2017-06-01

An Acoustical Analysis of the American English /l, r/ Contrast as Produced by Adult Japanese Learners of English Incorporating Word Position and Task Type

Braden Paul Chase
Brigham Young University

An Acoustical Analysis of the American English /l, r/ Contrast as Produced by Adult Japanese Learners of English Incorporating Word Position and Task Type

Braden Paul Chase

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Arts

Mark W. Tanner, Chair
Shawn L. Nissen
K. James Hartshorn

Department of Linguistics and English Language

Brigham Young University

ABSTRACT

An Acoustical Analysis of the American English /l, r/ Contrast
as Produced by Adult Japanese Learners of English
Incorporating Word Position and Task Type

Braden Paul Chase
Department of Linguistics and English Language, BYU
Master of Arts

Adult Japanese learners of English (JLEs) are often stereotyped as being unable to produce or perceive the English phonemes /l/ and /r/. This study analyzed acoustic samples of /l/ and /r/ obtained from intermediate-level Japanese speakers in two variable contexts: word positions (initial/final) and task type (controlled/free). These tokens were subjected to acoustic analysis which is one way of comparing oral productions of native and non-native English speakers. Previous research has identified a lowered third formant (F3) as the hallmark of an American English /r/ as produced by a native speaker, independent of word position or task type. The results indicate that participants can produce appropriate and statistically significant differences (p<.001) between these two phonemes across word position and task type. Other findings indicate that neither task type nor word position had a significant effect on F3 values. These results indicate that Japanese speakers of English may have the ability to distinguish /l/ from /r/ without specialized pronunciation training, but these differences are less dramatic as identified by F3 frequency values that those produced by native English speakers when producing these contrasting phonemes. In most tokens, however, large effect sizes remained between JLE productions and NES standards.

Keywords: acoustic analysis, Japanese, /l, r/, third formant (F3), task type,

ACKNOWLEDGEMENTS

I wish to express thanks to my thesis chair, Dr. Mark Tanner, and committee, Dr. Shawn Nissen and Dr. James Hartshorn, for their time, attention, and guidance during this thesis. I also wish to express gratitude to my mentor and good friend Judson Hart who help set me on this path. My parents, Paul Chase and Renee Adams, made education a priority and provided the desire to pursue my education. For that, I am extremely grateful. Finally, I would like to express appreciation to my wife, Luciane, for her love, support, trust, vision, and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Research into the select phonological difficulties that Asian speakers encounter when learning English as a second or foreign language typically identifies the /l/ and /r/ contrast as one which presents difficulty (Avery & Ehrlich, 1992). Very few Asian languages contain either the /l/ or the /r/ (Swan & Smith, 2001) and of the minute number that do, those productions are not acoustically identical to /l/ and /r/ productions in American English, nor are they perceived the same (Rochet, 1995).

Though English language learners from most Asian language backgrounds have difficulty with the perception and production of the American English /l, r/ contrast, it is well documented that native Japanese speakers experience extreme difficulty in accurately perceiving or producing the /l, r/ contrast in American English (Flege, 1995; Miyawaki, Strange, Verbrugge, Liberman, Jenkins, & Fujimura, 1975). As the Japanese phonological inventory does not contain this contrast, the struggle of these second language (L2) learners to master accurate production and perception of these sounds in English was originally explained through contrastive analysis which argued that second language learners will have more difficulty learning phonemes that are not already extant in their native language (Ellis, 1994). However, this hypothesis does not fully explain the seeming atypical difficulty that Japanese speakers experience while learning this sound contrast in English, especially as compared to learners from other major Asian language backgrounds which also do not have the same /l, r/ distinction as American English (Swan & Smith, 2001). Despite years of focused attention and practice, mastery of the native-like pronunciation of the /l, r/ contrast remains a constant challenge for most Japanese speakers learning English.

Rationale and Purpose

For over 40 years, the American English /l, r/ contrast has been studied specifically as it relates to Japanese speakers (Goto, 1971; McClelland, Fiez, McCandliss, 2002; Saito & Lyster, 2012). Most of that research has focused on accurate perception of the phonemes (Godfrey, 1983; Goto, 1971; Ingvalson, Holt, & McClelland, 2012; Miyawaki et al., 1975) with a smaller amount focusing on accurate production of the phoneme (Logan, Lively, Pisoni, 1991; Saito & Lyster, 2012). Even after decades of research and despite creative teaching methods and techniques (Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; McClelland et al., 2002), however, no consensus has emerged on how best to help adult Japanese learners of English (JLEs) to either perceive or produce the /l, r/ contrast (Takagi, 2002; Ingvalson et al., 2012). This uncertainty and variability regarding both perception and production of these phonemes has caused some researchers to suggest "that truly native-like identification of /r/ and /l/ may never be achieved by adult Japanese learners of English" (Takagi, 1995, 1996, 2002).

The difficulty posed by this contrast indicates that more empirical diagnoses of the /l, r/ contrast as produced by Japanese learners of English may be needed before appropriate prescriptions can be made. Recent technological innovations have created an opportunity for greater empiricism through use of acoustical analysis software, such as Praat, and technical computing languages, such as MatLab (Hisagi, Nishi, & Strange, 1998). As the prescriptive attempts to assist Japanese learners of English to acquire this contrast have only had limited success (Larsen-Hall, 2006), a descriptive approach using these new technologies may help researchers more precisely ascertain the nature of these JLE productions.

Previous research has indicated that the hallmark acoustic signature that distinguishes the /l, r/ contrast as produced by native speakers of American English (NESs) is a lowered third formant for the /r/ and a raised third formant for /l/ (Dalston, 1975; Saito & Lyster, 2012). Using the standards established by previous research as a starting point, the current study was designed to empirically assess the third formant frequency values of /l/ and /r/ as produced by female JLEs as a function of word position (initial and final) and task type (controlled versus spontaneous).

Definitions

As acoustical analysis contains terminology that is not common in other language research areas, this section outlines the following definitions and explanations to provide context for this study and to facilitate the discussion and interpretation of its results.

Formant – Sounds from the human vocal tract have a base pitch and several corresponding overtone pitches. These overtones give each sound its distinctive quality and are necessary for distinguishing them (Ladefoged, 2006). These overtones can be seen in a spectrogram as dark bands of energy and are called formants. Spectrograms are produced by acoustic analysis software such as Praat (Ladefoged, 2003; Boersma & Weenik, 2013).

In Figure 1, the dark horizontal bands represent the overtones that constitute the formants of one production of /i/. Starting from the bottom of the spectrogram, the darkest band is the first formant and is labeled F1. Each successive band is labeled sequentially, formant 2 (F2), formant 3 (F3), and formant 4 (F4). Thus, the value 3303 Hz identifies the 3rd formant (F3) of this production of /i/. Each dot represents a frequency value sampled approximately every six milliseconds by the Praat algorithm for that specific formant.

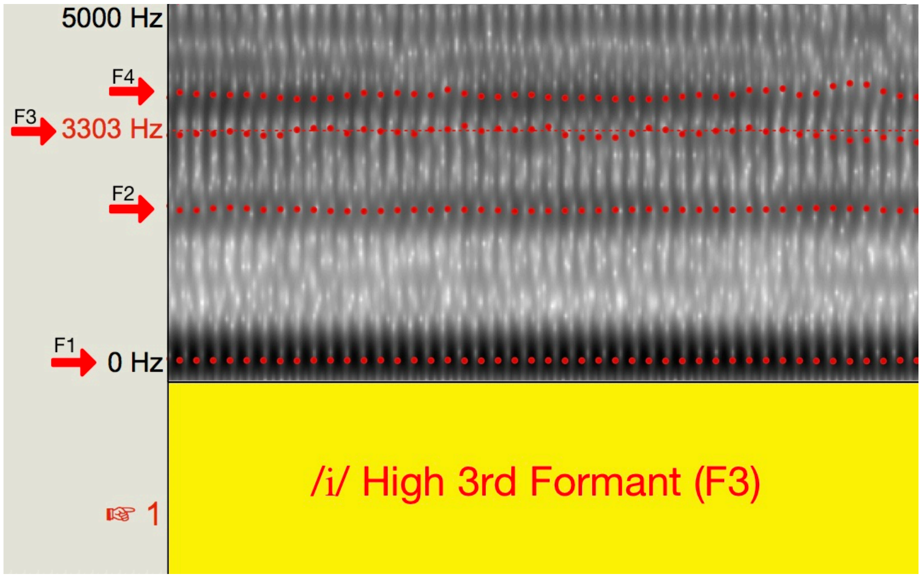

Figure 1. Spectrogram of NES-produced /i/ with labeled formant bands identifying the raised F3.

Segmental – Individual vowel and consonant sounds (Celce-Murcia, Brinton, and Snow, 2014).

Nucleus – The core sound of a syllable. In English and Japanese, the nucleus is almost always a vowel (Ladefoged, & Johnson, 2014).

Onset – The consonant sounds that occur before the nucleus (Ladefoged, & Johnson, 2014).

Coda – The consonant sounds that occur after the nucleus (Ladefoged, & Johnson, 2014).

Coarticulation – A pronunciation pattern where the vocal tract is producing sounds whose production overlaps acoustically (Hardcastle, & Hewlett, 2006).

Citation form – Language produced verbatim by the speaker while reading a list of target words in a carrier phrase. For example, the target word could be "leaf" which would be embedded in the carrier phrase "Say……again." and presented to the participant as "Say leaf again."

Read speech – Language produced verbatim by the speaker where the speaker is reading a paragraph of printed text.

Free speech – Language spontaneously produced by the speaker in response to authentic stimuli such as conversations, questions, and storytelling.

LITERATURE REVIEW

"The /l/ and /r/ sounds are just really difficult. You're going to have to work on it for years and years before you get it right." This disheartening sentiment is often heard by adult Japanese speakers learning English when their attempts to accurately and consistently perceive or produce the /l, r/ contrast repeatedly fail. Many teachers and researchers have repeatedly tried to craft methods and techniques to assist their native Japanese students in acquiring this contrast. However, English language teachers regularly report that both modern techniques and their traditional precursors, which have proven effective with students from other language backgrounds, are ineffective for Japanese students (Breitkreutz et al., 2001).

From an instructor's perspective, the /l, r/ contrast for Japanese students may be difficult to teach for a number of reasons. One of the standard techniques to teach the distinction is by visualizing the articulation points using diagrams or pictures. Another means of demonstrating this would be with a spectrograph or ultrasound machine. Most classrooms are not equipped with this kind of equipment which is expensive and requires years of training before the results can be accurately interpreted. Therefore, the responsibility of demonstrating these articulation points is often left to the teacher. Common classroom techniques include using images such as drawings, models, or digital tools to aid in the visualization; however, such techniques have produced little success with adult JLEs (Hazan, Sennema, Iba, & Faulkner, 2005). Some teachers alternately produce the /l/ and /r/ phonemes independently or in minimal pairs containing these phonemes while pointing inside their mouths for the students to see. Physically, this is very difficult as both /l/ and /r/ are articulated with a tightly constrained mouth thus impeding students from seeing past the teacher's teeth. Culturally, such demonstrations can be awkward and uncomfortable for

Japanese students as many feel that the inside of the mouth should not be displayed in public (Asakawa, 2015; Bakić-Mirić, 2011; Genzberger, 1994).

To further understand the difficulty that Japanese students have with the /l, r/ contrast, some researchers have attempted to apply theoretical explanations to this phenomenon. Those explanations tend to center around the consonant-vowel (CV) phonological structure that forms the foundation of the Japanese language (Aoyama, 2004). English, by comparison, has a consonant-vowel-consonant (CVC) pattern which can be extended to become CCCVCCC, such as in the word "strength." The Japanese CV pattern, by contrast, prohibits any consonant from ending any word with the arguable exception of nasalization, represented by the grapheme ん. This CV pattern has served as supporting evidence for the Contrastive Analysis Hypothesis which explained the difficulty of the /l, r/ contrast by indicating that these consonants in word-final position are novel for adult Japanese learners and therefore more difficult to acquire (Lado, 1957). According to this perspective, accurate productions of English words with an /r/ in coda position (i.e. /dear/) are more difficult to learn and so should receive more emphasis during instruction. In his Speech Learning Model (SLM), Flege (1995, 1996) disagrees with this assertion and demonstrates that the impossibility of word-final /r/ in Japanese makes it more noticeable and therefore more easily acquired. Aoyama (2004) further concludes that word-final /r/ is more "phonetically dissimilar" and therefore more easily acquired than word initial /r/ for adult Japanese learners. For English teachers and adult Japanese learners of English alike, these very contradictory conclusions are a source of confusion, not clarity.

To better understand why such a focus has been placed on /l/ and /r/, it may be valuable to consider the functional load of this contrast. Functional load, which is distinct from cognitive load, is a notion that identifies and ranks phonemic contrasts within a language based on the number of minimal pairs distinguished by that contrast. A careful discussion of each component used in the calculation of functional load is provided by Brown (1991) who outlines the roles of variables such as frequency, part of speech, inflections, and phonemic similarity. King (1967) defined functional load as “a measure of the work which two phonemes…do in keeping utterances apart.” For example, /l/ and /r/ minimal pairs such as “lap” and “rap” are more frequent in English than /ð/ and /d/ minimal pairs such as “then” and “den” (Catford, 1987; Munro and Derwing, 2006). That increased frequency is one factor in the calculation that determines the /l, r/ contrast to have a higher functional load than the /ð, d/ contrast.

In practice, incorrect productions of high functional load contrasts tend to cause communication breakdown because the utterances are more frequently misinterpreted by the listener (Brown, 1991). In contrast, incorrect productions of low functional load contrasts less frequently cause communication breakdown and could be interpreted by listeners as accent (Munro and Derwing, 2006). Thus, functional load is one way of empirically ranking the relative “importance” of any phonological contrast as compared to all other contrasts within that language. A concordance of the scales used by Catford (1987) and Brown (1991) ranks the functional load of the /l, r/ contrast at a 9 out of 10, or extremely high. This high functional load rating suggests that inaccurate productions of /l/ and /r/ have a high probability of negatively affecting perception and causing breakdown in communication.

Functional load is particularly valuable in this research context as it applies to both the perception of a contrast as well as to its production. The perception/production interdependency has been a major component of previous research (Hahn & Dickerson, 1999) and functional load supports the view that the ability to appropriately perceive and produce the /l, r/ contrast is fundamental to fluency in English. Furthermore, this support provides both justification for further research and guidelines for pronunciation instruction (Munro & Derwing, 2006). Though functional load can be used to empirically support the notion that both perception and production are important, it makes no clear delineation regarding when each of these two aspects of language learning should occur.

A widely-observed phenomenon in both first language (L1) and second language (L2) acquisition is that of perception preceding production (Hahn & Dickerson, 1999). Conclusions from this observation have led some researchers to train adult JLEs through perception-focused methodologies (Lively et al., 1993; Logan et al., 1991) in the hopes of a generalization of their perceptual abilities to both production as well as novel phonemic contexts. These perception-focused teaching methodologies have demonstrated that the /l, r/ contrast is not just an issue of articulation or motor control. Adult JLEs also struggle to perceptually distinguish /l/ and /r/. Furthermore, perception of /r/ appears to be more variable than /l/, indicating that /r/ is more difficult for adult JLEs to acquire (Bradlow et al., 1997). Intra-participant variability can also be inconsistent, indicating that some participants can acquire the contrast to a greater degree than others, with no apparent pattern. After the training, percent correct productions of /r/ in word initial position *decreased* after instruction while percent correct productions of /l/ in word initial position *increased* indicating a continued uncertainty in the word initial position. However, while Bradlow et al. (1997) did show significant improvement in almost all other phonemic contexts, Pruitt (1993) pointed out that “the study did not systematically examine the relative effects of types of stimulus…or changes in the training procedures” and so it is uncertain what, exactly, improved the participants’ ability to perceive the contrast or what decreased the accuracy of coda /r/ productions.

In response to this uncertainty in the literature, some researchers have developed production-focused methodologies that have helped some Japanese students are able to accurately produce the /l, r/ contrast even without being able to accurately perceive the same contrast (Aoyama et al., 2004; Flege et. al., 2003; Takagi, & Mann, 1996). Other research has shown that averaged intelligibility scores for adult JLEs improved after segmental instruction (Flege, Takagi, & Mann, 1996). A problem with this study, however, was that researchers did not control for differences in length of exposure to English with some Japanese having over 13 years of immersion experience and others having less than one month. Furthermore, Flege, Takagi, and Mann (1996) did not differentiate between male and female speakers which have been shown to vary significantly regarding modes of speech such as voice pitch, intonation, and word choice (Campbell, 2007). As noted by Ayoama (2004), "Further research (for [l] and [r]) will be needed to examine fine phonetic differences between the native and non-native productions of these segments (p. 274)." It seems then that innovative use of technology in language research is required for future research.

An important innovation in production-focused research methodology has been the introduction of electropalatography (EPG) which is an electronic device inserted into the mouth that can track tongue movement and articulation points. Possibly the only study to use EPG to describe the /l, r/ contrast relative to Japanese speakers has shown that the Japanese /r/ has significant variation in production (Gibbon, 1991). For the two speakers involved in this study, there was no discernable pattern to the tongue movements when producing /l/ or /r/, though the male participant articulated /l/ and /r/ in a greater variety of ways than the female participant. The variation in production between participants seemed to be determined primarily by vowel context and word position, but no pattern of production was discernable. Though EPG devices provide a level of specificity previously unattainable, the small number of participants utilized in most studies to date limits the generalizability of these results. Furthermore, the expense of EPG technology has been prohibitive to additional research.

Acoustic analysis software is another technology that has become more robust and more accessible in recent years and its use in this area as a diagnostic tool could be considered innovative. Applications such as Praat have allowed for empirical acoustic data to be analyzed for academic research by providing an empirical tool to measure oral productions (Boersma & Weenink, 2013). Japanese participants in a recent study demonstrated significant improvement after a form-focused approach to teaching /r/ to intermediate through advanced adult JLEs (Saito & Lyster, 2012). In that study, acoustic analysis was used to empirically measure the pre- and post-test /r/ production of the Japanese participants and identified statistical significance for the changes in production.

However, Saito and Lyster's study (2012) has limitations on several independent factors. First, no distinction was made between the productions of male and female adult JLEs. This is significant as the relative frequency values for males and females relative to /l/ and /r/ can vary as much as 700 Hz (Campbell, 2007; Erwan, 2013; Karlsson, 1992). Furthermore, the greater variability in productions noted by Gibbon (1991) may have affected the reliability of those measurements. A second criticism is that acceptable productions of /l/ and /r/ were determined by five native English speakers who listened to a random selection of participants' /l/ and /r/ productions and then judged each production using a Likert scale containing options such as "definitely /r/," "mostly /r/," "mostly not /r/" and "definitely not /r/." Those productions judged to be "definitely /r/" underwent formant analysis with a resulting F3 average of 2250 Hz judged as "good-enough exemplars." It is probable that had previous research on NES production standards for /l/ and /r/ (Dalston, 1975) been used in their calculations, which indicate an F3 average of 2078 Hz, their results may have been different. Thus, considerable insight may be gained by a comparison of JLE /l/ and /r/ productions to NES /l/ and /r/ productions.

Some aspects of those productions may change depending on task type. Task type can be defined as the sociolinguistic function for which language is being used and it is a construct for evaluating the similarity of language to authentic language use (Riney, Takada, & Ota, 2000). Increasingly complex tasks have been shown to affect phonetic articulations with more complex tasks tending to reduce phoneme production accuracy (Riney et al., 2000; Saito & Lyster, 2012). Citation form, read aloud paragraphs, and free speech are variations of task which place a progressively greater cognitive load on the speaker and allow for less control of the speech production. A complete picture of JLE productions of /l/ and /r/ would require variation of task type.

In summary, the functional load principle as established by Catford (1987) and Brown (1991) and further researched by Munro and Derwing (2006), the perception studies by Bradlow, et al. (1997) and Aoyama (2004), and the production studies by Flege (1996) and Saito and Lyster (2012) all reaffirm the need for specialized research and pedagogy for adult JLEs learning the /l, r/ contrast.

Acoustic Analysis

To better understand the important factors in acoustically analyzing the /l/ and /r/ productions of adult Japanese learners of English, it is valuable to first understand how these sounds are produced by native English speakers (NESs). Decades of previous acoustic research on the /l, r/ contrast has shown that across word position, vowel contexts, gender, and age, this contrast in American English is distinguished acoustically by a lowered third formant (F3) for the /r/ and a raised F3 for /l/ (Dalston, 1975). This finding has been reaffirmed by Ladefoged and

Johnson (2014) and Thomas (2016). Figure 2 shows one /r/ production with a significantly lower F3 of 1723 Hz while Figure 3 shows a raised F3 for /l/ of 2833 Hz. As the F3 is the distinguishing acoustic characteristic for the /l, r/ contrast in American English, the analyses used in this study will focus on comparisons of F3 values in adult JLE speech.

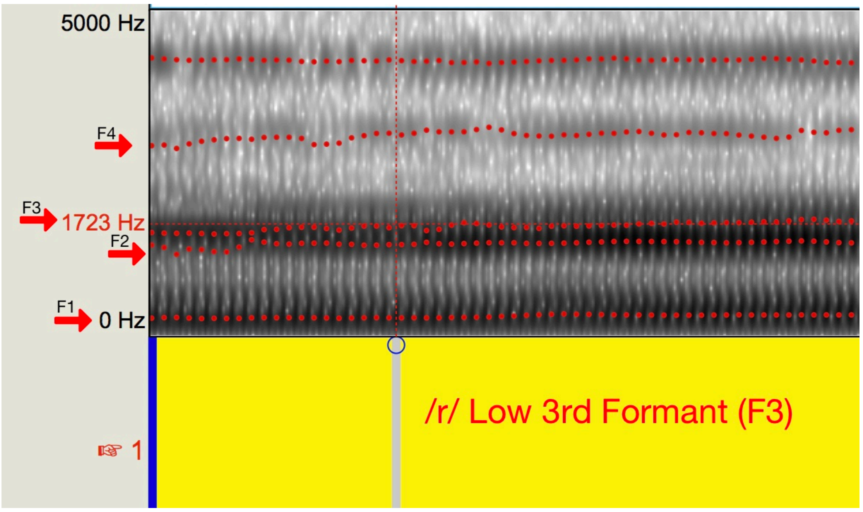

Figure 2. Spectrogram of NES-produced /r/ identifying the lowered F3 (1723 Hz).

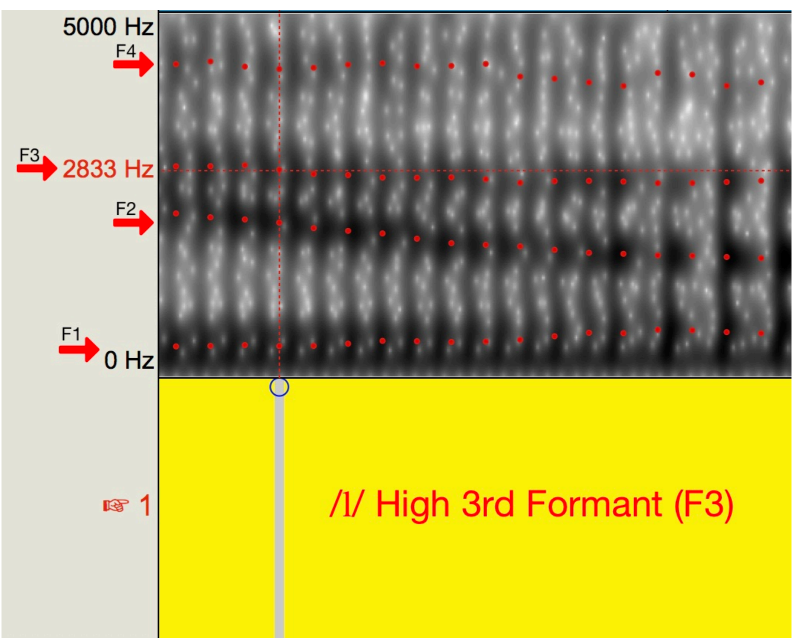

Figure 3. Spectrogram of NES-produced /l/ showing the raised F3 (2833 Hz).

Research Questions

This study was designed to provide empirical acoustic evidence that either substantiates or contradicts the common view that native Japanese speakers cannot accurately produce the American English /l, r/ contrast. With the approval of the Institutional Review Board (IRB), oral discourse was elicited to acoustically quantify the F3 trajectories of /l/ and /r/ articulations when produced by female adult JLEs as a function of word position and task type.

To more completely assess the relationships between each interaction, the research questions for this study have been condensed into the following null hypothesis and its correlates:

1. H_0 – There will be no significant difference between the F3 frequency values for female JLE when producing the /l/ and /r/ phonemes.
2. C_1 – Word position (initial or final) will have no significant effect on the F3 frequency values for /l/ or /r/.
3. C_2 – Task type (citation form, read speech, or free speech) will have no significant effect on the F3 frequency values for /l/ and /r/.

METHODS

To evaluate the null hypothesis regarding the F3 frequency values in each of the targeted contexts, a total of seven data-elicitation instruments were developed. This section of the paper will describe who the participants were, what information was gathered about their demographic make-up, what instruments and methods were used to elicit speech data from the subjects, and how the analyses were done.

Delimitations

The complex nature of acoustic analysis requires that many delimitations and controls be taken to ensure systematic study. The delimitations taken for this study are as follows:

- *Japanese participants* – To control for variance due to phonological background, only native Japanese speakers were selected to participate in this study (Henrich, Heine, & Norenzayan, 2010).
- *Female participants* – To control for gender differences in average voice pitch and speaking style, only productions by female speakers were analyzed (Hisagi, Nishi, & Strange, 1998; Riney, Takada, & Ota, 2000).
- *Age range* – To control for possible variation due to age, participant ages ranged from 23-36 with exact ages of 23, 23, 26, 27, and 36.
- *Relative proficiency* – To control for variation due to language proficiency, all participants were double-blind rated at the intermediate level according to the ACTFL Proficiency Guidelines (ACTFL, 2012), averaged across all four skills (listening, speaking, reading, and writing).

- *Vowel context* – To control for vowel co-articulation and to minimize vowel context anomalies, only productions of /l/ and /r/ immediately preceding or following an /ɪ/ or /i/ were analyzed (Hardcastle, & Hewlett, 2006).
- *Consonant context* – To control for possible consonant co-articulation and variation, consonant clusters with /l/ and /r/ were excluded from the analysis (Hardcastle, & Hewlett, 2006).
- *Word position* – To control for possible intervocalic variation caused by word-medial articulations, only word-initial and word-final tokens of /l/ and /r/ were analyzed (Hardcastle, & Hewlett, 2006).
- *Language instruction* – To control for variation due to language instruction or input exposure, all participants were chosen from the same English language school and had an average of nine years of English instruction (Guion, Flege, Akahane-Yamada, & Pruitt, 2000).
- *Concurrent treatment* – To our knowledge, these participants were not involved in any other study or treatment at the time of data collection (Washio & Houmanfar 2007).

Participants

Participants in this study were part of a larger pronunciation research project being conducted by a large-scale intensive English program (IEP). The data for this study was collected as part of a pre-test of proficiency. Because of the focus of this pronunciation study, participants included five intermediate-level Japanese females learning English as a second language in an IEP where they were in classes for 20 hours per week. All participants had completed a minimum of one semester at the IEP, but as traditional with many Japanese learners of English, they had had a minimum of eight years of English education in Japan (Aoyama, 2004; Kashihara, 2008). The participants of this study did not report any previous /l, r/-specific training. All participants showed normal maturation between semesters prior to acoustic sampling.

Pre-semester placement test scores put all participants at the intermediate proficiency level according to the ACTFL proficiency scale prior to their participation in the study. Participants ranged in age from 23-36 years old (*M* = 27 years old). All participants granted consent to participate in this study by signing a consent form that had been prepared according to IRB specifications and translated into Japanese.

Instruments

This section will describe the different instruments used to collect the speech data analyzed in this study. A total of seven instruments were designed to elicit both controlled and free speech from the participants. In addition to the data elicitation instruments, all participants completed a survey which captured demographic information regarding age, length of English study, and nationality.

Controlled speech. Previous studies (Boyce, 1997) have looked at non-sense words to gauge accurate production of individual phonemes. While such words may be nonsense in English, in a second language context those words could have meaning in the participants' L1 which could introduce an extraneous variable through L1 interference (Houmanfar et al., 2005; Osborne, 2001; Washio & Houmanfar, 2007). To avoid this issue, highly frequent, whole words in English containing the target phonemes were chosen as stimulus words.

Word lists. Citation form instruments took the form of three word lists of 36 words embedded in the carrier phrase "Say X again" where X was the word containing the target phoneme. This phrase was used to minimize prosodic declination. The sequence of stimulus words for each word list was also randomized among non-target words so that the participants "would not become aware" of the target phonemes (Dalston, 1975). The resulting instruments included a total of six productions of word initial /l/, three productions of word-final /l/, three productions of word initial /r/, and three productions of word final /r/ per participant. These word lists are included in Appendix A.

Read speech. The read-aloud speech stimulus took the form of two engineered paragraphs of 175 words each. Each paragraph contained all 36 words contained in the original word lists, including the five target words (dear, read, leash, leap, peel). The paragraphs were designed to be authentic examples of phrasing and real-use language. Paragraphs were included to simulate authentic readings in native language contexts; however, they were crafted for this study. These engineered paragraphs are included in Appendix B.

Free speech. It was predicted that spontaneous speech would have substantially lower production rates of the target phonemes thus two separate free speech contexts were developed to elicit a sufficient number of tokens. The two free speech contexts were 1) a fairy tale retelling and 2) a guided interview phase which utilized intermediate-level open-ended questions.

Fairy tale retelling. Participants were presented with a list of three common Japanese fairy tales plus the well-known European fairy tale, *Cinderella* from which they were asked to choose one fairy tale to retell. The names for all the stories were provided in both Japanese and English on the stimulus cards and a time limit of approximately five minutes was provided for the retelling task. The prompts given for this story-retelling task are given in Appendix C.

Open-ended questions. For the open-ended question (OEQ) stimulus, interviewers were previously trained in standard open-ended question interviewing techniques and instructed to continue probing students until the allotted time was fulfilled. The question instrument is included in Appendix D.

Data Collection Session

Setup and Equipment. Recording sessions were conducted in a 12x20 ft. room with carpeted walls, noise-isolating windows, and thick doors. Sound was captured using an Audio-Technica AT2020 Cardioid Condenser USB+ microphone connected to an Apple iMac ME086LL/A. All audio was recorded using Adobe Audition available in CS6. Background noise in the recording rooms was measured at -70db, a sufficiently low level so as not to adversely affect the recording of the elicited speech samples. As best as can be determined, no non-target sound influences were captured in the recorded audio.

Observations. The data collection sessions were allocated with three minutes for word lists, five minutes for paragraph reading, five minutes for fairy tale recall, and approximately seven minutes for OEQ for a total of approximately 20 minutes to complete the data elicitation portion of the study. The order and presentation of each instrument were in an effort to avoid an ordering effect in the data elicitation with some students doing word lists first, some doing the paragraphs first, some doing the OEQs first, and others doing the Fairy Tale recall first to avoid an ordering effect in the data elicitation.

Possible Hawthorne effects were minimized by having all sessions observed and recorded in the same way with all participants knowing that they were being recorded and observed. Participants demonstrated awareness of these observations by having them sign a consent form giving the researchers permission to elicit data under these conditions. Possible halo effects were also minimized by using non-teacher interviewers who had been trained in interviewing procedures and techniques. Participants had no contact with the interviewers before or after the interview and the three interviewers were randomized among the five participants. All tasks were explained to the participants before the start of the data collection session and each task was further explained and modeled before participants were asked to perform the task. While each session was randomly sequenced, the following section will outline general observations regarding the recording session protocol and procedures.

Observation 1 - Word lists. Prior to recording the data, participants were shown the list of words and were asked to read the words silently to themselves and ask any questions they had about meaning, pronunciation, structure, or usage. If a participant had been unfamiliar with one of the target words, which did not occur, the meaning would have been explained to her. After confirmation of familiarity with all the target words, participants were told to read the word list at a natural comfortable pace.

Observation 2 – Read speech. Participants were shown two paragraphs containing the target words. They were told to read the paragraphs to themselves and ask any questions they had about meaning, pronunciation, structure, or usage of both target and non-target words prior to recording the passage. A few of the non-target words used in the paragraphs were low frequency that some participants were unfamiliar with (e.g. thrift, kilts, fleece). The meaning, pronunciation, and structures of those words were explained to the participants. The target words in the paragraphs were identical to the word lists and participants did not report any concerns with the target words. After confirming comprehension, participants were asked to read the paragraphs at a natural, comfortable pace while their voice was recorded.

Observation 3 – Open-ended questions. During the open-ended question stimulus, participants answered questions posed by the interviewer. Question types and subject matter were randomized by the interviewers and no pattern of questioning was evident. Each participant was asked an average of ten questions with topics ranging from hometown stories to favorite foods to stressful situations in the USA. The average duration for each participant to respond to all the open-ended questions was nine minutes.

Observation 4 – Fairy tale recall. During the fairy tale retelling task, participants were given a choice on how they could respond. They could choose one fairy tale to retell from the list of four on the stimulus card or they could choose a different story to retell with which they felt more comfortable. All participants confirmed familiarity with all fairy tale options. Four participants chose to retell *Cinderella* and one chose to retell *The One Inch Boy*. Participants were given an average of one minute to prepare their response and then asked to retell the story in their own words within approximately five minutes.

SOUND TREATMENT AND ANALYSIS

To capture empirical data relating to the F3 formants for participants producing initial /l/ and /r/, and final /l/ and /r/ sounds, the elicited sound files were processed in two macro phases: (1) extraction of F3 formant frequency values and (2) data preparation for statistical analysis.

Extraction of F3 Formant Frequency Values

The data collection sessions produced five audio wave files for each participant with an average length of 19 minutes and 40 seconds (range of 14.5-22 minutes). Each audio file was then marked and segmented using the audio editing software Amadeus Pro V2.3.1 running on an Apple iMac. Files were segmented by the major sections that had been predetermined by the data elicitation instruments, namely the Word Lists 1-3, Paragraphs 1 and 2, Fairy Tale Recall, and Open-ended Questions. After segmenting the individual sections, each target word was identified and saved as separate, independent audio files totaling 108 /l/ token and 89 /r/ token (197 total tokens).

The TextGrid function of the acoustic analysis software Praat 6.0.19 was then used to create tiers which allowed for the isolation of onset and coda /l/ and /r/ phonemes. After the isolation of each phoneme, the "Show formants" algorithm in Praat was then used to sample F1-F4 frequency values approximately every six milliseconds which resulted in four distinct formant tracks for each of the 197 /l/ and /r/ tokens. Isolated onset or coda /l/ and /r/ were then saved as *.*collection* files with metadata regarding phoneme, number of formants, and maximum frequency settings. The "Formant listing" function of Praat was then used to export a table containing every formant frequency value (F1-F4) to a UFT-8 text file resulting in 33,864 individual formant measurements across all 197 tokens (/r/=15708 values; /l/=18156 values). The resulting raw formant measurements were comprised of three files: 1) the raw audio file, 2) the *.*collection* file containing tier separations and formant values, and 3) the exported formant frequency value tables in UTF-8 text files.

Data Preparation

The next stage in processing the acoustic data consisted of extracting the formant frequency values from the formant listing text files and transferring them to a spreadsheet software program. To do this, a custom-designed software script was developed using the MatLab technical computing language. MatLab is a proprietary computing language designed for numerical calculations (Attaway, 2016). This script was then used to extract the formant frequency values from the text file tables, sort those values into categories, run calculations on those values (e.g. slopes, means, maximums, minimums), and correctly place those values and calculations into an Excel spreadsheet similar to the method used by Fox & McGory (2007).

The audio files were carefully renamed to harmonize with the requirements of the MatLab software. The renamed files identified the randomized and anonymized participant number, word position, phoneme, task type, and token number. Figure 4 contains example file names for Participant 1 where "subj_num" refers to the participant number, "word_pos" refers to the position in each word in which the phoneme occurred (i=initial or f=final), "sound_type" refers to which phoneme was intended in production (/l/ or /r/), "task_type" refers to the task context where the token was elicited (ft=fairy tale recall, p1=paragraph reading 1, p2=paragraph reading 2, w1=citation form list 1, and oq=open-ended questions), and "tok_num" indicates the chronological order of each token produced within the specified contexts.

	A	B	C	D	E	F
1	in_file	subj_num	word_pos	sound_type	task_type	tok_num
2	1_f_l_ft_01.txt	1	f	l	ft	1
3	1_f_l_p1_01.txt	1	f	l	p1	1
4	1_f_l_p2_02.txt	1	f	l	p2	2
5	1_f_l_p2_03.txt	1	f	l	p2	3
6	1_f_l_w1_01.txt	1	f	l	w1	1
7	1_i_l_oq_01.txt	1	i	l	oq	1
8	1_i_l_oq_02.txt	1	i	l	oq	2
9	1_i_l_oq_03.txt	1	i	l	oq	3

Figure 4. Example column names foreach data point obtained from the MatLab software.

Sound File Segmentation. To control for variation in the length of the various audio files, each sound file was segmented into sequential, cascading "windows" based on a fixed percent of the total sound file duration. Segmentation of each token was performed in four sets of differing percentage values. The first set utilized the smallest duration window, 12.5% which created a total of eight windows. The second set utilized window durations of 25%, which created a total of four windows. The third set utilized window durations of 50%, which created a total of two windows. The fourth set utilized a single window duration of 100% which resulted in a single window. Figure 5 graphically represents these relationships.

Figure 5. Grid showing the four duration sets used to segment each sound file.

The cascading nature of these windows provided for both static and dynamic values to be generated and calculated across the duration of each token. The resulting spreadsheet contained anonymized mean data regarding the participant, task type, phoneme type, word position, token number, and duration of each sound file. The bulk of the spreadsheet consisted of columns containing the minimum, maximum, mean, slope, extent, and duration of each formant within each window.

Coarticulation. A secondary function of using windows was to allow for non-subjective and controlled adjustment of the F3 values due to the coarticulatory nature of the approximants /l/ and /r/. Previous research encountered technological limitations (Flege et al., 1995, 1996) which forced researchers to estimate the F3 formant values. One relevant aspect of these estimates was that the researchers visually determined the initial and final stages of coarticulation through manual measurement of the peak and trough of the formant track (Saito & Lyster, 2012). The use of percent-based windows in this study allowed for adjustments to be made in a more precise and objective manner. Figure 6 is a visualization of the averages for the /r/ phoneme in onset and coda positions segmented across the eight window durations of 12.5%

before controlling for coarticulation. Here the stable states seen in windows 1, 2, 3, 4, and 5 appear to be appropriate measures of /r/ while the rapid increase in frequency seen in windows 6, 7, and 8 seem to indicate coarticulation with the following vowel.

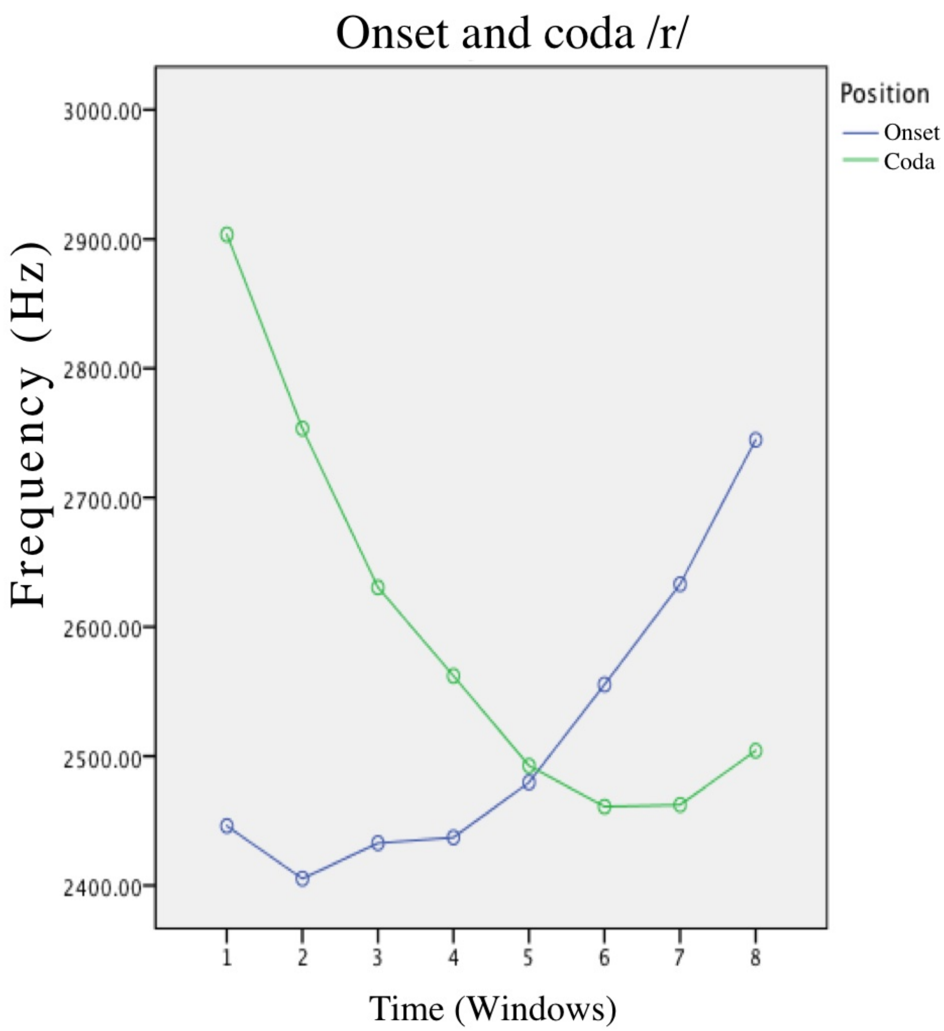

Figure 6. Line graph demonstraing coarticulation for /r/ in onset and coda across eight time segments.

These initial stages of analysis provided data which indicated that approximately 25% of each phoneme articulation could be identified as coarticulation. Thus, a percent-based strategy was formed to minimize the effects coarticulation in the data. For word-initial position, vowel coarticulation occurs from the onset to the nucleus. In contrast, coarticulation in word-final position occurs from the syllable nucleus to the syllable coda (Hardcastle, & Hewlett, 2006). To minimize the co-articulation effects on the /l/ or /r/ phonemes for word initial tokens, the initial 75% of each duration was used and the final 75% of each duration was used for word-final tokens. In other words, the final two data points of word initial tokens results were removed along with the initial two data points for word final data points. This resulted in less variable averages of /l/ and /r/ tokens which were then averaged. Figure 7 represents post-coarticulation adjustments.

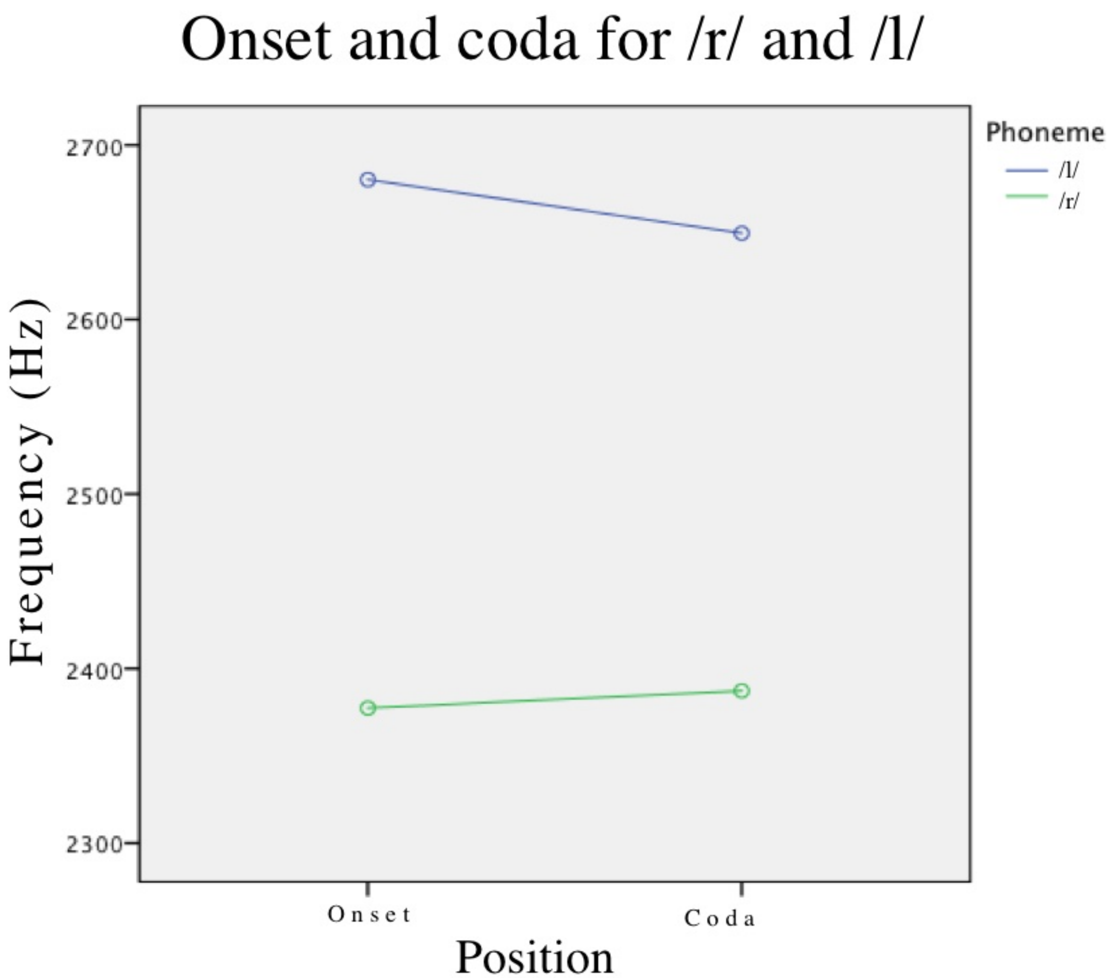

Figure 7. Line graph representing /l/ and /r/ values post-coarticulation removal.

Challenges in Data Preparation. In creating the data files for analysis from the audio samples, it was discovered that the formant algorithm used in Praat was unable to track formant values accurately 100% of the time. While causes for the inaccuracy are unclear, approximately 9% of the audio samples had missing F3 frequency values (marked as "undefined") or the algorithm inconsistently detected the F3, producing improbably large changes in raw frequency data over short periods of time (Baken, 1987).

Undefined formant values. The "undefined" formant values occurred when the Praat formant algorithm was unable to identify frequency values for a formant track. Occasionally, the error was visible in the spectrogram in the form of an unexplained ending or beginning of a formant track. The three arrows in Figure 8 indicate where the algorithm was unable to accurately track the formants. From the left, the first arrow indicates a new formant being detected by the algorithm even though no formant is visible. This new F2 shifts the formants and causes the formant tracks previously identified as F3 and F4 to become F4 and F5. The second arrow shows an unexplained jump in the formant values for F4 which rapidly becomes F5. The third arrow shows an abrupt ending of the F5 in the middle of the utterance. This most often occurred with the F3 and F4 formants, as shown by the formant listing in Figure 9 which is the formant listing for Figure 8 as exported by Praat and shows the frequency values of the spectrogram in Figure 8. The algorithms inability to identify the F3 and F4 to be labeled as "undefined" during the time framed identified by the red box.

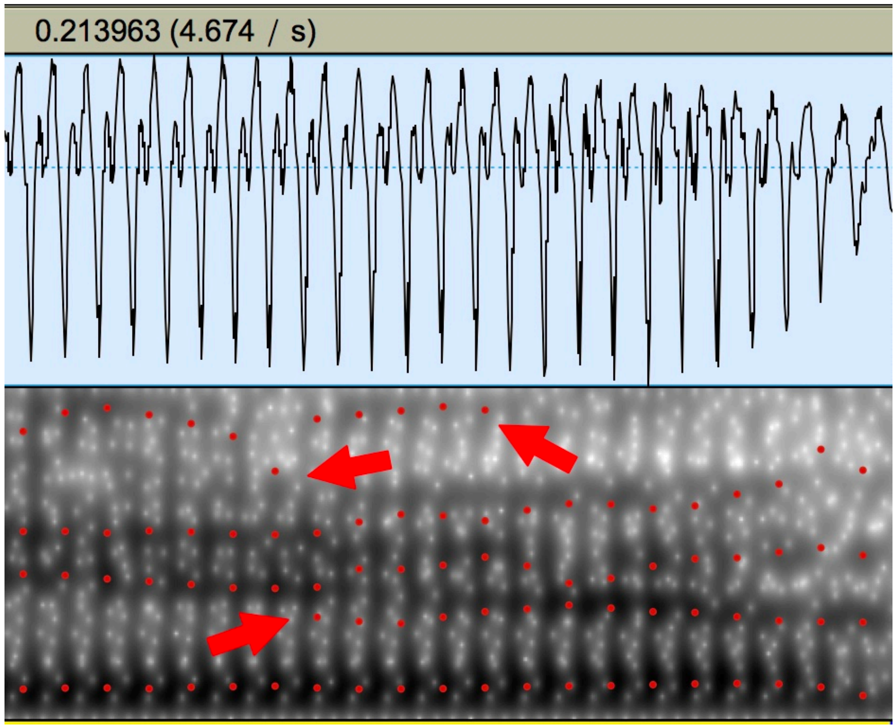

Figure 8. Sonogram of formant algorithm errors.

0.257350	451.969962	2269.681242	2853.377901	4532.936099
0.263600	453.026492	2164.383959	2769.495580	4660.609556
0.269850	452.507072	2065.034854	2732.358149	4699.007064
0.276100	457.551038	2017.983105	2792.941046	4586.563181
0.282350	465.882617	1998.685388	2794.237332	4529.083163
0.288600	472.330957	1952.179930	2774.338379	4388.055602
0.294850	481.199694	1987.770146	2841.919463	5093.258440
0.301100	484.083056	1926.851143	2728.601517	4432.964816
0.307350	492.939209	1918.792243	2649.293190	4423.765935
0.313600	504.844788	1953.995090	2740.293577	4435.835573
0.319850	509.674412	1905.574758	2750.190117	4501.664531
0.326100	514.876145	1841.442913	2695.171746	4354.991717
0.332350	537.012274	1847.967449	2894.868364	4772.933121
0.338600	550.423336	1799.057537	--undefined--	--undefined--
0.344850	560.036836	1792.640300	--undefined--	--undefined--
0.351100	571.496255	1774.486656	--undefined--	--undefined--
0.357350	574.214418	1732.644610	--undefined--	--undefined--
0.363600	587.580901	1720.151097	--undefined--	--undefined--
0.369850	565.276767	1649.235724	--undefined--	--undefined--
0.376100	483.609697	1469.211519	2636.920331	5263.442515
0.382350	345.643489	1519.854084	2700.517307	4934.451087

Figure 9. Formant listing identifiying “undefined” errors.

Formant detection. Approximately 5% of the formant values displayed formant detection errors which needed to be manually corrected. These errors seemed to occur most frequently when the Praat algorithm was unable to accurately distinguish between F2 and F3 thus producing dramatic increases or decreases in the F2 or F3 frequency of, on average, 1000 Hz within 12 milliseconds.

An example of this can be seen in Figure 11 which shows a spectrogram with no apparent variation in the F3. However, when compared to the formant listing exported for the same spectrogram, there is a significant dip in the F3 value. Previous research has found that formants produced by the human vocal tract follow progressive increases and decreases in frequency values over time and that, while those changes can be rapid, dramatic spikes or dips such as these are not possible (Baken, 1987).

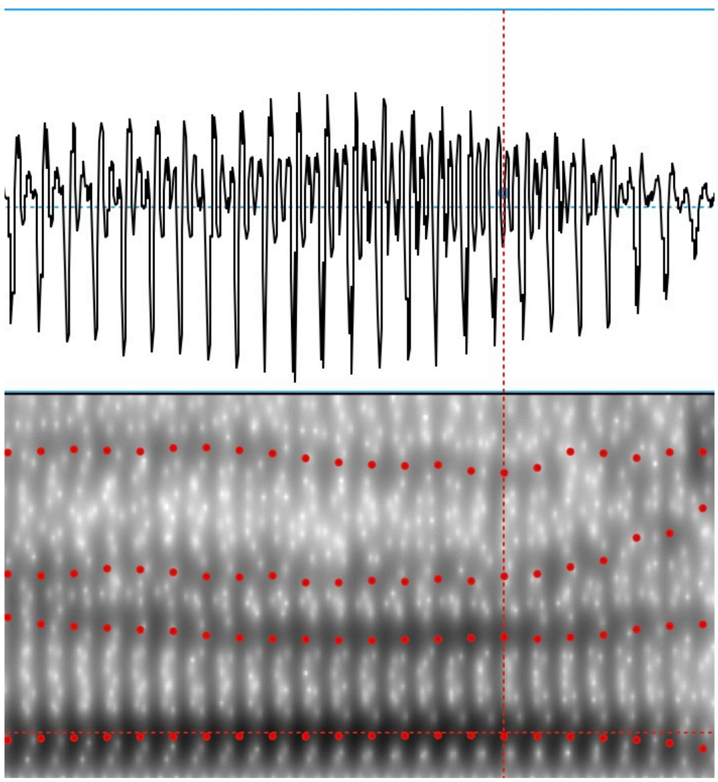

Figure 10. Sonogram with no visible formant errors.

Time_s	F1_Hz	F2_Hz	F3_Hz	F4_Hz
0.346215	502.600403	2260.028548	2647.035695	4299.091917
0.352465	518.351501	2096.805636	2637.193327	4313.307010
0.358715	527.221144	2051.401505	2710.617077	4335.547305
0.364965	543.869282	2012.117419	2839.617568	4314.997123
0.371215	555.236618	1964.583598	2804.114505	4281.092839
0.377465	554.782458	1941.107322	2728.226468	4310.431351
0.383715	553.108804	1891.936067	2661.678676	4327.139998
0.389965	556.994855	1861.195220	2663.103877	4303.158660
0.396215	560.566400	1837.911874	2692.648190	4258.338247
0.402465	555.535941	1837.610562	2633.938661	4223.272481
0.408715	557.807757	1824.389276	2621.536262	4194.244686
0.414965	561.191548	1817.681995	2637.041828	4149.999551
0.421215	561.280689	1828.999290	2645.179974	4118.253093
0.427465	568.787313	1830.885279	2670.798528	4134.184664
0.433715	562.110895	1854.105876	2700.756878	4093.384622
0.439965	550.783058	1874.197507	2162.735346	4005.308244
0.446215	545.848140	1838.835722	2730.405268	4206.355860
0.452465	542.542247	1859.162229	2809.405267	4490.201812
0.458715	525.905785	1889.852623	2929.620346	4409.951338
0.464965	503.882305	1970.169122	3322.964494	4372.160595
0.471215	452.366270	1974.886578	3260.255594	4314.008367
0.477465	402.417135	1981.240486	3196.713140	4189.178211

Figure 11. Formant listing with "impossible" dip in formant values shown in blue.

Resolution. Resolving "undefined" and impossible formant values proved a significant obstacle as the defining acoustic characteristic between /r/ and /l/ is /r/'s significantly lower F3. Recommendations from the Praat help files included instructions to modify the maximum frequency settings or the number of formants (Boersma, & Weenink, 2013). These settings were incrementally adjusted until the sampled formant track generated by the Praat algorithm accurately followed the visible formant path in the spectrogram. When the use of this technique was unsuccessful, the formant values immediately preceding and following the "undefined" values were incrementally averaged and added to the formant tables in place of the "undefined" labels.

RESULTS

The purpose of this study was to provide empirical data that explores the ability of adult JLEs to produce measurable differences in the /l, r/ contrast through detailed acoustical analysis. The current study examined the relationships of the /l, r/ contrast produced by five adult female JLEs. Each /l/ and /r/ production was processed acoustically resulting in comparisons of the F3 formant tract.

As shown in Table 1, there were a total of 197 tokens of /l/ and /r/ phoneme production from the five Japanese participants. Each task type produced an average of 66 tokens with the number of /l/ tokens being slightly more than the number of /r/ tokens. There was also a slight majority of the overall number of tokens occurring in the initial position.

Table 1

Between Subject Factors

Between-Subjects Factors					
	/l/		/r/		
Task	Initial	Final	Initial	Final	Total
Citation	31	14	14	16	75
Read	20	16	18	14	68
Free	14	16	11	13	54
Total	65	46	43	43	197

In order to fully analyze all relevant data and discuss each variable and interaction, a null hypothesis and two correlates were generated.

1. H_0 – There will be no significant difference between the F3 frequency values for female JLE when producing the /l/ and /r/ phonemes.
2. C_1 – Word position (initial or final) will have no significant effect on the F3 frequency values for /l/ or /r/.

3. C_2 – Task type (citation form, read speech, or free speech) will have no significant effect on the F3 frequency values for /l/ and /r/.

To determine whether the null hypothesis and its corresponding correlates could be rejected, the F3 formant values extracted through the previously described process were analyzed using a repeated multivariate analyses of variance (MANOVA). The MANOVA was used to determine variability of the F3 formant values between each of the three variables of task, phoneme, and position and the resulting four interactions. The descriptive statistics for the data collected – broken down by task type, position, and phoneme – can be found in Appendix E. The results of the MANOVA are in Table 2.

Table 2

Results of Between Variable Tests

Source	***F***	***df***	***p***	η_p^2
Task (T)	.218	1,189	.641	.001
Phoneme (PH)	71.575	1,189	.000	.275
Position (P)	.005	1,189	.944	.000
T * PH	1.485	1,189	.225	.008
T * P	.046	1,189	.831	.000
PH * P	.097	1,189	.756	.001
T * PH * P	.317	1,189	.574	.002

The MANOVA results showed that among the three variables occurring in this analysis (phoneme, task, position), only phoneme demonstrated significance at the .05 level and accounted for approximately 27% of the variance in F3 frequency values $F(1, 189) = 71.575$, $p < .001$, $\eta_p^2 = .275$. No other independent variable demonstrated significant results and no interaction among variables demonstrated significant results.

DISCUSSION

The purpose of this study was to add to the knowledge base regarding the /l, r/ contrast as produced by adult Japanese learners of English (JLEs) and provide empirical evidence either supporting or refuting the claim that JLEs are unable to accurately produce the /l, r/ contrast. While this contrast has been extensively researched using phonetic and perceptual methods, very little research seems to have been done on it acoustically. This study consisted of identifying and analyzing the F3 values of /l/ and /r/ as produced by five female JLEs. Each token of /l/ or /r/ was categorized according to the task type used to elicit each token (controlled speech vs. free speech) and the word position of each token (initial or final).

This discussion section will synthesize the results in relationship to the null hypothesis and its correlates, which are restated below:

1. H_0 – There will be no significant difference between the F3 frequency values for female JLE when producing the /l/ and /r/ phonemes.
2. C_1 – Word position (initial or final) will have no significant effect on the F3 frequency values for /l/ or /r/.
3. C_2 – Task type (citation form, read speech, or free speech) will have no significant effect on the F3 frequency values for /l/ and /r/.

In order to discuss word position and task type appropriately, corollary 1 (C_1) and corollary 2 (C_2) will be resolved before discussing the null hypothesis (H_0).

Resolving C_1 and C_2

Previous research has indicated that task type could affect the accuracy of /l/ and /r/ productions and that more complex tasks could cause a decrease in accuracy (Flege, Takagi, Mann, 1995). Word position has also been indicated as having statistically significant effects on the F3 frequency values for /l/ and /r/ (Saito & Lyster 2012). The F3 values obtained for /l/ and /r/ in this study demonstrated that neither task type nor word position had a significant effect on the F3 values for the /l, r/ contrast. Hence, it would appear that neither task type nor word position affect JLE productions of /l/ and /r/ which was an unexpected result.

Resolving H_0

While previous research indicates that adult JLEs can be trained to improve their accuracy in producing a distinction between the /l/ and /r/ phonemes (Saito & Lyster 2012; Sheldon & Strange, 1982), the results of this study are surprising in that they provide empirical evidence that at least some adult female JLEs are consistently producing a statistically significant distinction between /l/ and /r/ across word position and task type *without* any known targeted /l/ and /r/ coursework. Furthermore, these productions are meaningfully different from each other as measured by $\eta_p^2 = .275$, indicating that approximately 27% of the variance in the F3 frequency values can be attributed to intentional attempts by the participants to produce distinct phoneme productions. These results provide empirical evidence supporting the rejection of the null hypothesis and also provide more nuanced data regarding the variance of phoneme productions in this highly-controlled context.

Native English Speakers

With the null hypothesis rejected and a demonstrated ability to differentiate between /l/ and /r/, the next logical step would be to compare the results of this study to the results found in previous research which establishes F3 frequency standards for native speakers of American English (NES). A frequently cited study examining these values is Dalston's (1975) study which compared /l/ and /r/ productions for native English speaking adults and children using acoustical analysis of the F3 formant. Dalston's delineation between the F3 values of male and female adult

English speakers is useful in this context as it identifies female NESs as averaging 472 Hz higher in raw frequency values than males when producing either /l/ or /r/.

Table 3 shows the average values for the /l/ and /r/ productions of the five adult female JLEs as compared to the female NES standards proposed by Dalston (1975). The raw averages provided by Dalston for a NES /l/ and a NES /r/ are in columns four and five.

Table 3

Average F3 Values (Hz) for adult JLEs vs. NES Standards

	JLE		NES	
Participant	/l/	/r/	/l/	/r/
1	3107.64	2790.49	2935	2078
2	2869.95	2232.84	2935	2078
3	2692.07	2546.40	2935	2078
4	2700.94	2456.65	2935	2078
5	2672.06	2384.09	2935	2078

Figure 12 shows a line graph of the data in Table 5 which delineates the differences between each participant by phoneme.

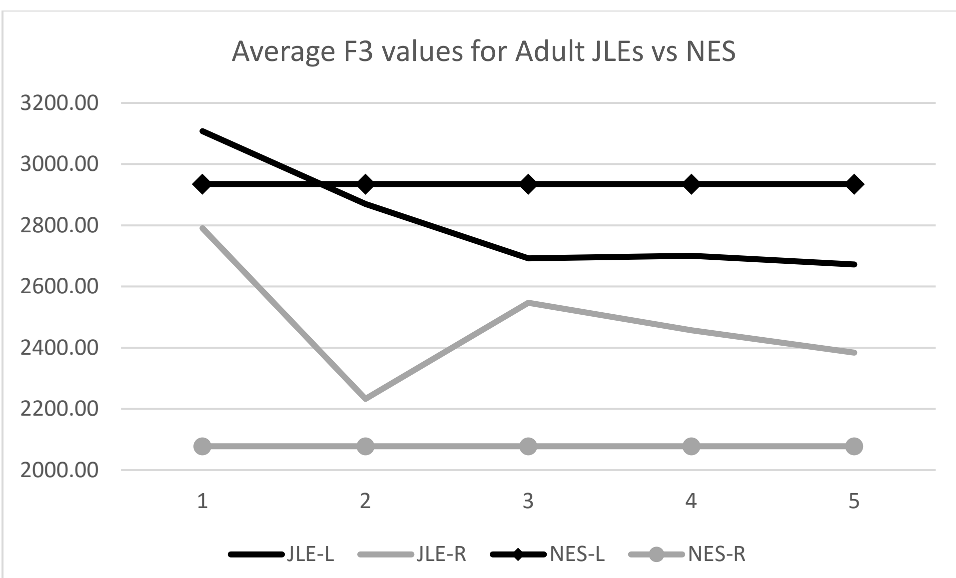

Figure 12. Average F3 values for adult JLEs vs. NES standards

The adult female JLEs who participated in this study, though able to produce differentiated tokens for the /l, r/ contrast, did not distinguish /l/ or /r/ to the same degree as did the NES females and it is interesting to note that the JLE productions are neither as high nor as low as NES standards. These values indicate that, though the participants are distinguishing the /l, r/ contrast, they are not distinguishing it to the same degree as NESs.

Table 4 shows that the effect size for the difference between NES /r/ and adult JLEs /r/ is meaningfully larger than for respective productions of /l/. This data provides empirical evidence supporting the claim that /r/ production is a greater challenge for adult JLEs than /l/ production. It also provides additional support for previous research which concluded that /l/ was easier for adult JLEs to accurately produce than /r/ (Aoyama, 2004; Flege et al., 1995) and it lends further support to specific training methods to teach production of the /r/ sound (Saito & Lyster, 2012) rather than focus on contrastive instruction of /l/ vs /r/.

Table 4

Effect Size of JLE Phonemes Compared to NES Standards

	JLE		NES		
Phoneme	*M*	*SD*	*M*	*SD*	*d*
/l/	2809.72	263.778	2935	174.4	-0.560
/r/	2491.81	226.413	2078	346.1	1.415

Implications

For the female adult Japanese JLEs who participated in this study, the major challenge did not seem to be an inability to produce distinct occurrences of the /l, r/ contrast, but instead to produce occurrences of /l/ or /r/ at NES standard F3 values which would make the sounds more easily distinguishable acoustically.

Curiously, the findings from this study indicating that task type did not produce discernable differences in the speakers' production of /l/ and /r/ contrast with previous findings which have found task type to affect language productions (Riney, Takada, Ota, 2000; Saito & Lyster, 2012). These results open up discussion regarding the effect of task type on language production as it seems that there could be certain aspects of language that may not be affected by task type. While it is yet unclear exactly which aspects could be unaffected, these results seem to indicate that the relationship between task type and language production may not be as uniform as previously believed. It may even be possible that an array of language variables is wholly unaffected by task type. If more unaffected aspects of language production were discovered, language assessment for this population might be affected.

The results presented in this study may also have certain implications for teaching, especially if task type does not seem to alter production, as shown in this study. Standard English as a second language teaching methodologies evaluate proficiency based on task types, also called functions. As task type appeared to have little effect on the speaker's ability to produce differentiated /l/ and /r/ tokens, it is probable that even highly proficient JLEs will continue to have difficulty with this contrast. Though the frequency of inaccurate productions tends to decrease as the speaker's proficiency increases, this may not be the case with JLEs.

Another possible implication of this research could be that a variety of task types might not be necessary for future research. If task type does not affect productions of the /l, r/ contrast as produced by this population, there may not be any need to include it as a variable in future research. This could allow for more focused research and a more detailed description of this contrast as produced by JLEs.

A beneficial implication for pedagogy could be that JLEs are able to distinguish between /l/ and /r/ however not to the same degree as NES. This would suggest that focused production and articulation training would be more beneficial than task-based, contextual, or fluency, training.

Limitations

There are several important limitations to consider when determining the generalizability of the results found in this study. Despite the relatively large number of total tokens collected and analyzed in this study (197), gathering large numbers of unique tokens for each context proved difficult. This phenomenon was most noticeable in the free speech task types where, even though prompts were given that would elicit the target sounds, an average of only five tokens per participant were collected during the seven minutes they produced spontaneous speech.

An additional limitation in this study was the small number of participants. While additional subjects would have provided a more robust data set from which to run the MANOVA, future researchers need to carefully weigh the additional time commitment required to prepare an increased number of tokens for acoustical analysis.

Directions for Further Research

This section will address the need for further diagnostic and descriptive research to identify other major causes of variability regarding adult JLE productions of the /l, r/ contrast. This study was designed to provide empirical evidence regarding JLE productions of /l/ and /r/ which, when gathered in a systematic way, would provide a more complete diagnosis of the challenges of adult JLEs. One avenue of future research could be to analyze the /l, r/ contrast in different vowel contexts or in clusters. Careful control of additional variables must be maintained in order to systematically identify all factors that contribute to the variance detected in JLE /l/

and /r/ productions. As the underlying reason for JLE difficulty with this phonemic contrast remains unknown, it is likely that only systematic and controlled research conditions will further develop our understanding of this topic.

One interesting phenomenon identified in this study relates to the fact that in approximately 7% of the word-initial /l/ productions, participants produced a flapped /r/ and NOT the lateral approximant /l/ which caused perceptual confusion for the researchers (e.g. /leap/ was perceived as /deep/). This observation is interesting because the F3 formant remained appropriately high (~2832 Hz) for a /l/ production even though the production was a flapped /r/ and not an /l/. This result could indicate that the F3 formant alone may not be a sufficiently robust measure for the diagnosis of the /l, r/ contrast for adult JLEs (Underbakke et al. 1988; Gibbon, 1991). This observation is supported by previous research (Miyawaki, 1975) which found that the Japanese /r/ was often perceived by NESs as /d/. One possible explanation could be that adult JLEs are producing the initial /l/ sound with the blade of the tongue rather than the tip against the alveolar ridge. This production could cause the F3 formant values to remain relatively high even though the manner of articulation is similar to /r/.

Another possibility is that the Japanese /r/ phoneme could share phonological space (Moulton, 1962) with more than just the /l/ and /r/ phonemes in English and may overlap with the phonemes /d/ (e.g. /deep/) and the /ɾ/ flap (e.g. /better/). If this were the case, then instruction focused only on the production of the /l, r/ contrast may not fully meet the needs of adult JLEs.

Another future application could be to replicate this study with adult male JLEs. Previous research has indicated that adult male JLEs may have a more variable pronunciation of the /l, r/ contrast than females (Gibbon, 1991). However, neither the degree of variability, nor the extent of this variation is known. Focused research in this area may reveal patterns previously undetected (Hisagi et al., 1998) which could assist in the development of training methods for the adult JLEs.

Given the small number of participants in this study, future research could also attempt a replication of this study with larger numbers of adult female JLEs to build a more robust data set and possibly uncover more nuances within the data. It is hoped that the detailed explanation and procedure outlined in this study could make replication plausible.

Finally, the results from this study also seem to indicate that, with regards to production of the target phonemes, a further benefit could be gained through simultaneously using multiple technologies in the description and diagnosis of Japanese productions of /l/ and /r/. One such tool could be electropalatography (EPG) which allows researchers to empirically capture /l/ and /r/ tongue articulations in combination with acoustic data. EPG has been shown to significantly enhance the level of detail for description and diagnosis (Alwan, Narayanan, & Haker, 1997) and shows potential as a tool for second language instruction.

CONCLUSION

The purpose of this study was to construct a detailed acoustic analysis for the F3 formant frequencies for /l/ and /r/ as produced by adult JLEs and, with it, provide empirical evidence either supporting or refuting the null hypothesis that there will be no significant difference between the F3 frequency values for the /l/ and /r/ phonemes as produced by adult female JLEs. The accuracy of those productions was also discussed relative to the NES /l/ and /r/ standards presented by Dalston (1975).

In this study, five female Japanese intermediate-level learners of English produced speech elicited from controlled speech and free speech which contained the initial and final phonemes of /l/ and /r/. The data were collected during five independent interviews that elicited controlled and free speech samples in random order. Speech samples were processed using Praat, Amadeus Pro, MatLab, and SPSS to identify and analyze F3 frequency values for /l/ and /r/.

Based on the results of this study, at least some female adult JLEs can consistently produce a statistically significant difference between the English phonological contrast of /l/ and /r/, as measured through a lowered F3 formant for /r/ and a raised F3 formant for /l/. Within the vowel and word position contexts outlined in this paper, they are also able to produce a significant distinction between the /l/ and /r/ phonemes as early as the intermediate proficiency level without any known /l/ or /r/ focused instruction. In contrast to previously observed NES patterns (Flege 1996), increased complexity introduced through altered task type did not have a significant effect on production of these phonemes for these participants. This finding supports the conclusion that the observed difficulties of JLEs with the /l, r/ contrast is more complex than an inability to produce the F3. Future diagnostic research may provide further insights into the highly complex /l, r/ contrast as produced by Japanese learners of English.

REFERENCES

ACTFL. (2012). *Proficiency guidelines 2012*. Alexandria, VA: Author.

Alwan, a, Narayanan, S., & Haker, K. (1997). Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part II. The rhotics. *The Journal of the Acoustical Society of America*, *101*(2), 1078–1089. https://doi.org/10.1121/1.418030

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: the case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, *32*(2), 233–250. https://doi.org/10.1016/S0095-4470(03)00036-6

Asakawa, G. (2015). *Being Japanese American: A JA sourcebook for Nikkei, Hapa ... and their friends* (2nd ed.). United States: Stone Bridge Press. Retrieved from https://books.google.com/books?id=w-1HCgAAQBAJ

Attaway, S. (2016). *Matlab: A practical introduction to programming and problem solving*. Elsevier Science. Retrieved from https://books.google.com/books?id=2vQqCgAAQBAJ

Author, R., Review, D. B., Source, D. B., Society, L., & Url, A. S. (2015). Linguistic Society of America Review, *84*(3), 655–658. https://doi.org/10.1002/wcs.1163

Baken, R. J., & Orlikoff, R. F. (1987). *Clinical measurement of speech and voice*. Boston: College-Hill Press. Retrieved from https://books.google.com/books?id=ElPyvaJbDiwC

Bakić-Mirić, N. (2011). *An integrated approach to intercultural communication*. Newcastle upon Tyne: Cambridge Scholars Publisher. Retrieved from https://books.google.com/books?id=mC0sBwAAQBAJ

Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer [Computer program]. Version 6.0.19. Retrieved from http://www.praat.org/

Boyce, S., & Espy-Wilson, C. Y. (1997). Coarticulatory stability in American English /r/. *Acoustical Society of America*, *101*(6), 3741.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America*, *101*(4), 2299–310. https://doi.org/10.1121/1.418276

Brunner, J., Fuchs, S., & Perrier, P. (2009). On the relationship between palate shape and articulatory behavior. *The Journal of the Acoustical Society of America*, *125*(6), 3936–49. https://doi.org/10.1121/1.3125313

Burmeister, H. (1987). Applied psycholinguistics. *System*, *15*, 234–236. https://doi.org/10.1016/0346-251X(87)90080-7

Campbell, N. (2007). Differences in the Speaking Styles of a Japanese Male According to Interlocutor; Showing the Effects of Affect in Conversational Speech. *Computational Linguistics and Chinese Language Processing*, *12*(1), 1–16.

Celce-Murcia, M., Brinton, D. M., & Snow, M. A. (2014). *Teaching English as a second or foreign language*. National Geographic Learning. Retrieved from https://books.google.com/books?id=_kJ4MQEACAAJ

Chen, H. C. (2015). Acoustic Analyses and Intelligibility Assessments of Timing Patterns Among Chinese English Learners with Different Dialect Backgrounds. *Journal of Psycholinguistic Research*, (10), 749–773. https://doi.org/10.1007/s10936-014-9315-5

Dalston, R. M. (1975). Acoustic characteristics of English /w, r, l/ spoken correctly by young children and adults. *Journal of the Acoustical Society of America*, *57*(March 1974), 462–469.

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, *39*, 379–397. https://doi.org/10.2307/3588486

Ellis, R. (1994). *The study of second language acquisition*. Oxford University Press. Retrieved from https://books.google.com/books?id=3KglibyrZ5sC

Flege, J. E., Takagi, N., & Mann, V. (1996). Lexical familiarity and English-language experience affect Japanese adults' perception of /ɹ/ and /l/. *The Journal of the Acoustical Society of America*, *99*(2), 1161–1173. https://doi.org/10.1121/1.414884

Flege, J. E., Takagi, N., & Mann, V. (1995). Japanese adults can learn to produce English /r/ and /l/ accurately. *Language and Speech*, *38*, 25–55.

Gass, S. M., & Mackey, A. (2007). *Data elicitation for second and foreign language research*. Lawrence Erlbaum Associates.

Genzberger, C. (1994). *Japan business: The portable encyclopedia for doing business with Japan*. San Rafael: World Trade Press. Retrieved from https://books.google.com/books?id=Ofvka_i08CAC

Gibbon, F. E. (1991). An electropalatographic study of the /r/, /I/ distinction for Japanese learners of English. *Computer Assisted Language Learning, 4,* 153–171.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "L" and "R." *Neuropsychologia*, *9*(3), 317–323. https://doi.org/10.1016/0028-3932(71)90027-3

Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *The Journal of the Acoustical Society of America*, *107*(5), 2711. https://doi.org/10.1121/1.428657

Hahn, L. D., & Dickerson, W. B. (1999). *Speechcraft: Discourse pronunciation for advanced learners*. University of Michigan Press. Retrieved from https://books.google.com/books?id=iuZiItGhnVQC

Hardcastle, W. J., & Hewlett, N. (2006). *Coarticulation: Theory, data and techniques*. Cambridge University Press. Retrieved from https://books.google.com/books?id=SOrbRH5p5h4C

Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication*, *47*, 360–378. https://doi.org/10.1016/j.specom.2005.04.007

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*, 61–135. https://doi.org/10.1017/S0140525X0999152X

Hillenbrand, J. M. (2011). Acoustic Analysis of Voice: A Tutorial. *Perspectives on Speech Science and Orofacial Disorders*, *21*(2), 31–43. https://doi.org/10.1044/ssod21.2.31

Hisagi, M., Nishi, K., & Strange, W. (1998). Acoustic properties of Japanese and English vowels: Effects of phonetic and prosodic context. *Japanese and Korean Linguistics (Proceedings)*, *13*.

Houmanfar, R., Hayes, L. J., & Herbst, S. A. (2005). An Analog Study of First Language Dominance and Interference over Second Language. *The Analysis of Verbal Behavior*, *21*(1), 75–98. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22477315

Ide, S., Motoko, H., Kawasaki, A., Ikuta, S., & Haga, H. (1986). Sex Difference and Politeness. *International Journal of the Sociology of Language*, *58*, 25–36.

Ingvalson, E. M., Holt, L. L., & McClelland, J. L. (2012). Can native Japanese listeners learn to differentiate /r–l/ on the basis of F3 onset frequency? *Bilingualism: Language and Cognition*, *15*(2), 255–274. https://doi.org/10.1017/S1366728911000447

Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. *The Journal of the Acoustical Society of America*, *118*(5), 3267. https://doi.org/10.1121/1.2062307

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition* (Vol. 87). https://doi.org/10.1016/S0010-0277(02)00198-1

Kashihara, T. (2008). Globalisation and Linguistic Competencies: Responding to Diversity in Language Environments; In *English Education and Policy in Japan.* Seminar conducted at the meeting of the 12th OECD-Japan Seminar, Tokyo, Japan.

Ladefoged, P. (1996). *Elements of Acoustic Phonetics*. University of Chicago Press. Retrieved from https://books.google.com/books?id=5l6pDpjMpzwC

Ladefoged, P. (1999). Instrumental Techniques for Linguistic Phonetic Fieldwork. In W. J. Hardcastle, J. Laver, & (eds) (Eds.), *The handbook of phonetic sciences*. Blackwell Publishing Ltd.

Ladefoged, P., & Johnson, K. (2014). *A course in phonetics*. Cengage Learning. Retrieved from https://books.google.com/books?id=NOYbCgAAQBAJ

Lambacher, S. (1999). A CALL tool for improving second language acquisition of English consonants by Japanese learners. *Computer Assisted Language Learning*, *12*(2), 137–156. https://doi.org/10.1076/call.12.2.137.5722

Larson-Hall, J. (2006). What does more time buy you? Another look at the effects of long-term residence on production accuracy of English /r/ and /l/ by Japanese speakers. *Language and Speech*, *49*, 521–548. https://doi.org/10.1177/00238309060490040401

Lefkowitz, N., & Hedgcock, J. (2002). Sound barriers: influences of social prestige, peer pressure and teacher (dis)approval on FL oral performance. *Language Teaching Research*, *6*, 223–244. https://doi.org/10.1191/1362168802lr107oa

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the Acoustical Society of America*, *94*(3), 1242. https://doi.org/10.1121/1.408177

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English /r/ and /l/. III. Long-term retention of new phonetic categories. *The Journal of the Acoustical Society of America*, *96*(4), 2076–2087. https://doi.org/10.1121/1.410149

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English/r/and/l/: A first report. *The Journal of the Acoustical Society of America*, *89*(2), 874–886. https://doi.org/10.1121/1.1894649

Lust, B., Flynn, S., & Foley, C. (1998). What children know about what they say: Elicited imitation as a research method for assessing children’s syntax. *Methods for Assessing Children’s Syntax*, (1), 55–76.

McClelland, J. L., Fiez, J. A., & McCandliss, B. D. (2002). Teaching the /r/–/l/ discrimination to Japanese adults: behavioral and neural aspects. *Physiology & Behavior*, *77*(4), 657–662. https://doi.org/10.1016/S0031-9384(02)00916-2

Mennen, I., Scobbie, J. M., de Leeuw, E., Schaeffler, S., & Schaeffler, F. (2010). Measuring language-specific phonetic settings. *Second Language Research*, *26*, 13–41. https://doi.org/10.1177/0267658309337617

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics*, *18*(5), 331–340.

Moulton, W. (1962). Dialect Geography and the concept of phonological space. *Word*, *18*(March), 23–32. https://doi.org/10.1080/00437956.1962.11659763

Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System*, *34*, 520–531. https://doi.org/10.1016/j.system.2006.09.004

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *45*, 73–97. https://doi.org/10.1111/j.1467-1770.1995.tb00963.x

Nishi, K. (2002). *Perceptual characteristics and intelligibility of Japanese-accented English*. University of South Florida.

Nishi, K., & Kewley-Port, D. (2005). Training Japanese listeners to identify American English vowels. *The Journal of the Acoustical Society of America*, *117*(4), 2401–2401. https://doi.org/10.1121/1.4786064

Pisoni, D. B., Lively, S. E., Yamada, R. A., Tohkura, Y., & Yamada, T. (1993). Training Japanese listeners to identify English /r/ and /l/: A replication and extension. *The Journal of the Acoustical Society of America*, *93*(4), 2391. https://doi.org/10.1121/1.406051

Pruitt, J. S. (1993). Comments on "'Training Japanese listeners to identify English /r/ and /l/: A first report'" [J. S. Logan, S. E. Lively, and D. B. Pisoni, J. Acoust. Soc. Am. 89, 874–886 (1991)]. *The Journal of the Acoustical Society of America*, *94*(2), 1146. https://doi.org/10.1121/1.406962

Reetz, H., & Jongman, A. (2003). *Acoustic characteristics of speech sounds*. *15th International congress of Phonetic Sciences (ICPHS)*. https://doi.org/10.1007/s00424-011-1058-7

Riney, T. J., Takada, M., & Ota, M. (2000). Segmentals and global foreign accent: The Japanese flap in EFL. *TESOL Quarterly*, *34*, 711–737. https://doi.org/10.2307/3587782

Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɹ/ by Japanese learners of English. *Language Learning*, *62*, 595–633. https://doi.org/10.1111/j.1467-9922.2011.00639.x

Saito, K., & Shintani, N. (2016). Foreign accentedness revisited: Canadian and Singaporean raters' perception of Japanese-accented English. *Language Awareness*, *25*(4), 305–317. https://doi.org/10.1080/09658416.2016.1229784

Sheldon, A. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics*, *3*, 243–263. https://doi.org/10.1207/s15327752jpa8502

Stone, M., & Lundberg, a. (1996). Three-dimensional tongue surface shapes of English consonants and vowels. *The Journal of the Acoustical Society of America*, *99*(6), 3728–3737. https://doi.org/10.1121/1.414969

Strange, W., Bohn, O. S., Trent, S. A., & Nishi, K. (2004). Acoustic and perceptual similarity of North German and American English vowels. *Journal of the Acoustical Society of America*, *115*(4), 1791–1807. https://doi.org/10.1121/1.1687832

Suzuki, N., Wakumoto, M., Michi, K., Dent, H., Gibbon, F., & Hardcastle, W. J. (1995). Cross-linguistic study of lateral misarticulation using electropalatography. *European Journal of Disorders of Communication: The Journal of the College of Speech and Language Therapists, London*, *30*(2), 237–45. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7492854

Tabain, M. (1998). Coarticulation in CV syllables: A locus equation and EPG perspective. *The Journal of the Acoustical Society of America*, *103*, 2980. https://doi.org/10.1121/1.422431

Takagi, N. (2002). The limits of training Japanese listeners to identify English /r/ and /l/: Eight case studies. *The Journal of the Acoustical Society of America*, *111*(6), 2887–2896. https://doi.org/10.1121/1.1480418

Thomas, E. R. (2016). Sociophonetics of Consonantal Variation. *Annual Review of Linguistics*, *2*(1), 95–113. https://doi.org/10.1146/annurev-linguistics-011415-040534

Tsukada, K., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Flege, J. (2005). A developmental study of English vowel production and perception by native Korean adults and children. *Journal of Phonetics*, *33*, 263–290. https://doi.org/10.1016/j.wocn.2004.10.002

Underbakke, M., Polka, L., Gottfried, T. L., & Strange, W. (1988). Trading relations in the perception of /r/–/l/ by Japanese learners of English. *The Journal of the Acoustical Society of America*, *84*(1), 90–100. https://doi.org/10.1121/1.396878

Washio, Y., & Houmanfar, R. (2007). Role of Contextual Control in Second Language Performance. *The Analysis of Verbal Behavior*, *23*(1), 41.

Weirich, M., & Fuchs, S. (2013). Palatal Morphology Can Influence Speaker-Specific Realizations of Phonemic Contrasts. *Journal of Speech, Language & Hearing Research*, *56*(6S), S1894-908. https://doi.org/10.1044/1092-4388(2013/12-0217)

Wilson, I. L., Fujinuma, J., Horiguchi, N., & Yamauchi, K. (2009). Acoustic analysis of the English pronunciation of Japanese high school teachers and university students. *Journal of the Acoustical Society of America*, *126*(4), 2312.

APPENDIX A – CITATION FORM WORD LISTS

Word List #1

Say beef again.

Say kilts again.

Say clean again.

Say teeth again.

Say thrift again.

Say think again.

Say cream again.

Say preach again.

Say leash again.

Say seat again.

Say freeze again.

Say beach again.

Say dear again.

Say feet again.

Say dirt again.

Say tilt again.

Say spring again.

Say please again.

Say leap again.

Say screen again.

Say cheek again.

Say fleece again.

Say split again.

Say shirts again.

Say street again.

Say treat again.

Say read again.

Say steep again.

Say sheep again.

Say piece again.

Say speak again.

Say deep again.

Say skin again.

Say peel again.

Say peek again.

Say sleep again.

Word List #2

Say shirts again.

Say cream again.

Say dear again.

Say spring again.

Say street again.

Say clean again.

Say seat again.

Say read again.

Say think again.

Say deep again.

Say steep again.

Say dirt again.

Say leash again.

Say sheep again.

Say beef again.

Say treat again.

Say tilt again.

Say fleece again.

Say kilts again.

Say split again.

Say screen again.

Say thrift again.

Say please again.

Say freeze again.

Say preach again.

Say speak again.

Say beach again.

Say piece again.

Say sleep again.

Say peek again.

Say leap again.

Say teeth again.

Say peel again.

Say skin again.

Say feet again.

Say cheek again.

Word List #3

Say clean again.

Say dirt again.

Say cream again.

Say deep again.

Say beach again.

Say sleep again.

Say leap again.

Say teeth again.

Say beef again.

Say feet again.

Say thrift again.

Say speak again.

Say skin again.

Say please again.

Say spring again.

Say dear again.

Say preach again.

Say cheek again.

Say seat again.

Say treat again.

Say read again.

Say peek again.

Say split again.

Say freeze again.

Say kilts again.

Say screen again.

Say steep again.

Say piece again.

Say think again.

Say sheep again.

Say tilt again.

Say leash again.

Say shirts again.

Say street again.

Say peel again.

Say fleece again.

APPENDIX B – PARAGRAPH READINGS

Paragraph Reading 1

Last spring Tom and Kate went to the beach with their dog. After removing his leash, they watched him climb a steep mound of dirt and leap in the air. Tom split a piece of beef and threw it for the dog to catch in his teeth. They found a good seat in the sand and began to read a letter from a dear friend. After taking a peek at one page, Tom's head began to tilt back and he began to sleep. Kate started to think, I wonder if Tom covered his skin to screen out the sun. She decided to speak, "Please wake up". It was too late. Tom had deep burns on his cheek, the skin had started to peel. They decided to clean the sand from their feet and walk to the thrift shop at the top of the street. The shop not only sold cream to treat his burns, but also had kilts and shirts made of fleece. Tom began to preach, "Those poor sheep, I hope they don't freeze."

Paragraph Reading 2

One spring evening, mother made a fruitcake to sell at the thrift shop on our street. She split a peach in half and began to peel away the skin. She smiled with her teeth as she began to read the instructions. Next, she told me to please clean the fruit to remove any dirt. With the bowl at a tilt she added cream to the treat. My dear father rose to his feet and kissed her on the cheek. She then gave him a leftover piece of beef to eat. He said, “I need to feed those sheep before I go to sleep.” He quickly rose from his seat, took a leap across the room, and rushed through the screen door. Mother said “I think we should put him on a leash”. With the cake in the oven, she began to speak and preach about her memories swimming in the deep ocean far from the beach and climbing a steep mountain peak to watch men freeze while dancing in shirts and kilts made of fleece.

APPENDIX C – FAIRY TALE RECALL

Please retell one of the following stories in your own words. Try to include as much detail as possible.

一寸法師	The One-inch Boy
金太郎	The Golden Boy
桃太郎	The Peach Boy
シンデレラ	Cinderella

APPENDIX D – OPEN-ENDED QUESTIONS

- Describe for me what happened on the day that you arrived in the United States.
-
    - Tell me more about what happened when...
- Share with me why learning English is so important to you.
    - Tell me more about how English will create opportunities for you.
    -
- Tell me about a favorite memory when you were young.
- Tell me about your hometown.
- Tell me about a recent vacation.
- Tell me how living in the United States is different than your home country.
- Tell me a story about a time you had fun with your family or friends?
- Tell me what you enjoy doing in your free time. Why?
- Tell me about your education before coming to the ELC.
- Tell me about your favorite subject in school and why.

APPENDIX E – DESCRIPTIVE STATISTICS

Descriptive Statistics

Task			Mean	SD
Controlled	/l/	Initial	2782.80	260.595
		Final	2804.56	265.478
		Total	2790.86	260.967
	/r/	Initial	2499.32	216.136
		Final	2500.64	256.586
		Total	2499.96	234.606
	Total	Initial	2673.51	279.889
		Final	2652.60	300.806
		Total	2664.73	287.993
Free	/l/	Initial	2882.67	208.129
		Final	2841.42	318.624
		Total	2860.67	269.012
	/r/	Initial	2454.70	184.304
		Final	2484.33	230.979
		Total	2470.75	206.958
	Total	Initial	2694.36	290.908
		Final	2681.34	331.533
		Total	2687.37	310.536
Total	/l/	Initial	2804.31	252.122
		Final	2817.38	282.090
		Total	2809.72	263.778
	/r/	Initial	2487.91	207.258
		Final	2495.71	246.494
		Total	2491.81	226.413
	Total	Initial	2678.33	281.238
		Final	2661.97	309.551
		Total	2670.94	293.714