究目的：
　この研究は、視聴覚からの情報が、各個人の知覚、発話に影響を与えているかを実験します。様々な録音された音を聞き、何が聞こえたかを文字で入力、もしくはボタンを押すか画面で選択します。また、何を聞いたか 1～5 の範囲でランク付けすることも求められます。全ての被験者の応答は記録され、どの言語を使用したことがあるかをアンケートで答えていただきます。

　これらの実験には20～30分を要し、0.5単位までこの実験の終了時には認定されます。もし、途中で実験の参加をやめられた場合、15分毎に0.25単位が認定され、次の参加していただいた15分に加算されます。例えば、1～15分実験に参加していただいた場合、0.25単位が認定され、16～30分参加した場合、0.5単位が認定されます。もし、参加の予約をしていながら、参加を撤回される場合は、0.25単位が認定されます。

権利：
　被験者の参加は任意です。同意されない、もしくは同意を撤回された場合でも、オレゴン大学、UO言語学科、および/またはアメリカ英語研究所（AEI）と

被験者の関係に不利が生じることはありません。 同意した後でも、罰則なしにいつでも取り消すことができます。心理学と言語学科では、参加を希望しない研究テーマであれば、代替課題を設けています。 代替の研究テーマに参加を希望する場合は、担当者に申し付けください。

　個人情報は常にプライバシーが守られており、この研究から発表されたデータに被験者の個人情報が載ることはありません。被験者から得られたデータの管理はコード番号で行われ、実験や分析はこのコードで識別されます。そして、これらのデータはセキュリティーのかかったコンピュータ内で管理され、他の研究結果と一緒になることはありません。被験者に正しく単位が授与されるまで、被験者の個人情報は保護され、被験者の個人情報が含まれた全てのデータは、その後削除されます。

参加により生じる利益、並びに予測されるリスク：
この研究に参加することによる利益として、研究の結果はが、第二言語獲得と非母国語の会話と言語学習の知識をさらに深めるが期待されます。この研究はまた、言語病理学のさらなる見識を提供し、異なる言語学習機関を通じて、より良い言語学習ツールおよびプログラムの開発、発展のために行われるものであることをご理解ください。

問い合わせ：
担当者 ＿＿＿ が説明を行いました。
この研究について知りたいことや、ご心配なことがありましたら、以下の担当者にご相談下さい。
Chia-Ni (Jennie) Shen, jshen@uoergon.edu (971) 506-2149
Dr. Melissa Baese-Berk, mbaesebe@uoregon.edu (541) 346-3899.
Research Compliance Services, 5237 University of Oregon, Eugene, OR 97403, (541) 346-2510, email ResearchCompliance@uoregon.edu.

データ結果記録の許可：

次の項目について詳しい説明を受け、十分に理解し、参加することに同意します。
- 実験で行われる発話等の録音
- 録音されたものは個人名が関連付けされないようにコード化されること

署名：
日付：

実験結果共有の許可:

次の項目について詳しい説明を受け、十分に理解し、参加することに同意します。
- 記録されたデータを公開リポジトリの他の科学者が利用できるようにすること
- 大学関係者がそのデータにアクセスする可能性があること
- 記録されたデータは、個人情報がこれらに関連付けられないようにコード化されること
- 部分によっては記録されたデータが、研究者の卒業論文のためのプレゼンテーション等で使用されること

署名：
日付：

参加への同意：
- 被験者になること
-　研究への参加に同意しないことまたは同意を撤回することによって不利益な扱いを受けないこと
- この同意書のコピーを必要であれば受け取ること
- 法的な請求、権利、または救済を放棄していないこと

署名：

日付：

Appendix A: Informed consent form (Arabic)

أنت مدعو للمشاركة في أبحاث أطروحة يكرم أطروحة التي أجراها شيا ني (جيني) شين وميليسا M. بيس بيرك، دكتوراه. يطلب منك المشاركة في هذا البحث لأنك متحدث باللغة الإنجليزية الأصلية لا تقل عن 18 عاما من العمر في جامعة أوريغون. ويهتم هذا البحث بالتحقيق في تأثير التعلم السمعي البصري على تنمية الكلام واللغة. نطلب منك قراءة هذا النموذج وطرح أي أسئلة قد تكون لديك قبل الموافقة على أن تكون في الدراسة.

الدراسة:

سوف تدرس تجارب هذا البحث كيف تؤثر المدخلات السمعية البصرية على تصور الفرد وإنتاجه. سنلعب سلسلة من التسجيلات الصوتية ونطلب منك الاستماع واتخاذ الأحكام حول ما تسمعه عن طريق الكتابة، الضغط على أزرار أو النقر على الشاشة. قد يطلب منك أيضا ترتيب ما تسمعه على مقياس من 1 إلى 5. سيتم تسجيل جميع ردودك. سوف نطلب منك أيضا إكمال استبيان تجربة اللغة.

لن تستمر مشاركتك اليوم أكثر من 20 إلى 30 دقيقة. سوف تحصل على ما يصل إلى ½ علامه عند الانتهاء من هذه الدراسة. إذا توقفت عن المشاركة في منتصف الدراسة، ستحصل على رصيد لكل 15 دقيقة من المشاركة، وتقريبها إلى ال 15 دقيقة القادمة. على سبيل المثال، إذا أكملت 1-15 دقيقة سوف تتلقى ¼ علامه، إذا كنت اكملت 16-30 دقيقة سوف تتلقى ½ علامه، وهلم جرا. إذا حافظت على موعدك المقرر ولكن قررت عدم المشاركة، سوف تتلقى ¼علامه.

حقوقك:

مشاركتكم طوعية. اعرف أن قرارك بشأن المشاركة أو عدم مشاركته لن يؤثر على علاقتك مع جامعة أوريغون أو قسم علم اللغويات أو و / أو معهد اللغة الإنجليزية الأمريكي (إي). إذا قررت المشاركة، أنت حر في سحب المشاركة في التجربة في أي وقت دون عقوبة. وقد أنشأت قسم علم النفس واللغويات مهام بديلة للطلاب الذين لا يرغبون في المشاركة كمواضيع بحثية. يرجى التواصل مع معلمك إذا كنت ترغب في إكمال المهمة البديلة بدلا من ذلك. سيتم حماية هويتك في جميع الأوقات ولن يتم ربط أي بيانات منشورة من هذا البحث مرة أخرى باسمك. سيتم تعيين لك رمز وسيتم تحديدها فقط من خلال هذه التعليمات البرمجية في كل من التجربة وتحليل البيانات. سيتم حفظ جميع البيانات الشخصية التي تم جمعها في جهاز كمبيوتر محمي بكلمة مرور، منفصلة عن غيرها من البيانات التي تم جمعها للتجربة. سيتم الاحتفاظ بالاسم فقط حتى نهاية المدة لإعلام مدربي مشاركتك بتعيين الرصيد الإضافي المناسب. سيتم حذف جميع الملفات، التي تربط إلى اسمك، في نهاية المدة.

المخاطر / الفوائد:

لاتوجد مخاطر متوقعة أو فوائد من المشاركة في هذه الدراسة، ومع ذلك، يمكن أن تسفر النتائج من هذه الدراسة. عن مزيد من المعرفة في اكتساب اللغة الثانية والكلام غير الأم وتعلم اللغه.وبالمثل، يمكن لهذا البحث أيضا أن يوفر المزيد من التبصر في علم الأمراض الكلام، وتطوير أدوات تعلم اللغة أفضل والبرامج في مختلف مؤسسات التعلم اللغوي المختلفة. معلومات الاتصال: الباحثون الذين أجروا هذه الدراسة هم شيا-ني (جيني) شين وميليسا M. بيس-بيرك، Ph.D. إذا كان لديك أسئلة أخرى أو مشاكل متعلقة بالبحوث، يمكنك الاتصال بشيا ني (جيني) شين، jshen@uoergon.edu‏ 2149-506 (971) أو الدكتور ميليسا بيس بيرك، mbaesebe@uoregon.edu‏ 3899- 346 (541). إذا كانت لديك أسئلة تتعلق بحقوقك كموضوع بحثي، يرجى الاتصال ببحوث خدمات الامتثال، 5237 جامعة أوريغون، أو يوجين، أو 97403، أو (541) 346-2510، أو البريد الإلكتروني ResearchCompliance@uoregon.edu. نسخة من نموذج الموافقة: يمكنك حفظ نسخة من نموذج الموافقة هذا إلى جهاز الكمبيوتر الخاص بك للحفاظ على السجلات لك والرجوع إليها في المستقبل.

تصريح مشاركة:

البيانات من خلال تحديد هذا المربع، تشير إلى أنك قرأت وفهمت المعلومات المقدمة أعلاه؛

- أنك توافق طوعا على إتاحة تسجيلاتك لعلماء آخرين في مستودع عام.
- أن مسؤولي الجامعة قد يكون الوصول إلى البيانات.
- أن يتم تسجيل هذه التسجيلات بحيث لن يتم ربط اسمك بهذه التسجيلات.
- أن مقتطفات من التسجيلات يمكن استخدامها في العروض المؤتمر أو الفصول الدراسية لأطروحة الشرف المحقق.

موافقتك:

من خلال تحديد هذا المربع، تشير إلى أنك قرأت وفهمت المعلومات المقدمة أعلاه؛
- أنك توافق طوعا على المشاركة؛
- أنك قد تسحب موافقتك في أي وقت وتوقف المشاركة دون عقوبة؛
- أن تتلقى نسخة من هذا النموذج إذا رغبت في ذلك؛ و
- أنك لا تتنازل عن أي مطالبات أو حقوق أو تعويضات قانونية.

Appendix B: Language Experience Questionnaire

<u>Language Experience Questionnaire</u>

Subject ID (for lab use only) _____

Year of Birth _____

Gender _____

Ethnicity (please circle one)

      Hispanic Latino   Non-Hispanic Latino

Race (please circle one)

        American Indian or Alaskan Native   Asian

        White

        Native Hawaiian or other Pacific Islander   African American or Black

Country of Origin (please circle one)

China   Taiwan   Japan   Saudi Arabia

Vietnam

Other _____

Demographic Information

Please list all the places you have lived for two years or longer.

Town, State/ Province, Country	Age (e.g. 0 to 12)

Linguistic Knowledge

Please list all languages aside from English you have learned/studied.

Language	Age you first start learning this language?	How long did you study this language?	Where did you learn this language?

Educational Information

Please enter the primary language used at each education level.

Elementary School	
Middle School	
High School	
University	
Graduate School	

Have you taken any linguistic and/or English classes? If yes, which ones?

Course Name	Year Taken

Learning Impairments
Have you ever had any speech, hearing, visual, or learning impairments? If yes, please explain…

_____

_____

_____

_____

_____

Are you…
A native speaker of American English? (Y/N)
A native speaker of another English dialect (Y/N)   If yes, which one?
_____
A non-native English speaker (Y/N)

English Proficiency
Please provide the most recent version of the information below if available

Test Type	Year Taken	Score
TSE		
TOEFL		
TOEIC		
IELTS		

Please list all the English speaking countries/places you have visited.

Location	Length of Stay	Initial Age

Have you studied abroad in another country? If yes, where? and for how long?

Location	Length of Stay	Initial Age

English Experience

At what age did you begin to study/learn English? _____
How many years did you study English in your home country? _____
Did you practice speaking English with a native English speaker in your home country? (Y/N)
       If yes, how many hours per week did you spend speaking English? _____
Are you currently living with a host family? (Y/N)
       If yes, how many hours per week do you spend speaking English? _____
Are you currently taking any ESL or AEI classes? _____
Do you have an ESL tutor? (Y/N)   If yes, for how long? _____
Do you have an English conversation partner? (Y/N)   If yes, for how long? _____

Appendix C: Debriefing form (English)

Background: In this study, we are interested in examining the effect audiovisual training, pairing sounds with video, has on second language learners' speech and language development. Language learning and mastery is essential in enabling an individual to express novel ideas and thoughts, and bridging relationships in and outside the classroom. While previous research has suggested audiovisual training to be beneficial in strengthening language perception among dyslexic children, little research has been done examining its effect for second language learners.

Purpose of This Study: We are hoping to understand whether audiovisual training will have a positive and/or negative influence on second language learners, especially in terms of whether it is an effective tool for strengthening speech production and perception. This research also aims to assess non-native speaker's ability to produce, perceive, and distinguish sounds made with your lips, like 'b', 'm' or 'f', and whether audiovisual training is a helpful tool for processing novel or unfamiliar speech input.

Your Part: Your participation is important as it gives us a better idea on the effects of audiovisual training on second language learning and development. In addition, it also gives us further insight on speech and language development across second language learners from various backgrounds and experience levels.

Follow Up: Better understanding of audiovisual training may yield further knowledge in second language acquisition and non-native speech perception and production. Likewise, results could also lead to improved development of speech and language tools across different learning institutions.

If you have questions or would like to learn more about this research, please contact Chia-Ni (Jennie) Shen jshen@uoregon.edu (971) 506-2149 or Dr. Melissa Baese-Berk, mbaesebe@uoregon.edu (541) 346-3899. If you have any questions regarding your rights as a research participant, please contact Research Compliance Services, 5237 University of Oregon, Eugene, OR 97403, 541-346-2510, or email ResearchCompliance@uoregon.edu. If you are interested in more psychology and linguistics research please contact the Human Subjects Coordinator at hscoord@uoregon.edu.

Thank you for participating!

Appendix C: Debriefing form (Simplified Chinese)

研究理念：
在这项研究中，我们将探讨视听训练的实质效果，视频与声音的搭配对于第二语言学习者的口语表达能力和语言提升的影响。语言的学习和掌握，对于使个人能够表达出新颖的观点和想法，亦或是在课堂内外间的沟通交流都有着相当重要的影响，虽然以前的研究显示视听训练有助于加强阅读障碍儿童的语言感知，但很少有研究检测其对第二语言学习者的影响。

研究目的：
我们希望能了解视听训练是否会对第二语言学习者产生积极的和/或消极的影响，尤其它是否是增强言语表达和感知的有效工具。这项研究旨在评估非母语者的表达能力，感知和区分出用嘴唇发出的声音（如 "b", "m" 或 "f"）的能力，以及视听训练是否有助于学习者面对新的或不熟悉的词语。

关于受测者：
您的参与非常重要，因为您的受测结果让我们能够更加的了解视听训练对于第二语言学习和发展的影响。此外，它还使我们能够进一步了解来自不同背景和经验水平的第二语言学习者的口语和语言发展。

期许：
对视听训练的更好理解，可能会让我们对『第二语言的学习和非母语的口语感知和语言表达』产生更进一步的知识。同样地，结果也可能会带动或改进不同学习机构中的提升语言学习的工具。

如果您有任何问题或想了解更多关于这项研究的信息，请联系Chia-Ni (Jennie) Shen jshen@uoregon.edu (971) 506-2149 或 Dr. Melissa Baese-Berk, mbaesebe@uoregon.edu (541) 346-3899。如果您对于作为研究受测者的权力有任何疑问，请联系Research Compliance Services, 5237 University of Oregon, Eugene, OR 97403, 541-346-2510, 或 email ResearchCompliance@uoregon.edu。如果您对于更多的心理学和语言

学研究感兴趣，请联系Human Subjects Coordinator at
hscoord@uoregon.edu。

感谢您的参与和配合。

Appendix C: Debriefing form (Traditional Chinese)

研究理念：

在這項研究中，我們將探討視聽訓練的實質效果，視頻與聲音的搭配對於第二語言學習者的口語表達能力和語言提升的影響。語言的學習和掌握，對於使個人能夠表達出新穎的觀點和想法，亦或是在課堂內外間的溝通交流都有著相當重要的影響，雖然以前的研究顯示視聽訓練有助於加強閱讀障礙兒童的語言感知，但很少有研究檢測其對第二語言學習者的影響。

研究目的：

我們希望能了解視聽訓練是否會對第二語言學習者產生積極的和/或消極的影響，尤其它是否是增強言語表達和感知的有效工具。這項研究旨在評估非母語者的表達能力，感知和區分出用嘴唇發出的聲音（如 "b", "m" 或 "f"）的能力，以及視聽訓練是否有助於學習者面對新的或不熟悉的詞語。

關於受測者：

您的參與非常重要，因為您的受測結果讓我們能夠更加的了解視聽訓練對於第二語言學習和發展的影響。此外，它還使我們能夠進一步了解來自不同背景和經驗水平的第二語言學習者的口語和語言發展。

期許：

對視聽訓練的更好理解，可能會讓我們對『第二語言的學習和非母語的口語感知和語言表達』產生更進一步的知識。同樣地，結果也可能會帶動或改進不同學習機構中的提升語言學習的工具。

如果您有任何問題或想了解更多關於這項研究的信息，請聯繫Chia-Ni (Jennie) Shen jshen@uoregon.edu (971) 506-2149 或 Dr. Melissa Baese-Berk, mbaesebe@uoregon.edu (541) 346-3899。如果您對於作為研究受測者的權力有任何疑問，請聯繫Research Compliance Services, 5237 University of Oregon, Eugene, OR 97403, 541-346-2510, 或 email ResearchCompliance@uoregon.edu。如果您對於更多的心理學和語言學研究感興趣，請聯繫Human Subjects Coordinator at hscoord@uoregon.edu。

感謝您的參與和配合。

Appendix C: Debriefing form (Japanese)

英語が第二外国語話者に対する視聴覚学習の影響、効果と報告書

背景：
視聴覚学習、映像と音声のペアリングがどのように第二外国語を使用する場合、また言語発達に影響があるかについてこの実験を行なっています。言語の学習や獲得には、斬新なアイデアや考えを個人が表現できるようになること、クラス内外で関係の架け橋になることが重要です。過去の実験結果から、視聴覚学習は失読症の小児の言語認識強化に有益であると示唆されていますが、第二外国語学習者への影響を調べる研究はほとんど行われていません。

研究目的：
特に音声の生成や認識を強化するための効果的な方法であるかどうかに関して、視聴覚学習が第二外国語学習者に効果的なのか理解したいと考えています。また、この研究では、第二外国語話者が唇で生成される、「b」、「m」または「f」のような音を発声、認識または、区別することができるかどうか、もしくは視聴覚学習が不慣れな音声入力に効果的かどうかを評価します。

被験者の役割：
視聴覚学習が第二外国語の学習と向上に影響を与えるという仮説に対して、被験者の実験結果は重要な役割を果たします。さらに、被験者それぞれの様々な経歴や経験から、第二外国語学習への会話や言語発達についてさらなる理解が期待されます。

補足：
視聴覚学習へのさらなる理解は、第二外国語獲得や非母国語の発話、認識への知識がさらに深まることが期待されます。同様に、得られた結果から異なる学習状況において発話、言語を向上させるツールをより良いものにする可能性もあります。

この研究について知りたいことや、ご心配なことがありましたら、以下の担当

者にご相

談下さい。

Chia-Ni (Jennie) Shen, jshen@uoergon.edu (971) 506-2149
Dr. Melissa Baese-Berk, mbaesebe@uoregon.edu (541) 346-3899.
Research Compliance Services, 5237 University of Oregon, Eugene, OR 97403,
(541) 346-2510, email: ResearchCompliance@uoregon.edu.

他の心理学・言語学の実験に興味がある方は、Human Subjects Coordinator

hscoord@uoregon.edu　までご連絡ください。

実験ご参加ありがとうございました。

Appendix C: Debriefing form (Arabic)

خلفية: الاقتران الأصوات مع الفيديو، التدريب السمعي البصري تأثير دراسة في مهتمون ونحن الدراسة، هذه في
الأفكار عن للتعبير الفرد تمكين في ضروري والتمكن اللغة تعلم التنمية. واللغة اللغة الثانية اللغة على لديه
التدريب اقترحت قد السابقة البحوث أن حين في الدراسية. الفصول وخارج داخل العلاقات وسد الجديدة، والأفكار
من القليل إجراء تم وقد القراءة، عسر لديهم الذين الأطفال بين اللغة تصور تعزيز في مفيدة لتكون البصري السمعي
الثانية. اللغة لمتعلمي تأثيره دراسة البحوث.

الدراسة هذه من الغرض: أو / و إيجابي تأثير له سيكون البصري السمعي التدريب كان إذا ما نفهم أن نأمل نحن
والإدراك. الكلام إنتاج لتعزيز فعالة أداة هو كان إذا ما حيث من وخاصة الثانية، اللغة المتعلمين على سلبي
مثل شفتيك، من المصنوعة الأصوات وتمييز إدراك، و إنتاج، على الأم غير المتكلم قدرة تقييم إلى أيضا البحث هذا
مفيدة أداة البصري السمعي التدريب كان إذا وما و، أو م ب،
الكلام. إدخال مألوفة غير أو رواية لمعالجة.

بك الخاص الجزء: اللغة تعلم على البصري السمعي التدريب آثار عن أفضل فكرة تعطينا لأنها مهمة مشاركتكم
اللغة المتعلمين عبر واللغة اللغة تطوير في التبصر من مزيد أيضا يعطينا فإنه ذلك، إلى وبالإضافة والتنمية. الثانية
الخبرة. ومستويات الخلفيات مختلف من الثانية.

المتابعة: وإدراك الثانية اللغة اكتساب في المعرفة من مزيد إلى البصري السمعي للتدريب أفضل فهم يؤدي قد
الأصلي غير الكلام
تعليمية مؤسسات عبر واللغة الكلام أدوات تطوير تحسين إلى أيضا النتائج تؤدي أن يمكن وبالمثل، والإنتاج.
مختلفة.

شين (جيني) شيا-ني الاتصال يرجى البحث، هذا عن المزيد معرفة في ترغب أو أسئلة لديك كان إذا
jshen@uoregon.edu (971) 506-2149
أو الدكتور ميليسا بيس بيرك 346-3899 (541) mbaesebe@uoregon.edu
5237 الامتثال، مجال في البحث بخدمات الاتصال فيرجى بحثي، كمشارك بحقوقك تتعلق أسئلة أية لديك كانت إذا
الإلكتروني البريد أو ،2510-346-541 أو ،97403 يوجين، أو أوريغون، جامعة
ResearchCompliance@uoregon.edu. البحوث واللغويات النفس علم من مزيد في مهتما كنت إذا
hscoord@uoregon.edu. في البشري المنسق الاتصال يرجى
لمشاركتك إشكرا!

64

Appendix D: Recruitment Scripts (English)

Recruitment Scripts (E-mail):

Hello,

I am emailing about an opportunity to participate in research conducted by the Linguistic Department at the University of Oregon as part of my honors thesis. I am interested in how audiovisual learning affects second language acquisition and learning. This research will take place in the the CMER building at 722 E. 11ᵗʰ Avenue and 1600 Millrace after fall term. The experiment is simple and easy to complete. Students who choose to participate will be asked to listen and/or watch a series of words and be asked to make judgments about them. It will happen over the course of 2 days, with an hour commitment per day. After completion, participants will receive 2-3% extra credit toward their final grade. If you have any questions, or are interested in participating, please contact Chia-Ni (Jennie) Shen at jshen@uoregon.edu.

Thank you!

Recruitment Script (In Class)

Hello,

My name is Chia-Ni (Jennie) Shen, a senior in the Honors College majoring in Communication Disorders & Sciences. I'm here today to tell you about an opportunity to participate in research conducted by the Linguistic Department at the University of Oregon as part of my honors thesis project. I am interested in how audiovisual learning affects second language acquisition and learning. This research will take place in the CMER building at 722 E. 11ᵗʰ Avenue and 1600 Millrace after fall term. The experiment is simple and easy to complete. Students who choose to participate will be asked to listen and/or watch a series of words and be asked to make judgments about them. The experiment will happen over the course of 2 days, with an hour commitment per day. After completion, participants will receive 2-3% extra credit toward their final grade.

Does anyone have any questions?

If you're interested in participating please sign this sheet, or email me at jshen@uoregon.edu.

Appendix D: Recruitment Scripts (Simplified Chinese)

有偿志愿者招募启示(E-mail):

您好：
藉由这封电子邮件通知您一个可以参与目前正在University of Oregon语言学系进行的研究的测验机会，研究项目为『视听学习对第二语言的口语表达能力和语言提升的影响』；这项研究将在1600 Millrace举行。这个测验相当简单，并不困难，选择参加的学生将被要求听或者观看一系列的词语，并被要求对它们做出判断。测验将分成两天，每天大约一小时。完成测验后，参与者可获得20美元。如果您有任何问题或有兴趣参与，请联系研究人员jshen@uoregon.edu。

谢谢您。

有偿志愿者招募启示(In class):

您好：
我的名字是_____，在学校里的言语感知与语言表达实验室担任研究助理。我今天将在这里告知您有机会参与由University of Oregon语言学系正在进行的研究，研究项目为『视听学习对第二语言的口语表达能力和语言提升的影响』；这项研究将在1600 Millrace举行。这个测验相当简单并不困难，选择参加的学生将被要求听或者观看一系列的词语，并被要求对它们做出判断。测验将分成两天，每天大约一小时。完成测验后，参与者可获得20美元。有任何问题欢迎提出。
如果您还有任何问题或有兴趣参与，请在这张表格上签名或者联系研究人员jshen@uoregon.edu。

Appendix D: Recruitment Scripts (Traditional Chinese)

有償志願者招募啟示(E-mail):

您好：

藉由這封電子郵件通知您一個可以參與目前正在University of Oregon
語言學系進行的研究的測驗機會，研究項目為『視聽學習對第二語言
的口語表達能力和語言提升的影響』；這項研究將在1600 Millrace舉
行。這個測驗相當簡單，並不困難，選擇參加的學生將被要求聽或者
觀看一系列的詞語，並被要求對它們做出判斷。測驗將分成兩天，每
天大約一小時。完成測驗後，參與者可獲得20美元。如果您有任何問
題或有興趣參與，請聯繫研究人員jshen@uoregon.edu。

謝謝您。

有償志願者招募啟示(In class):

您好：

我的名字是 ＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿ ，在學校裡的言語感知與語
言表達實驗室擔任研究助理。我今天將在這裡告知您有機會參與由
University of Oregon語言學系正在進行的研究，研究項目為『視聽學
習對第二語言的口語表達能力和語言提升的影響』；這項研究將在
1600 Millrace舉行。這個測驗相當簡單並不困難，選擇參加的學生將
被要求聽或者觀看一系列的詞語，並被要求對它們做出判斷。測驗將
分成兩天，每天大約一小時。完成測驗後，參與者可獲得20美元。

有任何問題歡迎提出。

如果您還有任何問題或有興趣參與，請在這張表格上簽名或者聯繫研究人員jshen@uoregon.edu。

Appendix D: Recruitment Scripts (Japanese)

実験参加依頼（メール）：

こんにちは。
現在、オレゴン大学の言語学科では、視聴覚による学習が第二外国語の獲得と
学習に与える影響について研究を行っています。今回は下記の内容の実験への
ご参加をお願いしたく、メールさせていただきました。

****
****
概要：一連の単語を視聴し、それらについて簡単な質問に答えていただきま
す。
所要時間：約一時間　二日間
謝礼：２０ドル

場所: 1600 Millrace Dr, Eugene, OR 97403

****
****

ご不明な点がある場合や参加していただける方は、担当者jshen@uoregon.eduま
でご連絡ください。

よろしくお願いいたします。

実験参加依頼（クラス内）

こんにちは。
私は、大学の音声知覚＆プロダクション研究所の研究助手の＿＿＿です。現
在、オレゴン大学の言語学科では、視聴覚による学習が第二外国語の獲得と学
習に与える影響について研究を行っています。今回は実験へのご参加をお願い

したく参りました。実験内容は 一連の単語を視聴し、それらについて質問に答えていただくという、単純かつ時間のかからないものです。実験は二日に分けて行われ、それぞれ一時間の参加が必要です。謝礼として２０ドルお渡しします。場所は、秋学期以降はthe CMER building （722 E. 11th Avenue and 1600 Millrace）で行われます。

不明な点はありますか？

参加していただける方は、この募集用紙に署名していただくか、担当者 jshen@uoregon.eduまでご連絡ください。

よろしくお願いいたします。

Appendix D: Recruitment Scripts (Arabic)

مخطوطات التوظيف (البريد الإلكتروني):

مرحبا،

أبعث لكم رساله الكترونية عن فرصة للمشاركة في البحوث التي أجرتها الإدارة اللغوية في جامعة ولاية أوريغون. نحن مهتمون في كيفية تأثير التعلم السمعي البصري على اكتساب اللغة الثانية والتعلم. وسيجري هذا البحث في مبنى سمر في E 722. الحادي عشر شارع و 1600 ميلريس بعد فترة الخريف. التجربة بسيطة وسهلة لإكمال. سوف يطلب من الطلاب الذين يختارون المشاركة الاستماع و / أو مشاهدة سلسلة من الكلمات ويطلب منهم إصدار أحكام بشأنها. وسوف يحدث على مدى يومين، مع التزام ساعة في اليوم الواحد. بعد الانتهاء، سيتلقى المشاركون 20 دولارا. إذا كان لديك أي أسئلة، أو ترغب في المشاركة، يرجى الاتصال المحقق في jshen@uoregon.edu.

شكرا لكم!

سيناريو التوظيف (في الصف)

مرحبا،

اسمي (اسم را) _____ ، مساعد باحث في تصور الكلام والإنتاج مختبر في الحرم الجامعي. أنا هنا اليوم لأقول لكم عن فرصة للمشاركة في البحوث التي أجرتها الإدارة اللغوية في جامعة ولاية أوريغون. نحن مهتمون في كيفية تأثير التعلم السمعي البصري على اكتساب اللغة الثانية والتعلم. سيتم إجراء هذا البحث في مبنى سمر في E 722. الحادي عشر شارع و 1600 ميلريس بعد فترة الخريف. التجربة بسيطة وسهلة لإكمال. سوف يطلب من الطلاب الذين يختارون المشاركة الاستماع و / أو مشاهدة سلسلة من الكلمات ويطلب منهم إصدار أحكام بشأنها. سوف تحدث التجربة على مدى يومين، مع التزام ساعة في اليوم الواحد. بعد الانتهاء، سيتلقى المشاركون 20 دولارا.

هل عند أحدكم أية أسئلة؟

إذا كنت ترغب في المشاركة يرجى التوقيع على هذه الورقة، أو البريد الإلكتروني للمحقق في jshen@uoregon.edu.

Appendix E: Research tokens

Below is a list of example tokens used for the experiment, which are various minimal pairs related to the consonants in question. Both real and nonsense words are used to test for participants' overall production and perception abilities.

Initial Consonant:
- Ban, pan, man, fan, van /æ/
- Bat, pat, mat, fat, vat   /æ/
- Beat, peat, meat, feat, veat   /i/
- Bit, pit, mit, fit, vit   /ɪ/
- Bore, pore, more, fore, vore /oʊ/
- Bond, pond, mond, fond, vond /oʊ/
- Back, pack, mack, fack, vack   /æ/
- Bye, pie, my, fie, vie   /aɪ/
- Bay, pay, may, fay, vay   /eɪ/
- Base, pace, mace, face, vase   /eɪ/
- Boat poat, moat, foat, vote   /oʊ/
- Bed, ped, med, fed, ved   /ɛ/

Final Consonant
- Lab, lap, lam, laugh, lav
- Cab, cap, cam, calf, (cav)
- Cob, cop, com, cough, cov
- Koob, coop, coom, coof, coov
- Robe, rope, roam, roaf, rove
- Rub, rup, rum, rough, ruv
- Ab, app, am, af, av.
- Lib, lip, lim, lif, live
- Gib, gip, gim, gif, give
- Lub, lup, lum, luf, luv
- Lube, loop, lume, loof, loov
- Gabe, gape, game, gafe, gave

Appendix F: Research sentences

Pre/post-test sentence examples:
1. He is driving a very fancy van.
2. He likes to have fun by playing basketball at the park.
3. She is poorly dressed for the party.
4. My plan for this paper is to bypass the boring introduction and focus on the main point.
5. I call dibs on the potato chips and fruit platter.
6. My little brother likes to stuff his face into a full plate of pie.
7. The boy likes to rub his tummy when he feels full.
8. Don't be a bore, pour some more!
9. Please turn off the fan and oven to save more power.
10. You have a beautiful voice.
11. The mansion is just up the road, past the bakery, between the factory and public library.
12. Can you bring potato chips to the potluck on Friday?

The rating scale for native speech perception will be
1: very hard to understand
2: hard to understand
3: somewhat okay to understand
4: easy to understand
5: very easy to understand

BIBLIOGRAPHY

Anderson, S., (2012). How many languages are there in the world? *Linguistic Society of America: Advancing the Scientific Study of Language*. Retrieved from https://www.linguisticsociety.org/content/how-many-languages-are-there-world

Baese-Berk, Melissa M., & Samuel, Arthur G. (2016). Listeners Beware: SpeechProduction May Be Bad for Learning Speech Sounds. *Journal of Memory and Language, 89*, 23-36. doi: 10.1016/j.jml.2015.10.008

Best, Catherine. (1995). 'A direct realist view of cross-language speech perception,'. Speech Perception and Linguistic Experience: Issues in Cross Language Research. 171-203.

Best, C., & Avery, R. (1999). Left-Hemisphere Advantage for Click Consonants Is Determined by Linguistic Significance and Experience. *Psychological Science, 10*(1), 65-70.

Best, C., Mcroberts, G., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America, 109*(2), 775-94.

Binturki, T. (2001) Analysis of Pronunciation Errors of Saudi ESL Learners. (Master's Thesis). Available from ProQuest Dissertations and Theses database. (UMI No. 1456863)

Bradlow, A., Pisoni, D., Akahaneyamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English vertical bar r vertical bar and vertical bar l vertical bar .4. Some effects of perceptual learning on speech product. *Journal Of The Acoustical Society Of America, 101*(4), 2299-2310.

Duanmu, S. (2000). The Phonology of Standard Chinese. Oxford: Oxford University Press

Edwards, H.T. (1992). Applied phonetics: The sounds of American English. San Diego: Singular Publishing Group, Inc.

Faris, Best, Tyler (2016). An examination of the different ways that non-native phones may be perceptually assimilated as uncategorized. *The Journal of Acoustical Society of America 139*(1):EL-EL5.

Fenwick, Best, Davis, & Tyler. (2017). The influence of auditory-visual speech and clear speech on cross-language perceptual assimilation. *Speech Communication, 92*, 114-124.

Flege, J. & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language & Speech , 24 ,* 125-146.

George, D., & Mallery, M. (2010). SPSS for Windows Step by Step: A Simple Guide and Reference, 17.0 update (10a ed.) Boston: Pearson.

Hanulíková, Adriana, Dediu, Dan, Fang, Zhou, Bašnaková, Jana, & Huettig, Falk. (2012). Individual Differences in the Acquisition of a Complex L2 Phonology: A Training Study. *Language Learning: A Journal of Research in Language Studies,62*(2[Supplement]), 79-109. doi: 10.1111/j.1467-9922.2012.00707.x

Hattori, K. (2010). Perception and production of English /r/-/l/ by adult Japanese speakers. Doctoral thesis, UCL.

Hazan, Sennema, Iba, & Faulkner. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication, 47*(3), 360-378.

Hirata, Yukari, & Kelly, Spencer 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. doi: 10.1044/1092-4388

Jamieson, D., & Morosan, D. (1986). Training non-native speech contrasts in adults: Acquisition of the English /ð/-/θ/ contrast by francophones. *Perception & Psychophysics,40*(4), 205-215.

Kartushina, N., Hervais-Adelman, A., Frauenfelder, U., & Golestani, N. (2015). The effect of phonetic production training with visual feedback on the perception and production of foreign speech sounds. *The Journal of the Acoustical Society of America, 138*(2), 81732. doi: 10.1121/1.4926561

Kujala, Karma, Ceponiene, Belitz, Turkkila, Tervaniemi, & Näätänen. (2001). Plastic Neural Changes and Reading Improvement Caused by Audiovisual Training in Reading Impaired Children. *Proceedings of the National Academy of Sciences of the United States of America, 98*(18), 10509-10514. doi: 10.1073/pnas.181589198

Liberman, Harris, Hoffman, & Griffith. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54*(5), 358-368.

Lidestam, B., Moradi, S., Pettersson, R., & Ricklefs, T. (2014). Audiovisual training is better than auditory-only training for auditory-only speech-in-noise identification. *The Journal of the Acoustical Society of America, 136*(2), EL142-7.

Lim, S., & Holt, L. (2011). Learning Foreign Sounds in an Alien World: Videogame
Training Improves Non-Native Speech Categorization. *Cognitive Science, 35*(7),
1390-1405.

Lively, S., Logan, J., & Pisoni, D. (1993). Training Japanese listeners to identify
English r and l. 2. The role of phonetic environment and talker variability in
learning new perceptual categories. *Journal Of The Acoustical Society Of
America, 94*(3), 1242-1255.

Lively, Se, Pisoni, Db, Yamada, Ra, Tohkura, Y, & Yamada, T. (1994). Training
Japanese listeners to identify English /r/ and /l/.3. Long term retention of new
phonetic categories. *Journal Of The Acoustical Society Of America, 96*(4), 2076-
2087.

Logan, J., Lively, S., & Pisoni, D. (1991). Training Japanese listeners to identify
English /r/ and /l/: A first report. *The Journal of the Acoustical Society of
America, 89*(2), 874-86.

MacKain, K., Best, C., & Strange, W. (1981). Categorical perception of English /r/ and
/l/ by Japanese bilinguals. Applied Psycholinguistics, 2(4), 369-390.

Magnan, Annie, Ecalle, Jean, Veuillet, Evelyne, & Collet, Lionel. (2004). The Effects
of an Audio-Visual Training Program in Dyslexic Children. *Dyslexia, 10*(2), 131
140. doi. 10.1002/dys.270

Mathew, Nalliveettil George, & Alidmat, Ali Odeh Hammoud. (2013). A Study on the
Usefulness of Audio-Visual Aids in EFL Classroom: Implications for Effective
Instruction. *International Journal of Higher Education, 2*(2), 86-92.

Mcgurk, H., & Macdonald, J. (1976). Hearing lips and seeing voices. *Nature*,
264(5588), 746-8.

Nasukawa, K. (2010). No consonant-final stems in Japanese verb morphology. *Lingua*,
120(10), 2336-2352.

Ortega, L. (2009). *Understanding second language acquisition* (Understanding
language series). London: Hodder Education. Print.

Peng, Xiaolan, Chen, Hui, Wang, Lan, & Wang, Hongan. (2018). Evaluating a 3-D
virtual talking head on pronunciation learning. *International Journal of Human
– Computer Studies*, 109, 26-40.

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H.,
Kastman, E., Lindeløv, J. (2019). PsychoPy2: experiments in behavior made
easy. *Behavior Research Methods*. 10.3758/s13428-018-01193-y

Sekiyama, Burnham, Tam, Erdener (2003). Auditory-visual speech perception development in Japanese and English Speakers. In *AVSP* 2003, 43-47.

Shinohara, & Iverson. (2018). High variability identification and discrimination training for Japanese speakers learning English /r/–/l/. *Journal of Phonetics, 66*(C), 242-251.

Sumby, W., & Pollack, I. (1954). Visual Contribution to Speech Intelligibility in Noise. *The Journal of the Acoustical Society of America, 26*(2), 212-215.

Tateishi, M. (2013). Effects of the Use of Ultrasound in Production Training on the Perception of English /r/ and /l/ by Native Japanese Speakers. (Master's Thesis). Available from University of Calgary Library. uri: hdl.handle.net/11023/1097

Thompson & Hazan. (2010). The impact of visual cues and lexical knowledge on the perception of a non-native consonant contrast for Colombian adults. *Speech, Hearing and Phonetic Sciences, UCL, London, UK.*

Tyler, M., Best, C., Faber, A., & Levitt, A. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica, 71*(1), 4-21.

Veuillet, E., Magnan, A., Ecalle, J., Thai-Van, H., & Collet, L. (2007). Auditory processing disorder in children with reading disabilities: Effect of audiovisual training. *Brain : A Journal of Neurology, 130*(Pt 11), 2915-28. doi: 10.1093/brain/awm35

Wang, Y., Spence, M., Jongman, A., & Sereno, J. (1999). Training American listeners to perceive Mandarin tones. *The Journal of the Acoustical Society of America, 106*(6), 3649-58.

Wang, X. (2013). Perception of Mandarin Tones: The Effect of L1 Background and Training. *Modern Language Journal, 97*(1), 144-160.