March 2020

Acoustic Characteristics of Word-Final American English Liquids Produced by L2 Adult Speakers

Judith A. Espinal
Louisiana State University at Baton Rouge

ACOUSTIC CHARACTERISTICS OF WORD-FINAL AMERICAN ENGLISH LIQUIDS PRODUCED BY L2 ADULT SPEAKERS

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Arts

in

The Department of Communication Sciences and Disorders

by
Judith Alexis Espinal
B.A., Bowdoin College, 2013
May 2020

ACKNOWLEDGEMENTS

I wish to acknowledge and thank my thesis committee, Dr. Yunjung Kim, Dr. Hyunju Chung, and Dr. Todd Gibson for their guidance and support. I would especially like to thank my thesis advisor, Dr. Kim, for her continued dedication, direction, and patience throughout this project. In addition, I would like to thank lab members' support, especially Austin Thompson and Ashlyn Lee, for all of their support and assistance throughout this project. I would also like to thank all of the speakers who gave their time in order to participate in this study. Finally, I would like to thank my family; their love and support has meant the world to me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

In this study, the acoustic differences between native English speakers' (L1) and native-Korean speakers' (L2) production of American English liquids /ɹ/, /l/ and /ɹl/ were examined among 14 Korean speakers and 13 English speakers. Temporal measures included (1) relative timing of maximum constriction and (2) duration of vocalic nuclei. Spectral measures included (1) Euclidean distance between /ɹ/ and /l/ and (2) frequency difference between F2 and F3. The results indicated a significant interaction between speaker group and phonetic stimuli. That is, L2 speakers produced a similar degree of constriction across semivowels, whereas L1 speakers produced varying degrees of F2-F3 constrictions across phonetic stimuli. In addition, the relative timing of maximum constriction occurred earliest in /ɹl/ and latest in /ɹ/ production for L1 speakers. The opposite pattern was observed for L2 speakers. Furthermore, the two speaker groups exhibited significantly different results concerning the Euclidean distances between /ɹ/ and /l/. The Euclidean distances between the two sounds were significantly closer for L2 speakers compared to L1 speakers which indicates reduced acoustic distinction between the two liquids in L2 speakers. The same results were revealed for both measurement points, temporal midpoint and maximum constriction point. However, the speaker group difference was more apparent when measured from the point of maximum F2-F3 constriction compared to the temporal midpoint. The findings provide acoustic data on liquid production in L2 speakers and support the use of these measures in a clinical setting.

CHAPTER 1. INTRODUCTION

With the increasing number of native speakers of a language other than English living in the United States (US), the exposure to accented speech is growing; so is the need for the development of effective, data-driven accent management or accent neutralization approaches. In this study, we examined the acoustic characteristics of liquid sounds of American English produced by L2 adult speakers with a Korean language background.

We chose native Korean speakers, because of a relatively large, Korean-speaking population in the United States (approximately 1.1 million according to the US Census Bureau, 2015). In addition, numerous studies have reported that L2 speakers of English have difficulties learning /ɹ/ (as in car) and /l/ (as in call), especially in making clear differences between the two sounds (Borden et al., 1983; Han, 2002). For example, this population of L2 speakers tend to produce an [r]-like sound for /l/. Even after achieving a good degree of fluency in English, production of the two sounds still remains less distinctive and is perceived as confusing by native speakers of English. This challenge has been primarily attributed to the lack of the sounds in Korean and different occurrence of these sounds between the two languages (Iverson & Sohn, 1994). For example, there is no retroflexed [r] in Korean; only an apical flap occurring in intervocalic positions. In addition, although an apical /l/ exists in Korean, it occurs only in postvocalic positions. In addition to the two singletons, /ɹ/ and /l/, the current study included the combination /ɹl/ (as in Carl), in recognition that this complex sequence may increase articulatory difficulties for L2 speakers based on the observation that sounds that require more refined articulatory skills and a complex pattern of inter-articulator coordination are difficult to acquire (Kent, 1992).

The current study aimed to examine the acoustic differences between the production of American English liquids, /ɹ/, /l/, and its combination /ɹl/, by native English speakers and native speakers of Korean. With a more in-depth and comprehensive knowledge of the acoustic properties of L2 speakers, the current study further aimed to provide more clinically relevant research concerning the acoustic basis of these sounds and how they differ from Mainstream American English (MAE) acoustic patterns. With this research, we hoped to guide better clinical management or intervention for L2 speakers learning English. We sought to identify acoustic measures that are sensitive to these differences between L1 and L2 as acoustic data may inform clinicians not only about placement but also about the timing effect of constriction and its impact on production.

CHAPTER 2. REVIEW OF LITERATURE

A literature review was conducted to identify current knowledge or methodologies that provide information on the effect of first, native language (L1) on learning the sound system of a second language (L2) and to learn about acoustic characteristics of English liquid sounds, /ɹ/ and /l/. The results of the literature review are presented below in sections that address the acoustic and perceptual properties of English liquids /ɹ/ and /l/ produced by native English speakers, the perceptual characteristics of English liquids /ɹ/ and /l/ produced by Japanese and Korean learners of English, the articulatory characteristics of the Korean liquid /ㄹ/, and the characteristics of liquids produced by Korean and Japanese learners of English.

Liquids across different languages: Korean, Japanese, and English

Approximants, also referred to as semivowels, are sounds that do not involve any kind of closure of the vocal tract (Ladefoged, 2001) and more specifically, are made in such a way that one articulator is close to another without narrowing the vocal tract to create any friction (Yavas, 2006). Approximants have liquids and glides as a subcategory. In English, there are four approximants: /w, j, l, and r/ that can be divided into two categories: liquids (e.g., /ɹ/ and /l/) and glides (e.g., /w/ and /j/). Approximants exist in Japanese and Korean as well but are phonetically and phonologically different.

It is well documented in previous research that non-native speakers have significant difficulty with both perception and production of certain non-native phonetic contrasts (e.g., Flege, 1988; Goto, 1971). For example, native English speakers have little to no difficulties making the distinction between /ɹ/ and /l/. However, for learners of English, this distinction can be difficult and can lead to persistent pronunciation differences. Japanese and Korean are of particular interest due to speakers of these languages frequently misidentifying /ɹ/-/l/ contrasts and frequently reporting difficulties with their pronunciation of English /ɹ/ and /l/. These phonological and phonetic differences in liquids between these two languages result in English language learners experiencing difficulties in production and perception of English liquids. Han (2002) proposed that interference from one's native language (L1) is the primary phonological cause for the presence of foreign accents. Han (2002) proposes that a second language (L2) sound that is misidentified as a sound in their L1 category will be replaced by the L1 sound, even if L1 and L2 sounds differ phonetically. They further assert that contrasts between sounds in the L2 that do not exist in the L1 will not be registered by the speaker (Han, 2002). Therefore, it is important to understand the underlying differences between English /ɹ/ and /l/ and Japanese and Korean liquids.

In terms of English /ɹ/, there are three constrictions necessary for correct production. These places of constrictions are the pharynx, palatal vault, and the lips (Espy-Wilson et al., 2000). Research investigating the specific articulatory placement of the tongue found two distinct forms. For example, the tongue can be either bunched or in the retroflexed position to produce the English /ɹ/. For bunched, the tongue tip must be down to form a bunch near the palate. This creates contact between the sides of the tongue and the insides of the back molars. However, in the retroflex position, the tongue tip is raised and may be curled back (Espy-Wilson et al., 2000). The most common position for /ɹ/ consists of the tongue tip and blade turned up toward the hard palate with the tip pointing to, but not touching, the area behind the alveolar ridge (Han, 2002). English /l/ has two allophonic variants (i.e., light and dark /l/). The light /l/ most commonly occurs before a front vowel or glide (e.g., alive); however, the light /l/ can also occur before a low back too as well (e.g., law). The dark [ɫ] occurs in final positions, if it is syllabic, or when it precedes a back vowel. Articulatory studies have shown that the light /l/ is produced with the tongue top position being reached before the tongue dorsum position is achieved (Espy-Wilson, 1992). There is a rapid release of the tongue tip from the roof of the mouth, which creates a distinctive spectral discontinuity between /l/ and the following vowel (Espy-Wilson, 1992). However, the dark [ɫ] (e.g., full) has a more retracted tongue body and the tongue dorsum position is achieved before the tongue tip reaches the maximal position. In contrast with the light /l/ variation, there is no abrupt change.

American English liquids /ɹ/ and /l/ can occur in various phonological positions. For example, both /ɹ/ and /l/ can occur in the word initial position (e.g., right, light), in consonant clusters (e.g., pray-play), intervocalic position (e.g., arrive, alive), as well as the word-final position (e.g., poor, pool) (Han, 2002). However, Korean and Japanese liquids have particular constraints that limit phonological positions. Japanese has a syllable initial liquid, which is usually phoneticized as /ɾ/. This liquid has been described in literature as an alveolar flap, a retroflexed flap, a combination, a lax alveolar stop, and an alveolar lateral (Ingram & Park, 1998). This liquid can only be found in syllable onset positions and either word internally or word initially.

With respect to the /l-r/ distinction in onset clusters, Japanese constraints are identical to Korean. For example, neither language permits consonants + liquid clusters. English loan words containing consonants + liquid clusters are typically resyllabified by vowel epenthesis (Ingram & Park, 1998). Korean also has only one single liquid phoneme, ㄹ, that is transcribed phonetically as [l] (Han, 2002). However, ㄹ has two distinct allophones: /ɾ/ and /l/. When produced word-initially, the Korean liquid is perceived as an apical flap /ɾ/ (Han, 2002; Yune, 2016). In the in postvocalic position, the flap is produced (e.g., 달: "moon," 불: "fire," 달과: "moon and"). If produced in the intervocalic position, the alveolar lateral is used (e.g., 우리: "we," 나라: "country"). When this final lateral is immediately followed by another /l/ in the onset of the following syllable, a geminate alveolar lateral is produced (e.g., 달리: “differently”). A geminate occurs when a liquid is in a coda position and is followed by another in the onset position in the next syllable. Because Korean only has one liquid phoneme, the language does not have contrastive /ɹ/ and /l/ in the same way as English (Borden et al., 1983). Furthermore, Korean orthography does not distinguish these liquids, and Korean speakers learning English report difficulty in distinguishing the /ɹ/ and /l/ contrast in English (Borden et al., 1983). In Korean, there are several phonological constraints, also referred to as the law of initials, that influence how the Korean /l/ is produced. When in a word-initial position or after a consonant, the /l/ is produced as an [ɾ]. The Korean /l/ is only produced as an /l/ in syllable-final positions. When found in word-final positions, the Korean /l/ is produced as a light /l/ in English. Liquids appearing in the absolute word initial position in Korean do not exist and Korean morphology does not allow it. For example, the allophone of the Korean /l/, /ɾ/, is only found in well-known loanwords initially (Yune, 2016; Ingram & Park, 1998). For this reason, it is significantly more difficult for Korean speakers to perceptually identify /l-r/ in word initial positions (Ingram & Park, 1998).

Articulatory Properties of the Korean liquid /ㄹ/

Very few studies have been conducted examining the articulatory properties of the Korean liquid, /ㄹ/. It is still unclear what articulatory characteristics drive the lateral vs. flap percepts of this liquid phoneme. One study attempted to examine these characteristics. In a study conducted by Lee, Goldstein, & Narayanan (2015), the articulatory composition of this liquid was examined using real-time MRI. The phoneme /ㄹ/ is unique as it is the only Korean liquid phoneme. This phoneme has multiple allophones that alternate between lateral and rhotic articulations (Lee et al., 2015). Traditionally, this liquid allophony has been described as a single phoneme which changes perceptually at the phonological level. /ㄹ/ becomes a flap [ɾ] syllable initially, but a lateral [l] syllable finally. The onset and coda liquids in Korean are perceived as laterals, but their production involves the tongue body raising rather than retracting, whereas the English flap is distinguished articulatorily through tongue side contact and tongue body retraction. Lee et al. (2015) hypothesized that there would be at least two underlying phonological gestures comprising the single liquid phoneme — tongue tip and tongue body gestures in every context. They also hypothesized that the underlying phonological gestures would be influenced by phonological contexts (i.e., syllable positions) so that the articulatory variability of /ɾ/ would be caused by gestural reduction of the tongue body.

Previous research on English laterals show that even if the tongue tip gesture is reduced (e.g., which occurs in inter-vocalic positions) the liquid is still perceived as an /l/. Therefore, Lee et al. (2015) hypothesized that, in intervocalic positions, the production of the Korean liquid flap will be influenced by the movement of the tongue body and the tongue tip resulting in gestural overlap and shorter durations. Research showed that the production of the Korean liquid, /ㄹ/, depended on different syllable position contexts. The flap articulation (intervocalic contexts) was significantly different from the lateral (onset and coda contexts) articulation. The flap articulation showed a significantly reduced gestural movement of the tongue tip and tongue root. Furthermore, in the flap context, the tongue body did not raise at all. Lee et al. (2015) proposed that this difference signifies that there is not only a difference in gestural movements but also a complete difference in gestural composition. However, the exact difference in the gestures between the flap and lateral contexts remain unclear.

Although the two allophones of the Korean /ㄹ/ are articulatorily distinct depending on its position in a word, neither sound requires the tongue retraction required for the English flap.

Because native Korean speakers are unaccustomed to the degree of tongue retraction required to produce the English flap, it may be motorically more difficult for them to produce as it is a new sound. Furthermore, the dark [ɫ] requires significantly more tongue retraction than the light /l/, which is not required of either allophone of /ㄹ/ in Korean. Because /ㄹ/ showed significantly decreased tongue tip and body elevation and retraction, it may be articulatorily more difficult for native Korean speakers to both perceptually identify and physically produce the English /ɹ/ and the dark [ɫ].

Acoustic Properties of English /ɹ/ and /l/

Historically, acoustic models of complex liquids, such as /ɹ/ and /l/, have been less developed than models for vowels and obstruent consonants (Espy-Wilson et al., 2000). The need for further research is essential as the contrast between English /ɹ/ and /l/, especially for native speakers of Japanese and Korean, is considered to be the most crucial teaching and learning target (Saito & Lyster, 2011). Understanding the acoustic properties of /ɹ/ and /l/ of native speakers elucidate how to improve non-native speakers' productions. It is therefore necessary to examine how /ɹ/ and /l/ differ acoustically. Findings suggest that the acoustic difference between /ɹ/ and /l/ depends primarily on the frequency values of the third formant frequency (F3). Saito & Lyster (2011) found that listeners tended to perceive the stimulus sound as an /ɹ/ when its F3 dipped below 2000 Hz and as /l/ when its F3 exceeds 2400 Hz or more (Saito & Lyster, 2011). The importance of F3, particularly for the English /ɹ/, is well documented. Past studies investigating the spectral differences between all English sonorants (e.g., /ɹ/, /w/, and /l/) determined that the first formant, F1, did not reliably differentiate between any of the sounds as /ɹ/, /l/, and /w/ all had low frequency values. However, the second formant frequency (F2) values were able to discriminate /w/ from /l/ and /ɹ/, as /w/ had a consistently lower F2 value across gender and age when compared to F2 values of /l/ and /ɹ/. However, to distinguish /ɹ/ from /l/, researchers found that F3 was crucial. Dalston (1975) found that the third formant frequency, F3, value of /ɹ/ was very low when compared to both /w/ and /l/.

When examining the temporal characteristics of English sonorants, Dalston (1975) found that both F1 and F2 steady states (e.g., place of contact) for /l/ were longer than those of /w/ and /ɹ/. Dalston (1975) hypothesized that this was due to the fact that /l/ is the only liquid that requires physical contact between the tongue and another articulator. Dalston (1975) further observed that the identification of /l/ relies on the sudden shift up of F1 from the lateral to the following vowel. This sudden shift in F1, caused by the rapid movement of the tongue tip away from the roof of the mouth to the position of the next vowel, is accompanied by a transient click that can further help identify /l/ (Dalston, 1975). However, this transient click is present in the light /l/ allophone only (Espy-Wilson, 1992). This is most likely due to the fact that when producing the dark [ɫ] allophone, the tongue body is more retracted than in the light /l/, which causes a lower F2 value. Espy-Wilson (1992) found that apical contact for the dark [ɫ] is not as strong and that an increased speaking rate caused articulatory undershoot, causing a more gradual and less robust shift from F1 to F2. This is in direct contrast with the light /l/ as the transient click is often accompanied by a rapid release of the tongue tip from the roof of the mouth.

However, English /ɹ/ does not require contact between the tongue and another articulator, which makes its F3 values increasingly crucial to its identification. Furthermore, acoustic models of /ɹ/ must take into account three different constriction sites along the vocal tract, adding to its complexity. The English /ɹ/ is even more interesting acoustically, as it is often cited as, “An example of a many-to-one articulatory-acoustic relationship because different configurations produce essentially equivalent acoustical profiles (Espy-Wilson et al., 2000). English /ɹ/ is characterized by a stable acoustic pattern of F3 lowering close to the value of F2 (Espy-Wilson et al., 2000). Espy-Wilson et al. (2000) sought to outline specific acoustics of the English /ɹ/ in various word positions. The researchers found that both a syllable nucleus and consonantal position involved a F1-F2 pattern similar to a central, rounded vowel together with a very low F3 (Espy-Wilson et al., 2000). In the intervocalic position, the /ɹ/ showed a severe dip in the F3 value, causing it to approach or actually merge with its F2.

The perturbation theory is one prominent theory that attempts to account for this severe dip in F3 values that distinguish /ɹ/ from other sounds. This theory hypothesizes that certain points of constriction along the vocal tract where standing waves have maximum volume velocity can lower the natural resonance of the tract (Espy-Wilson et al., 2000). This theory proposes that lowering the effect of all three regions of constriction (e.g., pharyngeal, palatal, and labial regions) can produce the low F3 found in the typical /ɹ/. However, studies have shown that the Perturbation Theory does not make appropriate predictions of constriction location for /ɹ/. Espy-Wilson et al. (2000) found that the effects of eliminating the pharyngeal constriction of F3 were minimal and that the addition of the sublingual space was crucial for achieving F3 values representative of an English /ɹ/. The authors further found F3 is a front cavity resonance where the front cavity includes a lip constriction formed by tapering gradient of teeth and lips, with or without rounding, and a large volume cavity behind it that includes the sublingual space. Another theory posits that speakers actually place their constrictions for /ɹ/ at extremely specific points along the vocal tract (i.e., the points of maximum volume velocity). A finding that speakers place constrictions at other points would be evidence against the Perturbation Theory. This wide variation in appropriate constriction locations for the production of /ɹ/ creates a unique issue in the instruction of accent management therapy. Without a specific “correct” placement, teaching English language learners to accurately produce /ɹ/ is significantly more difficult.

Korean and Japanese Perception of English /ɹ/ and /l/

Theories of L2 speech perception posit that the perceived relation between phonetic segments encountered in an L2 play a key role in how those phonetic segments will be discriminated (Aoyama et al., 2004). According to the Speech Learning Model (SLM), discrimination may become difficult when two L2 sounds are both identified as instances of a single L1 category (Flege, 1995; Hattori & Iverson, 2009). This difficulty is particularly salient for speakers of Japanese and Korean as both languages have only one liquid phoneme, whereas English has two. In contrast, if the perceived phonetic dissimilarity between an L2 sound and the closest L1 sound is greater, SLM posits that learners will be able to discern the differences between the sounds. This is particularly crucial for English learners as Flege (1995) asserted that a native-like L2 speech perception is a necessary foundation for native-like production.

The Perceptual Assimilation Model (PAM) attempts to account for L2 perception by positing that L2 phonemes are perceived based on their articulatory similarity to the listener’s closest L1 phonemes (Best, 1995; Best et al., 2001; Hattori & Iverson, 2009). PAM uses the example of the Japanese apico-alveolar tap (e.g., /ɾ/) to support its claim. This flap has been described as being related to the English /ɹ/ and /l/. It is more rapid than English but has a range of F2 and F3 frequencies that overlap with those of English /ɹ/ and /l/ (Lotto et al., 2004). According to PAM, Japanese adults find /ɹ/ and /l/ perceptually difficult to distinguish because both the English /ɹ/ and /l/ are the same with regard to the Japanese phonological system (Best, 1995; Best et al., 2001). However, recent research found that PAM cannot fully explain Japanese difficulty with English /ɹ/ and /l/. Hattori & Iverson (2009) found that Japanese adult listeners could hear acoustic variations in the /ɹ/ and /l/ that did not affect perceived "goodness." The English /l/ was assimilated into the Japanese /r/ category to some degree and the assimilation of /ɹ/ and the Japanese /ɾ/ was comparatively weak, indicating that both /ɹ/ and /l/ are not automatically perceived to be the same.

While PAM is able to explain some aspects of L2 phoneme perception, the results showed that it was unable to account for varying F3 values. Furthermore, Hattori & Iverson (2009) found that individuals who had more nativelike exemplars along the F3 dimension were more accurate at identifying /ɹ/ and /l/, which further solidifies the importance of F3 in the perception and production of English liquids. F3 has been known to be problematic for Japanese and Korean speakers for a long time (Miyawaki et al., 1975). Recent work suggests that Japanese adults have difficulty learning English /ɹ/ and /l/ because they are overly sensitive to acoustic cues that aren't reliable for /ɹ/-/l/ categorization (e.g., F2 frequency). Researchers hypothesized that Japanese speakers were more likely to form category representations based on cues that are not necessary to native listeners due to perceptual sensitivities (Iverson et al., 2003).

However, the Speech Learning Model suggests that Japanese speakers will show greater learning for /ɹ/ than /l/ because the English /ɹ/ is more dissimilar from the Japanese /ɾ/ than the English /l/ (Aoyama et al., 2004). However, Korean listeners had trouble with l/r perception in all three positions. Perception in the word-initial singleton is more difficult for Korean listeners than in other environments, as word-initial liquids do not occur in Korean. A study conducted by Jun (2003) posited that Koreans are not normally aware of the difference between [l] and [ɾ] and will perceive both /l/ and /ɹ/ as the phoneme /l/. The results found that Korean speakers were able to perceive /l/ better than /ɹ/ and that the perception of /ɹ/ in word final positions were the most difficult for Koreans among all positions of /l/ and /ɹ/. Jamieson and Yu (1996) further supported these findings as the results from their study found that Korean listeners identified English /ɹ/ and /l/ sounds mostly accurately when these target sounds were presented in the final singleton position.

Korean and Japanese Production of English /ɹ/ and /l/

Previous research has shown that foreign accents persist even for highly proficient speakers of a non-native language (Bradlow et al., 1999). Many researchers believe that these accents persist due to production and perception being intricately connected processes. For example, Han (2002) hypothesized that accurate pronunciation of L2 requires the speaker to reactivate the continuous mode of perception as they believed that L2 production errors had a perceptional basis. According to research conducted by Han (2002), the critical age for language learning does not extend beyond the age of six. Beyond this age, it was unlikely for an individual to achieve complete mastery over a language. Many studies have examined the Critical Age Hypothesis, which posits that older learners may have neurological or motor skill constraints as their articulatory habits may be firmly established or their perception of foreign accents may rely heavily on sound contrasts of their mother language (Han, 2002). However, the exact age one's neurological or motor skills inhibit the ability to learn a new language varies among studies. While some research has shown that the Critical Age for language learning could be as early as five years, other research has indicated that the age could extend past twelve. The Speech Learning Model posits that a native-like L2 perception is a necessary foundation for native-like production. However, it is not fully known whether perception precedes production or if production precedes perception (e.g., Flege et al., 1995; Flege et al., 1996; Sheldon & Strange, 1982, Kang, 1999). Bradlow et al. (1997) found that some Japanese subjects were able to produce identifiable /ɹ/ and /l/ tokens even though they were unable to identify native English /ɹ/ and /l/ tokens reliably. However, in a later study, Bradlow et al. (1999) found that improved /ɹ/-/l/ production by Japanese speakers occurred after perceptual training procedures. The researchers further found that the improved /ɹ/-/l/ productions were retained several months after the perceptual training procedure, indicating that these processes are intricately linked.

Native-Japanese speakers often have difficulties identifying and producing English consonants /ɹ/ and /l/. According to Hattori & Iverson (2009), Japanese speakers may improve their production and perception of English /ɹ/ and /l/ through the role of assimilation. They hypothesized that a causal relationship between L1 assimilation and L2 category learning exists. Meaning that, those who assimilate phonemes less strongly into L1 categories will be better at identification and production. Alternatively, it is possible that those who are more accurate with English /ɹ/ and /l/ have changed their patterns asymmetrically. For example, past research examining the effects of auditory training on L2 production found that Japanese listeners learning English /ɹ/-/l/ may increase assimilation of the English /l/ into the Japanese /r/ category but not change their assimilation of English /ɹ/ (Hattori & Iverson, 2009). With regards to Korean production of English stop consonants (e.g., /p/, /b/, /t/, /d/, /k/, /g/), accuracy rates indicated a production-leading-perception pattern (Hao & de Jong, 2015). This pattern may be due to the fact that learners could be more easily trained in articulatory gestures, whereas perception training requires learners to cope with the high variability of acoustic cues.

As discussed, Korean has two liquid phones [r], an apical flap, and [l], a light apical lateral. Because Korean does not allow liquid phones to appear in word-initial position (with the exception of loanwords) and word-initial consonant clusters, the production of English liquids is significantly difficult for Korean native speakers. Kang (1999) found a distinct error bias in production of /l/ that appeared to be closely related to the fact that modern Korean allows flap [ɾ] in word-initial positions when it borrows words from foreign languages. Both English /ɹ/ and /l/ are substituted by Korean [ɾ]. For example, [ɾailak] ‘lilac,’ [ɾemon] ‘lemon,’ [ɾadio] ‘radio,’ and [ɾum] ‘room’ are the forms in which these words are borrowed into Korean (Kang, 1999). Kang (1999) proposes that this nativization process has caused the /l/ to /ɹ/ bias in production.

A study conducted by Lee and Hwang (2016) sought to identify priorities in pronunciation learning for Korean speakers learning English. Until the 1960s, pronunciation training encouraged L2 learners of English to acquire native-like pronunciation. However, this shift in pronunciation training protocol caused many phoneticians and teachers to attempt to create a set of priorities to help L2 learners acquire intelligible pronunciation with minimum time and effort (Lee & Hwang, 2016). Lee and Hwang (2016) established a learnability gradient in young Korean subjects trained on American English phonemes through the use of High Variability Phonetic Training (HVPT). HPVT presents naturally produced recordings of several different speakers in multiple phonetic contexts. Furthermore, immediate feedback is given to the trainee (Lee & Hwang, 2016). The results showed that with HVPT training, Korean learners of English performed better on consonant identification and had an increase in identification accuracy for most consonant sounds. The findings also suggested that the discrimination of /ɹ/ and /l/ sounds was not very difficult for Korean listeners. Upper level learners distinguished /ɹ/ and /l/ fairly successfully in both pretest and posttest, and lower level learners showed large improvements in /l/ identification (Lee & Hwang, 2016). However, production of /ɹ/ and /l/ remained difficult for both groups.

These results are further supported by a study conducted by Kang (1999) who found significantly more errors on production of /l/ than of /ɹ/. For speakers with lower production abilities, more errors were made for /ɹ/. However, for higher production abilities, the speakers made more frequent errors in /l/ to [r] direction. Kang (1999) also found that more errors were produced for /ɹ/ in consonant clusters than when in singleton regardless of group. Based on these results, they established the English liquids /r-l/ as a high priority for Korean speakers as these phonemes have a high functional load in the English language.

Present Study

The present study examined acoustic characteristics of the American English liquids, /ɹ/, /l/, and its combination /ɹl/, produced by native speakers of Korean. Given the relative lack of such data compared to other languages, such as Japanese, the findings of the study provided empirical data outlining the specific acoustic differences between native English speakers' and native-Korean speakers' production of American English liquids /ɹ/, /l/ and /ɹl/ (Kang, 1999; Lee & Hwang, 2016). One strength of the present study was the use of connected, spontaneous speech (passage reading) to elicit the sounds of target. Previous studies have used non-real words or target sounds in isolation. Furthermore, the present study included the /ɹl/ blend in the stimuli (e.g., Carl, girl, curl, etc.) which had not been previously studied. The current study addressed this gap in knowledge by incorporating the sentence: Carl got a croaking frog. Second, the study determined whether the acoustic data of Korean speakers varied between phonetic stimuli. Many studies focus on target-words in isolation. However, this study used a reading passage, The Caterpillar passage (Patel et al., 2013) to facilitate natural speech found during conversation. The present study proposed the following research questions:

(1) Do the selected measures exhibit significant differences in word-final /ɹ , l, ɹl/ between the two speaker groups? (group effect)

(2) Do certain measures, spectral or temporal, exhibit significant acoustic differences among the three phonetic stimuli? (phonetic stimuli effect)

(3) Do L2 speakers show greater differences compared to L1 speakers for certain phonetic stimuli? (Interaction between group and phonetic stimuli)

Based on prior studies, it was expected that (1) temporal and spectral measures would reveal differences in F2 and F3 trajectories between the two speaker groups, with English speakers having a shorter vowel duration and word duration than Korean speakers of English, (2) temporal and spectral measures of would reveal differences among the three phonetic stimuli, and (3) temporal and spectral measures would reveal differences between L1 and L2 speakers for each phonetic stimuli.

CHAPTER 3. METHODS

Participants and Speech Tasks

A total of 27 speakers, 14 native Korean speakers (7 men, 7 women) and 13 monolingual speakers of English (3 men, 10 women) participated in the current study. Native English speakers (L1) were either graduate students or students enrolled in Louisiana State University in Baton Rouge, Louisiana. They were adults with a median age of 20 years (range: 19-23 years). All were born in the United States. L2 speakers were recruited through email and were paid for their participation. Both groups of L1 and L2 speakers reported no history of hearing, speech, or language difficulties. All participants were adults with a median age of 37 years (range: 23-58 years). Prior to recording, additional background information of L2 speakers was obtained including the age of age of arrival (as often indexed by age of arrival, AOA, in the US) and time of residence in the US. L1 and L2 speaker data are summarized in Table 1.

All participants were asked to read The Caterpillar passage (Patel et al., 2013) (See Appendix A). Because one the phonetic sequences of our interest, /ɹl/, is not included in the passage, the participants were additionally asked to read the following sentence, Carl got a croaking frog. The word stimuli that were selected for acoustic analyses are summarized in Table 2 (next page).

Table 1. L1 Speaker Data and L2 Speaker Data (age, AOA, time of residence).

Median Age (Range)		Average Age of Arrival (Range)	Average Time of Residence (Range)
L1	L2	L2	
20 years (19-24 years)	37 years (23-58 years)	28 years (23-51 years)	14 years (3 weeks - 32 years)

Table 2. List of Word Stimuli.

	/ɹ/	/l/	/ɹl/
Word-final	Caterpillar x 6 Coaster Car Roller	Well Tall	Carl x 5

Data Collection

Acoustic data were simultaneously collected along with kinematic data in a sound-attenuating booth using an electromagnetic articulography system, the WAVE (NDI, Canada). An AKG C1000S microphone positioned approximately 30 cm from the speaker was used to record the speech stimuli with a sampling rate of 20 kHz and 16-bit quantization. Sensors were attached to the tongue back and tongue front, upper and lower lips, and jaw. As the application of sensors may alter the speaker's natural articulatory movements, the participants were instructed to speak for ten minutes prior to the introduction of stimuli in order to adapt to the presence of sensors (Dromey & Hunter, 2016). Kinematic data were collected but not used in the current study. Future studies will examine the kinematic data in relation to the acoustic differences of /ɹ, l, ɹl/ between L1 and L2 speakers.

Data Analyses

Acoustic analyses included temporal and spectral measures. Temporal measures included (1) relative timing of maximum constriction and (2) duration of vocalic nuclei, instead of duration of liquids, in consideration of the difficulty of segmenting liquids from the surrounding transition (Dalston, 1975). The relative timing of maximum constriction was determined by locating the time point in which F2 and F3 were closest during the duration of the vocalic nucleus. We included this measurement as currently no studies have examined the relative timing of maximum constriction or its potential to identify acoustic differences between L1 and L2. Two spectral measures included

(1) the Euclidean distance (kHz^2) between /ɹ/ and /l/ and (2) the frequency difference between F2 and F3. Temporal and spectral measures are summarized in Table 3. These measures were selected based on prior research on acoustic characteristics of liquids that focused on the measurement of F2 and F3 values for both /ɹ/ and /l/ (e.g., Dalston, 1975; Espy-Wilson, 1992). In particular, the two measures, relative timing and degree of maximal constriction, were explored in the current study as an attempt to develop a set of acoustic measures that are sensitive to foreign accent, especially in native speakers of Korean. In order to account for gender differences, frequencies were normalized using the Bark Scale (Smith & Abel, 1999). To account for durational differences of vocalic nuclei, the time was normalized to fit a scale of 100. Spectrogram and waveform displays were used to identify the target stimuli using TF 32 (Milenkovic, 2005). The formant extraction was based on linear predictive coding (LPC) analysis and manual correction was performed when errors of automatic extraction were found. Data were analyzed using a custom R script (R Core Team, 2013).

Table 3. Summary of Selected Acoustic Measures.

Temporal	Spectral
Duration of vocalic nuclei	Euclidean distance between /ɹ/ and /l/ at (1) the temporal midpoint and (2) the relative timing of maximum constriction
Relative timing of maximum constriction	Frequency difference between F2 and F3 at (1) the temporal midpoint and (2) the maximum constriction

CHAPTER 4. RESULTS

The following results are divided into three sections according to the three aforementioned research questions. Therefore, each section of the results reports the (1) effects of speaker group, (2), effects of phonetic stimuli, and (3) interaction effect between group and phonetic stimuli. Duration (ms), maximum F2-F3 constriction (kHz), and relative timing of maximum constriction (%) was calculated for both speaking groups and for all three phonetic stimuli (i.e., /ɹ/, /l/, and /ɹl/). A two-way multivariate analysis of variance (MANOVA) was performed to examine the main effects for group and phonetic stimuli, as well as their interaction between the two.

Reliability of Acoustic Measurements

Because the temporal segmentation was the only factor affecting spectral measures selected in the study, around 20% of the data were randomly selected and the duration of vocalic nuclei of the data was remeasured approximately four months after initial measurement. Using a random number generator, five subjects from L1 and five subjects from L2 were randomly selected with each target word from the Caterpillar Passage (Patel et al., 2013) and the sentence Carl got a croaking frog remeasured. The intra-measurer reliability was determined by calculating a Pearson Product Moment Correlation coefficient; $r = .98$ ($p < .05$). The inter-rater correlation coefficient was .95 ($p < .05$) computed with the second measurer blinded to the original measurement. Both correlation coefficients were considered a strong agreement.

A. Effects of Speaker Group

Table 4 summarizes the results of the two-way MANOVA indicating the main effect of speaker group (e.g., native English speakers (L1) and Korean speakers (L2) on each measure. A significant main effect of duration was found. Although significant main effects of maximum F2-

F3 constriction (kHz) and relative timing of maximum constriction (%) were not found, significant interactions between group and phonetic stimuli for these two dependent measures were found.

Table 4. Results of the two-way MANOVA for speaker group effects.

	L1		L2			
Measure	Mean	Std. Dev.	Mean	Std. Dev.	F	p-value
Duration (ms)	226.61	58.15	302.56	88.39	90.28	< .001
Maximum F2-F3 constriction (kHz)	0.64	0.39	0.80	0.37	1.54	n.s.
Relative timing of maximum constriction (%)	58.93	36.40	54.15	38.06	3.03	n.s.

Note. Std. Dev. = standard deviation.

L2 speakers produced significantly longer semivowels compared to L1 speakers. However, the durational pattern among the three stimuli was the same with native L1 speakers (Figure 1).

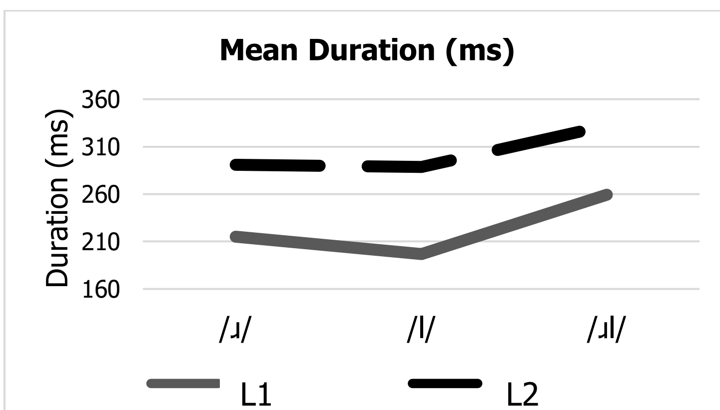

Figure 1. Vocalic durations for L1 and L2 speakers.

A separate statistical analysis, a one-way analysis of variance (ANOVA), was conducted to examine the difference in the Euclidean distance between /ɹ/ and /l/ between the two speaker groups. Table 5 summarizes the results of the one-way ANOVA indicating the Euclidean distance between /ɹ/ and /l/ between groups. The Euclidean distance between the two sounds on the F2-F3 plane was measured from two time points as previously addressed, (1) the temporal midpoint of vocalic nuclei and (2) the maximum constriction. Speaker group exhibited significantly different Euclidean distances between /ɹ/ and /l/ for both time points (Table 5) That is, L1 speakers had greater Euclidean distances between the two sounds than L2 speakers. However, the Euclidean distances were more apparent when measured from the point of maximum F2-F3 constriction compared to the temporal midpoint. Figure 2 (next page) displays the Euclidean distance between /ɹ/ and /l/ for L1 and L2 speakers.

Table 5. Results of the one-way ANOVA for the difference in Euclidean distance between /ɹ/ and /l/ for the two speaker groups.

	L1		L2			
Measure	Mean	Std. Dev.	Mean	Std. Dev.	F	p-value
Euclidean distance (kHz_2) measured at the temporal midpoint	1.48	0.47	1.05	0.49	5.41	0.02
Euclidean distance (kHz_2) measured at the relative timing of maximum constriction (%)	3.00	2.25	1.26	0.52	6.83	0.01

Note. Std. Dev. = standard deviation.

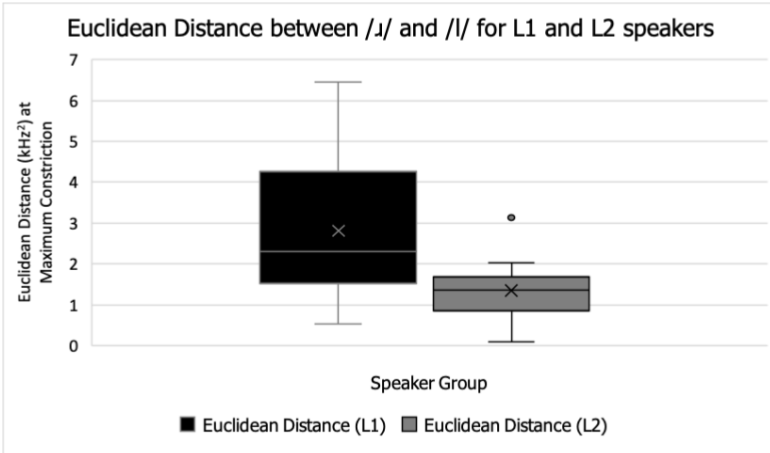

Figure 2. Euclidean distance (kHz2) between /ɹ/ and /l/ for L1 and L2 speakers.

B. Effects of Phonetic Stimuli

Table 6 summarizes the results of the two-way MANOVA indicating the main effect of phonetic stimuli (e.g., /ɹ/, /l/, and /ɹl/) on each measure. These results indicate that the three phonetic stimuli were significantly different for all three dependent measures across speaker groups.

Table 6. Results of the two-way MANOVA for phonetic stimuli effects.

	/ɹ/		/l/		/ɹl/			
Measure	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	F	p-value
Duration (ms)	254.50	87.10	244.36	87.06	294.86	68.44	17.23	< .001
Maximum F2-F3 constriction (kHz)	0.61	0.38	1.00	0.43	0.81	0.28	38.20	< .001
Relative timing of maximum constriction (%)	60.87	37.93	51.80	41.77	50.12	32.90	5.00	.007

Note. Std. Dev. = standard deviation.

C. Interaction Between Speaker Group and Phonetic Stimuli

Table 7 summarizes the results of the MANOVA indicating the effect of speaker group (e.g., L1 and L2) on each of the phonetic stimuli (e.g., /ɹ/, /l/, and /ɹl/). The interaction between speaker group and phonetic stimuli was significant with the exception of duration.

Table 7. Summary of the interactions for the two-way MANOVA.

	Measure	F	p-value
Group x Phonetic Stimuli Interaction	Duration (ms)	0.33	n.s.
	Maximum F2-F3 constriction (kHz)	13.42	< .001
	Relative timing of maximum constriction (%)	35.43	< .001

Note. Std. Dev. = standard deviation.

The interaction between speaker group and phonetic stimuli reached significance in every context (i.e., maximum F2-F3 constriction, relative timing of maximum constriction) except for the duration (ms). The durational pattern among the three stimuli was the same between L1 speakers and L2 speakers as shown in Figure 1. However, in terms of the degree of constriction, L1 speakers produced varying degrees of F2-F3 constrictions across phonetic stimuli, whereas, L2 speakers produced a similar degree of constriction across semivowels (Figure 3, left, next page). The relative timing of maximum constriction occurred latest in /ɹl/ and earliest in /ɹ/ production for L2 speakers, while L1 speakers produced a relative timing of maximum constriction earliest in /ɹl/ and latest in /ɹ/ (Figure 3, right, next page).

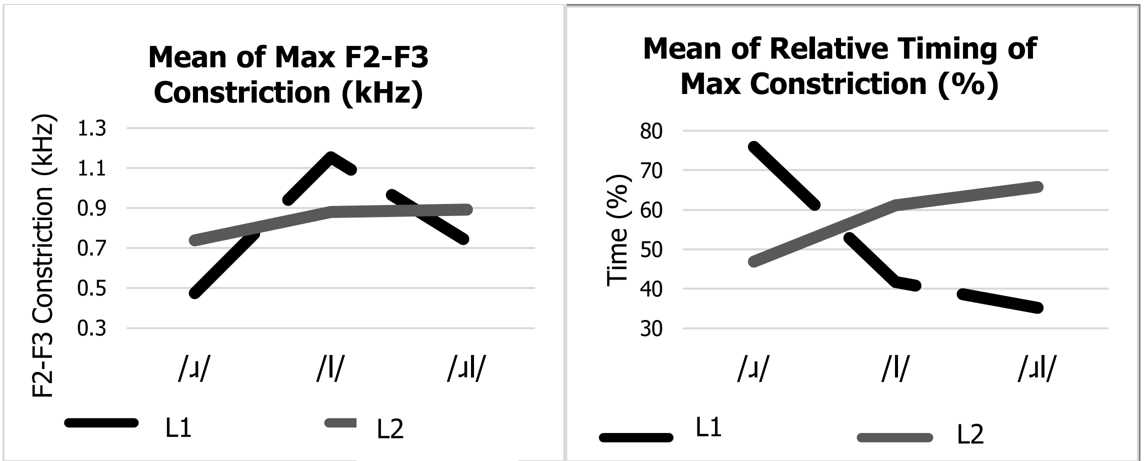

Figure 3. Interaction effects between L1 and L2 Speakers across phonetic stimuli.

CHAPTER 5. DISCUSSION

General Discussion

With an ever-growing population of English language learners (ELLs) in the United States, it is increasingly important to understand the acoustic differences which affect the perception and production of difficult sounds. Many studies have examined the acoustic differences in American English liquids (e.g., /ɹ, l, ɹl/) as produced by Japanese speakers; however, few studies currently exist that examine these differences for Korean speakers. The main focus of this study, therefore, was to investigate the acoustic differences of word-final American English liquids as produced by Korean speakers, focusing on identifying acoustic measures that are sensitive to differences between L1 and L2

The results of this study show that American English liquids produced by L2 speakers are acoustically different from liquids produced by L1 speakers. When compared to L1 speakers, the duration of the liquids, relative timing of maximum constriction patterns, and the degree of constriction (i.e., F2-F3 difference) showed opposite patterns.

Although L2 speakers produced longer liquids when compared to L1 speakers, the durational pattern did not vary between the liquids themselves. To date, there has been no attempt to directly compare the "typical" (normal) rate characteristics between English and Korean. In a study examining the acoustic predictors of speech intelligibility with Parkinson's Disease, Kim & Choi (2017) used articulation rate (AR) as a measure across languages. The researchers found that ARs from individual studies of Korean and English revealed similar rates for spontaneous speech (4-5 syllables/sec.). Many L2 speakers reported significant difficulty in the production of /ɹ/, /l/, and /ɹl/ in particular. Given the similar speaking rate between L1 and L2 speakers, the significant longer duration of L2 speakers may reflect their cautious production of these sounds. It is possible that the L2 speakers, in an attempt to sound more nativelike, spoke more slowly in order to potentially improve perceptual clarity and understanding.

Other possible explanations for these differences between L1 speakers and L2 speakers can be found when taking phonological and articulatory differences between English and Korean into account. Korean liquids, [r] and [l], are significantly different from the liquid sounds of English in both articulatory and acoustic characteristics (Kang, 1999). As reported by Lee et al. (2015), the Korean liquid does not require the tongue retraction required for the English /ɹ/. Furthermore, the dark [ɫ] requires significantly more tongue dorsum retraction than the light /l/, which is not required of either allophone of /ㄹ/ in Korean. These differences between languages due to articulatory differences and phonological constraints, may account for the extended duration seen in Korean speakers. Because American English liquids are motorically different and phonologically distinct from the Korean liquid, it was expected that an unfamiliar or non-native phoneme would be more difficult for L2 speakers to produce.

L2 speakers produced similar degrees of constriction across the phonetic stimuli, whereas L1 speakers produced different degrees of constriction for each of the phonetic stimuli. Acoustically, specifically for /ɹ/ and /ɹl/, F3 values for L2 speakers did not dip significantly when compared to L1 speakers. An explanation for these findings can be found when examining the results of past research on the acoustic characteristics of American English liquids. As reported by Espy-Wilson et al. (2000), English /ɹ/ is characterized by a stable acoustic pattern of F3 lowering close to the value of F2. In the intervocalic position, the /ɹ/ showed a severe dip in the F3 value, causing it to approach or actually merge with its F2. However, both the light /l/ and the dark [ɫ] allophones do not require a strong dip in F3 (Dalston, 1975). Instead, the difference between F2 and F3 was expected to be the greatest among the phonetic stimuli for L1 speakers.

Interestingly, for L1 speakers, the combination /ɹl/ showed components of both /ɹ/ and /l/. Similar to /ɹ/, there is a distinct dip in F3, which caused a significant constriction. While it is not as pronounced as the constriction found in /ɹ/, it is significantly more marked than the constriction found in /l/. However, similar to /l/, the F3 value increases as the duration continues. At the end of the vocalic nucleus (e.g., Carl), the F3 value has reached a value high enough to be perceptually identified as an /l/.

Another interesting component of /ɹl/ is its relative timing of maximum constriction. In terms of the timing of maximum constriction, /ɹl/'s maximum constriction occurred earliest in the vocalic nucleus, with the maximum constriction occurring latest in /ɹ/. This difference between /ɹl/ and /ɹ/, for L1 speakers, in the timing of maximum constriction can be accounted for by /ɹl/ containing the additional /l/ at the end of the vocalic nucleus. As discussed earlier, the maximum constriction of /l/ does not show a significant constriction due to its F3 remaining stable throughout production. For L1 speakers, these results indicate that /ɹl/ contains components of both /ɹ/ and /l/ in that the initial dip in F3 reaches similar constriction levels as /ɹ/ but shows an increase in F3 values toward the end of the vocalic nucleus to reflect levels closer to F3 values found in /l/. Furthermore, the maximum constriction occurs earlier in the vocalic nucleus due to /ɹ/ occurring earlier than /l/ in the phonetic sequence. For L2 speakers, F3 values for /ɹl/ did not change significantly when compared to L1 speakers in terms of both the relative timing of maximum constriction and the maximum F2-F3 constriction.

In general, L2 speakers produced opposite patterns in terms of the dependent variables for /ɹl/. Possible explanations for this difference can be found when examining the acoustic and articulatory requirements for the production of this sound. As discussed earlier, /ɹl/ contains acoustic characteristics of both /ɹ/ and /l/. Because these sounds in isolation have been identified as difficult sounds for Korean speakers, the combination of these sounds adds additional acoustic and, most likely, articulatory demands for Korean speakers. Further research is needed to elucidate the specific articulatory differences in the production of /ɹl/ between L1 and L2 speakers.

Additional Analysis

Based on the reduced acoustic distinction between /ɹ/ and /l/ as reported in Figure 3, and prior research on its relationship with perceptual estimation of accentedness, a Pearson Correlation was analyzed as post-hoc analysis. The correlation between the degree of accentedness and the Euclidean distance between /ɹ/ and /l/ was computed for the maximum F2-F3 constriction (Figure 4) as well as the temporal midpoint of the vocalic nucleus (Figure 5). The perceptual rating was obtained by 8 listeners on a 10-point scale interval. A score of 1 indicated a strong accent, while a score of 10 indicated no accent. This analysis did not reveal a significant relationship between the Euclidean distance and the degree of accentedness when measured at the temporal midpoint ($r(14) = 0.38$, $p = 0.18$) or at the relative timing of maximum constriction ($r(14) = 0.26$, $p = 0.36$). However, participant 7 appears to be an outlier. When removed from the post-hoc analysis, a significant relationship between the Euclidean distance and the degree of accentedness when measured at the relative timing of maximum constriction was found ($r(13) = 0.59$, $p = 0.03$). However, the relationship between the Euclidean distance and the degree of accentedness when measured at the temporal midpoint remained insignificant ($r(13) = 0.38$, $p = 0.18$). Therefore, measuring the Euclidean distance between /ɹ/ and /l/ at the relative timing of maximum constriction may have the potential to be an indicator of accentedness.

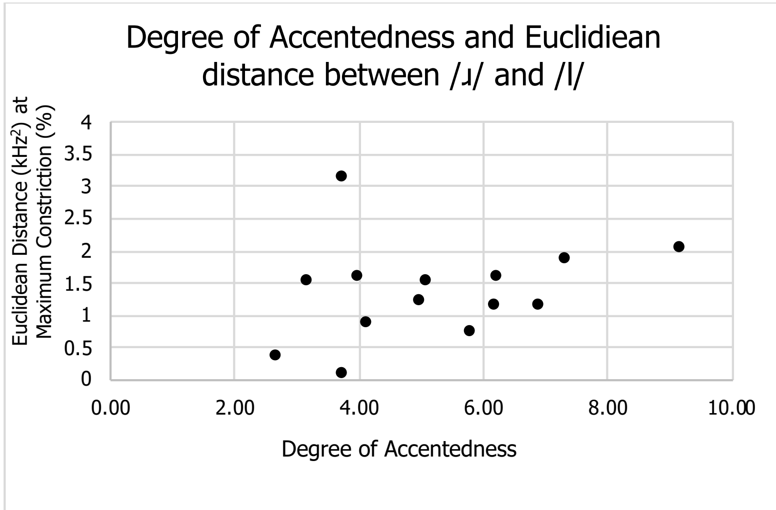

Figure 4. Degree of Accentedness on Euclidean distance between /ɹ/ and /l/ (maximum constriction).

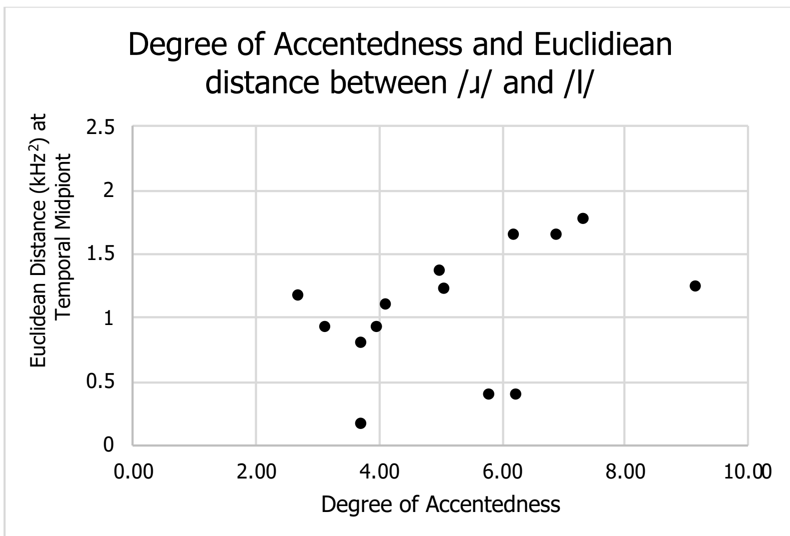

Figure 5. Degree of Accentedness on the Euclidean distance between /ɹ/ and /l/ (temporal midpoint).

Clinical Application

As this study shows, L1 speakers and L2 speakers are acoustically different in terms of duration, in the relative timing of maximum constriction, and in the degree of F2-F3 constriction. When considering target words in the use of accent management therapy for Korean speakers, the clinician should take into account several factors: the acoustic differences found in the current study, the phonological constraints of the Korean language, the nativization process described by Kang (1999), and feedback type.

The results of the current study yielded significant results that served to not only identify acoustic differences between L1 speakers and L2 speakers but also identify measures that are sensitive to these acoustic differences between languages. The results discussed above can serve as empirical, acoustic data for accent management for Korean speakers. These results show that Korean speakers often do not produce adequate degrees of constriction for /ɹ/, /l/, or /ɹl/ at the correct time during the vocalic nucleus. During accent management therapy, the clinician can use the acoustic measures, relative timing of maximum constriction and maximum F2-F3 constriction, to identify patterns of error and monitor changes for the duration of therapy.

It is important for the clinician to be aware that the Korean language has only one liquid phoneme, whereas American English has two liquid phonemes. Because Korean only has one liquid phoneme, the language does not have contrastive /ɹ/ and /l/ in the same way as English (Borden et al., 1983). Furthermore, Korean orthography does not distinguish these liquids, and Korean speakers learning English report difficulty in distinguishing the /ɹ/ and /l/ contrast in English (Borden et al., 1983). In Korean, there are several phonological constraints, also referred to as the law of initials, that influence how the Korean /l/ is produced. When in a word-initial position or after a consonant, the /l/ is produced as an [ɾ]. The Korean /l/ is only produced as an [l] in syllable-final positions. When found in word-final positions, the Korean /l/ is produced as a light /l/ in English. Liquids appearing in the absolute word initial position in Korean do not exist and Korean morphology does not allow it. For example, the allophone of the Korean /l/, /ɾ/, is only found in well-known loanwords initially (Yune, 2016; Ingram & Park, 1998). For this reason, it is significantly more difficult for Korean speakers to perceptually identify /l-r/ in word initial positions (Ingram & Park, 1998). Therefore, if the client is unable to perceptually distinguish /ɹ/ and /l/, it would be beneficial to train the client on the perceptual differences between these two phonemes.

Also, Korean does not allow liquids in the word-initial position, therefore, targeting words with initial liquids will also be beneficial. Because the combination, /ɹl/, only occurs word-finally, it is important for the clinician to incorporate such words into the accent management program. Such words include, but are not limited to earl, curl, pearl, swirl, snarl, etc. The nativization process described by Kang (1999) describes a distinct error bias in production of /l/ that appeared to be closely related to the fact that modern Korean allows flap [ɾ] in word-initial positions when it borrows words from foreign languages. Both English /ɹ/ and /l/ are substituted by Korean [ɾ]. Kang (1999) proposed that this nativization process has caused the /l/ to /ɹ/ bias in production. This nativization process could influence the production of both /ɹ/ and /l/ during accent management therapy and should be taken into account by the clinician when determining therapy goals and directions.

Findings in the literature concerning accent management therapy for L2 speakers reinforce the importance for clinicians to be aware that the client may not be able to perceptually distinguish /ɹ/ from /l/ or be able to reliably produce /ɹ/ and /l/ as distinct phonemes. Furthermore, /ɹl/ combines these two difficult phonemes and can present as a distinct and unique challenge for L2 speakers. Therefore, it is important for the clinician to incorporate a variety of feedback types in order to determine which is most effective. Acoustical analysis of speech samples with computer programs such as TF32 (Milenkovic, 2005), is one method that could provide the clinician with measurable data and the client with visual feedback.

Limitations and Future Directions

A primary limitation for this study is the target phonetic stimuli were only in the word-final position. In future studies, it would therefore be important to investigate the acoustic characteristics of American English liquids in the word-initial and the word-medial positions. Because the Korean language does not allow /ɹ/ in word-initial positions, results may show a more distinct pattern from word-final positions and can guide clinicians in target word selection and accent management therapy. In addition, the phonetic context for each sound was not ideally balanced due to the nature of the speech task used in the study. Therefore, it would be beneficial to have the participants produce multiple repetitions of each target word. Having multiple repetitions of each word would provide the clinician with more reliable data in terms of average duration, average F2-F3 difference, and the relative timing of maximum constriction. Additionally, a limitation of the current study was the inclusion of only one word, Carl, to investigate the combination /ɹl/. Therefore, it would be beneficial for future research to examine the differences between L1 and L2 speakers in a variety of words containing /ɹl/.

Conclusion

The current study found a significant interaction between speaker group (L1 and L2) and phonetic stimuli (/ɹ/, /l/, and /ɹl/). L2 speakers produced a similar degree of constriction, whereas L1 speakers produced varying degrees of constriction across phonetic stimuli. Furthermore, the relative timing of maximum constriction occurred earliest in /ɹl/ and latest in /ɹ/ production for L1 speakers. The opposite was observed for the L2 speakers. These findings serve as important foundational work for future accent management research investigating the acoustic characteristics of American English liquids as produced by L2 speakers. These results identified both the relative timing of maximum constriction and maximum F2-F3 constriction as measures that are sensitive to the acoustic differences between L1 and L2 speakers. These results show that Korean speakers often do not produce adequate degrees of constriction for /ɹ/, /l/, or /ɹl/ at the correct time during the vocalic nucleus. During accent management therapy, the clinician can use these acoustic measures to identify patterns of error and monitor improvement throughout the duration of therapy. Furthermore, building on the distinction between /ɹ/ and /l/ will further improve L2 speakers' distinct production and perception of not only these phonemes, but also the combination, /ɹl/.

APPENDIX. SPEAKER STIMULUS

The Caterpillar Passage (Patel et al., 2013). Do you like amusement parks? Well, I sure do. To amuse myself, I went twice last spring. My most MEMORABLE moment was riding on the Caterpillar, which is a gigantic roller coaster high above the ground. When I saw how high the Caterpillar rose into the bright blue sky, I knew it was for me. After waiting in line for thirty minutes, I made it to the front where the man measured my height to see if I was tall enough. I gave the man my coins, asked for change, and jumped on the cart. Tick, tick, tick, the Caterpillar climbed slowly up the tracks. It went SO high I could see the parking lot. Boy was I SCARED! I thought to myself, “There’s no turning back now.” People were so scared they screamed as we swiftly zoomed fast, fast, and faster along the tracks. As quickly as it started, the Caterpillar came to a stop. Unfortunately, it was time to pack the car and drive home. That night I dreamt of the wild ride on the Caterpillar. Taking a trip to the amusement park and riding on the Caterpillar was my MOST memorable moment ever!

References

Aoyama, K., Flege, J.E., Guion, S. G., et al. (2004). Perceived phonetic dissimilarity and L2 speech learning: the case of Japanese /r/ and English /l/ and /ɹ/. Journal of Phonetics, 32, 233-250.

Borden, G., Gerber, A., & Milsark, G. (1983). Production and perception of the /ɹ/-/l/ contrast in Korean adults learning English. Language Learning, 33(4), 499-526.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English /ɹ/ and /l/: Long-term retention of learning in perception and production. Perception & Psychophysics, 61(5), 977–985.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /ɹ/ and /l/. IV: Some effects of perceptual learning on speech production. The Journal of the Acoustical Society of America, 101(4), 2299–2310.

Dalston, R. M. (1975). Acoustic characteristics of English /w, r, l/ spoken correctly by young children and adults. The Journal of the Acoustical Society of America, 57(2), 462-469.

Dromey, C., & Hunter, E. (2016). Speech adaptation to kinematic recording sensors. The Journal of the Acoustical Society of America, 140(4), 3220-3221.

Espy-Wilson, C. Y. (1992). Acoustic measures for linguistic features distinguishing the semivowels /wjrl/ in American English. The Journal of the Acoustical Society of America, 92(2), 736-757.

Espy-Wilson, C.Y., Boyce, S.E, Jackson, M, Narayanan, S., & Alwan, A. (2000). Acoustic modeling of American English /ɹ/. The Journal of the Acoustical Society of America, 108(1), 343-356.

Flege, J. E., Takagi, N., & Mann, V. (1995). Japanese adults can learn to produce English /r/ and /l/ accurately. Language and Speech, 38(1), 25-55.

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.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds “l” and “r.” Neuropsychologia, 9, 317-323.

Han, J. I. (2002). The effect of formal instruction on acquisition of the English /ɹ/ and /l/ by Korean speakers. 어학연구, 38(2), 691-711.

Hao, Y. & de Jong, K. (2015). Imitation of second language sounds in relation to L2 perception and production. Journal of Phonetics, 54, 151-168.

Hattori, K., & Iverson, P. (2009). English /ɹ/–/l/ category assimilation by Japanese adults: Individual differences and the link to identification accuracy. The Journal of the Acoustical Society of America, 125(1), 469–479.

Ingram, J.C.L, & Park, S-G. (1998). Language, context, and speaker effects in the identification and discrimination of English /ɹ/ and /l/ by Japanese and Korean listeners. The Journal of the Acoustical Society of America, 103(2), 1161-1174.

Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /ɹ/–/l/ to Japanese adults. The Journal of the Acoustical Society of America, 118(5), 3267–3278.

Jamieson, D.G., & Yu, K. (1996). Perception of English /ɹ/ and /l/ speech contrasts by native Korean listeners with extensive English-language experience. In Proceedings of ICSLP 96 the Fourth International Conference on Spoken Language Processing, 3, 1453-1456.

Jun, E. (2003). The perception of English /l/ and /ɹ/ by Korean speakers. Studies in Phonetics, Phonology, and Morphology, 9(2), 471-484.

Kang, H.-S. (1999). Production and perception of English /ɹ/ and /l/ by Korean learners of English: an experimental study. Korean Journal of Speech Sciences, 6, 7-23.

Kim, Y-J. & Choi, Y. (2017). A cross-language study of acoustic predictors of speech intelligibility in individuals with Parkinson’s Disease. Journal of Speech, Language, and Hearing Research, 60, 2506-2518.

Lee, H., & Hwang, H. (2016). Gradient of learnability in teaching English pronunciation to Korean learners. The Journal of the Acoustical Society of America, 139(4), 1859-1872.

Lee, Y., Goldstein, L., & Narayanan, S. (2015). Systematic variation in the articulation of Korean liquid across prosodic positions. Department of Linguistics and Department of Electrical Engineering, University of Southern California, PAGES.

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /ɹ/ 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), 1242–1255.

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

Milenkovic, P. (2005). TF32 [Computer Software]. University of Wisconsin-Madison.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A., Jenkins, J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of /ɹ/ and /l/ by native speakers of Japanese and English. Perception & Psychophysics, 18, 331-340.

Patel, R., Connaghan, K., Franco, D., Edsall, E., Forgit, D., Olsen, L., & Russell, S. (2013). “The caterpillar:” a novel reading passage for assessment of motor speech disorders. American Journal of Speech-Language Pathology / American Speech-Language-Hearing Association, 22(1), 1-9. http://doi.org/10.1044/1058-0360(2012/11-0134).

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Saito, K., & Lyster, R. (2011). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɹ/ by Japanese learners of English. Language Learning: a Journal of Research in Language Studies, 62(2), 595-633.

Schmidt, A.M. (1996). Cross-language identification of consonants. Part 1. Korean perception of English. The Journal of the Acoustical Society of America, 99(5), 3201-3211.

Sheldon, A. & Strange, W., (1982). The acquisition of /ɹ/ and /l/ by Japanese learners of English: evidence that speech production can precede speech perception. Applied Psycholinguistics, 3, 243-261.

Smith, J. O., & Abel, J. S. (1999). Bark and ERB bilinear transforms. IEEE Transactions on speech and audio processing, 7(6), 697-708.

Strange, W., & Dittmann, S. (1984). Effects of discrimination training on the perception of /r–l/ by Japanese adults learning English. Perception & Psychophysics, 36(2), 131–145.

US Census Bureau (2015). Retrieved from https://www.census.gov on March 1, 2019.

Yune, S. (2016). Perception of English approximants by Korean-speaking students. Department of English Education, Korea University, 64-85.

VITA

Alexis Espinal was born in Covington, Louisiana but has lived in San Antonio, Texas for half of her life. She graduated from Bowdoin College and earned a Bachelor of Arts degree in Art History and Psychology. Directly following graduation, Alexis obtained her Teaching English as a Foreign Language certification where she taught English to recently immigrated adults for one year. After deciding to return to graduate school, Alexis enrolled as a master's student in Communication Disorders at Louisiana State University. Alexis began her thesis under Dr. Yunjung Kim as a partial fulfillment of the requirements for a Master of Arts degree. Upon graduation, Alexis hopes to gain a clinical fellowship position as a hospital-based speech-pathologist and spend more time reading books and walking around the lakes with her dog Tarzan.