5-2015

Effects of Native Language on Perception and Neurophysiologic Processing of English /r/ and /l/ by Native American, Korean, and Japanese Listeners

LEE JUNG AN
Graduate Center, City University of New York

How does access to this work benefit you? Let us know!

Follow this and additional works at: https://academicworks.cuny.edu/gc_etds

Part of the Linguistics Commons

Recommended Citation
https://academicworks.cuny.edu(gc_etds/838

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Dissertations, Theses, and Capstone Projects by an authorized administrator of CUNY Academic Works. For more information, please contact deposit@gc.cuny.edu.
EFFECTS OF NATIVE LANGUAGE ON PERCEPTION AND NEUROPHYSIOLOGIC
PROCESSING OF ENGLISH /r/ and /l/ BY NATIVE AMERICAN,
KOREAN AND JAPANESE LISTENERS

by

LEE JUNG AN

A DISSERTATION SUBMITTED TO THE GRADUATE FACULTY IN
SPEECH-LANGUAGE-HEARING SCIENCES IN PARTIAL FULFILLMENT OF
THE REQUIREMENT FOR THE DEGREE OF DOCTOR OF PHILOSOPHY,
THE CITY UNIVERSITY OF NEW YORK

2015
This manuscript has been read and accepted for the Graduate Faculty in Speech-Language-Hearing Sciences in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Brett A. Martin, Ph.D.

Date

Chair of Examining Committee

Klara Marton, Ph.D.

Date

Executive Officer

Glenis R. Long, Ph.D.
Valerie L. Shafer, Ph.D.
Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK
Abstract

EFFECTS OF NATIVE LANGUAGE ON PERCEPTION AND NEUROPHYSIOLOGIC PROCESSING OF ENGLISH /r/ and /l/ BY NATIVE AMERICAN, KOREAN AND JAPANESE LISTENERS

By

LEE JUNG AN

Adviser: Professor Brett A. Martin

The perception of English liquids /r/ and /l/ is challenging for native Korean and Japanese adult speakers because these sounds are not phonemic in these languages. The Korean language has a partial phonetic model (intervocalic [ɾ]-[ll]) that could potentially facilitate processing of English /r/ and /l/ but the Japanese language does not. The purpose of this study was to compare the effects of native language on the neurophysiologic processing of English intervocalic /r/ and /l/ by native American, Korean and Japanese listeners using several event-related evoked potentials (ACC, MMN, & P3a) along with behavioral identification and discrimination. Three specific aims were investigated. The first aim was to examine the effects of native language on the perceptual identification and discrimination of English intervocalic /r/ and /l/. The second aim was to determine the effects of native language on the neurophysiologic encoding of English intervocalic /r/ and /l/ using the acoustic change complex (ACC). The third aim was to determine the effects of native language on the pre-attentive discrimination of and related attention shifting/orienting to English intervocalic /r/ and /l/ using the mismatch negativity (MMN) and P3a.
Stimuli were a synthetic vowel-consonant-vowel (VCV) continuum that generated percepts in American English listeners ranging from /iri/ to /ili/. Stimuli falling within- and across-phonetic category were presented using an oddball paradigm. The probability of occurrence of the deviant was 20%. Nine participants from each language group participated. Stimuli were presented via insert earphones at 70 dB SPL using an 1100 ms offset-to-onset interstimulus interval. The evoked potentials were recorded from surface electrodes using a Neuroscan system. Behavioral testing included a 2-alternative forced choice identification task and a 3-alternative forced choice oddity discrimination task.

English medial /r/ and /l/ were perceived in a categorical manner by Americans, in a categorical-like manner by Koreans and in a non-categorical manner by Japanese. The midline-central ACC P1-N1-P2 responses did not differ significantly between language groups, suggesting little effect of native language on the primary cortical encoding of these sounds. In contrast, the lateral-temporal ACC T-complex differed across language groups, suggesting that native language influences secondary cortical processing of these sounds.

The MMN also depended on native language, suggesting that native language influences automatic, pre-attentive discrimination of English medial /r/ and /l/. Both early (400-650 ms) and late (655-905 ms) responses were obtained for Americans and Koreans, whereas early responses were not detected in Japanese. Early MMN responses were significantly larger for across-category pairs than for within-category pairs only in Americans and Koreans. Late MMN responses were significantly larger for across-category pairs than for within-category pairs only in Americans. Additionally, late MMN responses for the across category pairs were significantly larger in Americans compared to other language groups. The P3a had both early (600-700 ms) and late (900-1000 ms) responses, similar to MMN responses. Early P3a responses were present
in Americans and Koreans and P3a latency for across-category pairs was shorter for Americans than for Koreans. Late P3a responses were obtained from all three language groups, and did not differ significantly across tokens or groups.

The partial language model available to Koreans appears to facilitate both neurophysiologic processing and behavioral perception of English /r/ and /l/. The absence of such a model in Japanese results in perceptual processing difficulties and alterations of the neurophysiologic processing of these sounds. Native language influences the neurophysiologic processing, including encoding at the level of secondary auditory cortex, pre-attentive discrimination, attention-related processing, and behavioral identification and discrimination.
Acknowledgements

I would like to express my deepest gratitude to my academic mentor, Dr. Brett Martin. Thank you for your tremendous help and support. This academic journey with you has been a great honor for me. Without your guidance and help, this dissertation would not have been possible, thank you.

My special thanks go to the members of my dissertation committee, Dr. Glenis Long and Dr. Valerie Shafer, for their devotion and generous support since my second examination. Also I thank my external examiner, Dr. Monica Wagner.

I would like to thank everyone in the Audiology and Auditory Evoked Potential Lab for their daily support and constant encouragement; specifically, big thanks to Wen-Jie Wang and Reethee Antony for their support and kindness. I would also like to thank my Korean friends at the Graduate Center, Jungmoon Hyun, Jungmee Yoon, and Youngmi Park, for kindly sharing your experience and life with me, thank you.
Dedication

I give thanks to my God who is always good and faithful.
I believe that this journey is one of your great plans toward my life.
My life is in your hand.

To my spiritual mentor, Grace Choi and my dearest friend in Jesus, Eunyoung Won
Thank for your prayer and support

To my family in Korea, my father, Si-young An, my mother, Sunhee Kang,
my brothers Jin-suk & Jin-hong, and my sister Yeon-jung
Thank for your love and support
I love you so much.
Table of Contents

Abstract .......................................................................................................................................... iv

Acknowledgements ....................................................................................................................... vii

Dedication .................................................................................................................................... viii

List of Figures ................................................................................................................................ xii

CHAPTER 1: Introduction ............................................................................................................. 1

1.1 Native Language Neural Commitment Model ...................................................................... 2

1.2 Effects of native language on perception of English /r/ and /l/ ............................................. 4

1.3 Effects of native language on neural processing of speech sounds ................................. 11

1.4 Preliminary data and rationale for the current study .......................................................... 15

1.5 Purpose of the current study and hypotheses ................................................................. 16

CHAPTER 2: Method ................................................................................................................... 19

2.1 Participants ......................................................................................................................... 20

2.2 Stimuli ............................................................................................................................... 21

2.3 Procedure .......................................................................................................................... 23

2.4 Data Analyses ................................................................................................................... 26

CHAPTER 3: Results ................................................................................................................... 30

3.2 Acoustic Change Complex (ACC) N1-P2 .......................................................................... 34
List of Tables

Table 1. A summary of English, Japanese, and Korean liquids ..................................................... 6
Table 2. Background information for the American, Japanese, and Korean participants ............ 21
Table 3. Paired t-test results on the grand mean MMN waveforms. .............................................. 51
Table 4. Summary of significant results as a function of language group................................. 57
List of Figures

Figure 1. Representation of ERP components indicating latencies and major generators .................. 10
Figure 2. Formant frequency and duration information for the synthesized /iri/ to /ili/ continuum ....................................................................................................................................................... 23
Figure 3. Grand mean behavioral identification and discrimination performance ....................... 31
Figure 4. Individual subjects’ categorical perception index (left) and grand mean categorical index (right) ........................................................................................................................................... 33
Figure 5. Grand mean ACC waveforms taken from electrode site FCz ........................................ 34
Figure 6. Grand mean ACC (N1 & P2) latencies and amplitudes at FCz ..................................... 35
Figure 7. Grand mean ACC T-complex waveforms taken from electrode sites T7 and T8 ........ 37
Figure 8. Grand mean ACC T-Complex Ta and Tb latencies and amplitudes at T7 and T8 ...... 39
Figure 9. Grand mean ACC T-Complex Ta and Tb amplitudes at T7 and T8 ............................. 40
Figure 10. Grand mean difference waveforms for across-category pairs (tokens 4S7D and tokens 7S4D) at Fz ......................................................................................................................... 42
Figure 11. Grand mean MMN latencies (top) and amplitudes (bottom) for across-category pairs at Fz ........................................................................................................................................... 43
Figure 12. Grand mean difference waveforms for within-category pairs perceived as /r/ (tokens 1S4D and tokens 4S1D) at Fz ............................................................................................................. 46
Figure 13. Grand mean MMN latencies (top) and amplitudes (bottom) for within-category pairs perceived as /r/ at Fz ................................................................................................................................. 47
Figure 14. Grand mean MMN waveforms for within-category pairs perceived as /l/ (tokens 7S10D and tokens 10S7D) at Fz .................................................................................................................. 49
Figure 15. Grand mean MMN latencies (top) and amplitudes (bottom) for within-category pairs perceived as /l/ at Fz ................................................................. 50

Figure 16. Grand mean early and late MMN latencies (top) and amplitudes (bottom) at Fz....... 52

Figure 17. Grand mean waveforms for across- and within-category pairs at Fz....................... 55

Figure 18. Grand mean early and late P3a latencies (top) and amplitudes (bottom) at Fz........... 56
CHAPTER 1: Introduction
Knowledge of the mechanisms underlying speech perception is fundamental to understanding the process of language learning of both first and second languages. It is well known that native language influences the perception of speech sounds. For example, adults have difficulty perceiving speech sounds not present in their native language system (Cutler, 2012). Behavioral evaluation of speech perception has been the most dominant research. Even though behavioral measures provide information useful for the understanding of speech perception, this information is often insufficient for answering questions of the mechanisms underlying speech perception. The general purpose of this dissertation was to examine the effects of native language on the neurophysiologic and behavioral processing of non-native speech sounds in adults with different native language backgrounds.

This dissertation consists of four chapters. Chapter 1 provides an introduction, which includes a brief overview of the native language neural commitment model (section 1.1), a review of previous research on the effects of native language on the perception (section 1.2) and neural processing (section 1.3) of English /r/ and /l/, along with a rationale for the research (section 1.4), the purpose of the study and hypotheses (section 1.5). Chapter 2 provides a detailed description of the method. The results of the experiment are presented described in Chapter 3. Finally, Chapter 4 presents a discussion of the results.

1.1 Native Language Neural Commitment Model

Adults have difficulty perceiving speech sounds which are not present in their native language system. Kuhl and her colleagues proposed the concept of native language neural commitment to partially explain this phenomenon. The native language neural commitment
model presents the idea that initial language exposure generates changes in the neural circuitry in the brain, committing the brain’s neural networks to the native language (Kuhl, 2004; Kuhl et al., 2008). Therefore, these initially established speech patterns are hypothesized to influence the acquisition of new language patterns. When established neural networks influence second language learning so that the learning process is efficient, the networks facilitate speech processing. In contrast, when established neural networks do not support the efficient learning of a second language, it is thought that there is interference from these networks. There are several factors that may facilitate or interfere with learning a second language, including age of first exposure to the new language, duration of residence a country in which the second language is spoken (Cochrane, 1980; Yamada & Tohukura, 1992; Yamada, 1995; Fledge, Munro, & MacKay, 1995), and lexical factors such as word frequency (Broadbent, 1967; Luce, 1985; Jusczyk & Aslin, 1995; Landauer & Streeter, 1973) and lexical familiarity (Flege et al., 1996). Many related factors also contribute to the ease or difficulty in which a second language is acquired, including prior language experience, novelty of the sounds in the new language and the learner’s phonological inventory (Flege, 1995).

According to this model, neural commitment is malleable in infancy. This greater plasticity allows infants to easily learn and perceive sounds from other languages (Kuhl, 2004). In contrast, since neural networks are almost fully developed in adults, the established neural networks formed by the native language interfere with (or at least, do not facilitate) the second language learning process (e.g., Hisagi, Shafer, Strange, & Sussman, 2010; Iverson et al., 2003). Adults learning a second language experience greater difficulties than infants or children because the phonemic categories for their native language are highly established and new speech categories from the second language have a tendency to be assigned to the existing native
phonemic categories. Therefore, simple exposure to a second language may not easily change established neural networks of the native language in adults. An example that is directly relevant to this dissertation is that the perception of English liquids /r/ and /l/ has been shown to be challenging for native Korean and Japanese adults, as described in the next section.

1.2 Effects of native language on perception of English /r/ and /l/

The English language has two liquids, /r/ and /l/, and these sounds are phonemic for native English speakers. The first three formants are usually needed to perceive /r/ and /l/ (O'Connor, Gerstman, Liberman, Delattre, & Cooper, 1957). In American English, the most important cue for differentiating /r/ and /l/ is the third formant frequency (F3). In general, the /r/ has a low F3 value whereas the /l/ has a high F3 value. English liquids can occur in 5 different word positions: initial single (rock-lock), initial consonant cluster (cram-clam), medial (berry-belly), final single (fear-feel), and final consonant cluster (sort-salt). The transition cues for /r/ and /l/ are critical for perception of English /r/ and /l/ (Iverson & Kuhl, 1996). The steady-state and transition frequency cues for /r/ and /l/ vary according to the position of the phoneme in a word or syllable.

In the Japanese language, there is only one liquid phoneme, which is transcribed as /r/. Japanese /r/ is perceived as /d/ or /t/ by Americans (Miyawaki, 1973; Price, 1981; Vance, 1987). Therefore, Japanese /r/ is phonetically dissimilar to English /r/. Further, English /r/ and /l/ are not phonemic to native Japanese listeners. The Japanese liquid /r/ occurs only in word initial position. According to Bloch (1950), /l/ is an allophone of /r/ in Japanese, but rarely occurs.

In the Korean language, since there is only one liquid /l/ (Ahn, 1985; Kim-Renaud, 1975), the sounds /r/ and /l/ are not phonemic. However, the Korean /l/ has a range of allophones in
different word positions. The Korean language basically forbids word-initial liquids except in loan words. In contrast, in word-final position, the Korean liquid is consistently realized as /l/ (e.g. [mal] for horse, [al] for egg). In medial position, the Korean liquid /l/ is realized as an alveolar flap (e.g. [kuri] for bronze), transcribed as /ɾ/. When spectral differences in the production of this Korean liquid allophone, [ɾ] were analyzed, results indicated that the Korean alveolar flap [ɾ] had a generally low F3 that is similar to English /r/ (Kim, 2007). When a liquid in syllable final position is followed by another liquid in syllable initial position in Korean, the two liquids are realized as a double /ll/ (e.g. [allida] for inform). Therefore, it was proposed that the Korean language has a partial phonetic model available (intervocalic [ɾ]-[ll]) for perceiving English /r/ and /l/ sounds in medial position (Ingram & Park, 1998).

In an intervocalic position, English /r/ and /l/ have a different phonemic status for these three languages. For native English speakers, English /r/ and /l/ are phonemic. For native Korean speakers, they are partially phonemic. For native Japanese speakers, they are not phonemic. Table 1 provides a summary of these liquids for the English, Japanese, and Korean languages.
<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th>Japanese</th>
<th>Korean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liquid</td>
<td>/r/ and /l/</td>
<td>/r/ (perceived as /d/ or /t/ by Americans)</td>
<td>/l/</td>
</tr>
<tr>
<td>Word position</td>
<td>- Occur in 5 positions</td>
<td>- occurs in word initial position only</td>
<td>- occurs in medial and final positions only</td>
</tr>
<tr>
<td>Allophones</td>
<td>- /l/ can be an allophone of /r/ but very rare (Bloch, 1950)</td>
<td>- /l/ in medial position is an alveolar flap (e.g. [kori] for chain, [kuri] for bronze) or is a double /ll/ (e.g. [allida] for inform, [nollida] for tease)</td>
<td>- therefore, Koreans have a partial model ([r]-[ll]) available, but only in medial position (Ingram &amp; Park, 1998)</td>
</tr>
<tr>
<td>Acoustic cue</td>
<td>- lower F3 for /r/ - higher F3 for /l/</td>
<td>- Japanese production: F2 distribution</td>
<td>- Korean alveolar flap [ɾ]; lower F3</td>
</tr>
<tr>
<td>Phonemic status of English medial /r/-/l/ in each language</td>
<td>Phonemic</td>
<td>Non-phonemic</td>
<td>Partial phonemic model available in medial position</td>
</tr>
</tbody>
</table>

Table 1. A summary of English, Japanese, and Korean liquids

The perception of English liquids /r/ and /l/ is difficult for native Japanese and Korean listeners because these sounds are not phonemic in these languages (Ahn, 1985; Goto, 1971; Kim-Renaud, 1975; Miyawaki et al., 1975). Miyawaki et al. (1975) investigated the perception of English initial /ra/ and /la/ by Americans and Japanese. A continuum from perceived /ra/ to perceived /la/ by native American English listeners was synthesized by varying the third formant (F3) and presented using identification and discrimination tasks. Americans showed a clear categorical boundary for the /ra/-/la/ continuum, meaning they could accurately discriminate...
speech tokens from across these categories, but not within category. In contrast, Japanese
listeners showed near-chance percent correct identification of the sounds and discrimination
performance for the /r/ and /l/ stimuli was also near chance. Therefore, initial English /r/ and /l/
were perceived in a categorical manner by native English speakers but in a non-categorical
manner by native Japanese listeners.

English /r/ and /l/ are not phonemic in Korean, but the findings from previous behavioral
studies have shown that identification and discrimination of English /r/ and /l/ were better in
Korean than in Japanese listeners (An et al., 2013; Hazan et al., 2006; Ingram & Park, 1998;
Komaki & Choi, 1999; Komaki, Tajima, Akahane-Yamada, & Choi, 2000). The perceptual
identification of /r/ and /l/ by native Japanese and Korean listeners, however, varied as a function
of word position (Dissosway-Huff, Port, & Pisoni, 1982; Goto, 1971; Henly & Sheldon, 1986;
Ingram & Park, 1998; Mochizuki, 1981; Sheldon & Strange, 1982). Japanese had greater
perceptual identification difficulty for the initial consonant cluster position and least difficulty in
final and initial position in the word (Mochizuki, 1981; Sheldon & Strange, 1982). Korean
listeners had greater identification difficulty in word initial position compared to either medial or
consonant cluster positions (Ingram & Park, 1998). Importantly, Koreans achieved highest
identification performance in word medial position and poorest performance in word initial
position. Similar identification patterns were also obtained when nonsense words were used
(Hazan et al., 2006). As reviewed previously, Japanese liquid /ɾ/ occurs only in syllable initial
position, so Japanese listeners identify the initial liquids more accurately. In contrast, in the
Korean phonological system, word-initial liquids are generally forbidden, leading Korean
listeners to make more errors in word initial position. However, the partial model (intervocalic
[r]-[ll]) present in Korean may explain the better identification and discrimination of /r/ and /l/
for Koreans compared to Japanese, particularly when they are presented in syllable medial position.

These behavioral identification and discrimination patterns described for English medial /r/ and /l/ by Koreans relative to Japanese might be predicted by Best’s perceptual assimilation model (Best, 1995; Best & Tyler, 2007). Koreans have a good exemplar of /l/ but a poorer exemplar of /r/ in medial position, giving potential for categorization for the contrast. In the case of Japanese listeners, however, English /r/ and /l/ are non-phonemic. It seems that Japanese listeners assimilate English /r/ and /l/ into a bad exemplar of the Japanese /r/ category. If this is the case, according to the model, it represents a single category assimilation, resulting in a poor categorization for the contrast, although a recent study suggested that native Japanese listeners differentially assimilate English /l/ sounds (not both English /r/ and /l/ sounds) into the Japanese /r/ category (Hattori & Iverson, 2009). The perceptual assimilation model does not fit perfectly to categorical perception; however, based on the model that clear categorical perception would be observed in the American group, one might predict categorical or categorical-like perception in the Korean group, and non-categorical perception in the Japanese group. However, limitations of the perceptual assimilation model limit these predictions because it does not provide detailed information to predict patterns of categorical perception (e.g., the boundary itself, boundary shift, boundary slope, discrimination peak or accuracy, etc.) corresponding to relevant native language as well as underlying neural mechanisms.

Since the perception of English /r/ and /l/ in medial position reveals a distinction between native English, Korean and Japanese groups, one might assume that there would also be a different pattern of neurophysiologic responses for each language group based on the native language neural commitment model (Kuhl et al., 2004). For instance, this model would predict
that since English has both /r/ and /l/ in the language, Americans would have distinct neural networks for these sounds. Japanese would not develop distinct neural networks for /r/ and /l/ and their perception of English /r/ and /l/ would be limited since the Japanese language has neither /r/ nor /l/ categories in the language system. Koreans might form distinct neural networks corresponding to their partial model in medial position. However, it is unclear how similar or different the neural networks would be between language groups because of limitations in the information provided by behavioral measures.

Therefore, a combination of neurophysiologic measures along with behavioral measures may provide more specific and detailed information about the mechanisms underlying the processing of /r/ and /l/ in these language groups. Behavioral measures provide information about perception, whereas electrophysiological measures such as event-related potentials (ERP) provide information about how the brain processes information while providing good temporal resolution. Since ERPs provide information regarding the timing and sequence of information processing, they are sensitive to the overlapping stages of processing leading to perception. Behavioral measures represent the endpoint of this processing continuum. The combination of ERP and behavioral measures of speech processing may improve our understanding of the mechanisms underlying speech perception. Furthermore, evidence from different ERP components might be used to extend or refine models of the perception of non-native sounds, such as the native language neural commitment model, in terms of the nature of the neural networks. Also, it might elucidate how native language influences brain responses to nonnative speech contrasts in terms of the relative contribution of acoustic- and phonemic/phonologic processing. Figure 1 provides an overview of representation of event-related potential components with latencies and major generators that was examined in this dissertation, along
with a description of the P1-N1-P2 elicited by sound, using onset latency as a point of reference.

The Figure 1 is a simplified example of various ERP components relevant to this dissertation study arranged by their time course as well as major generators between input sound and perception.

![Figure 1. Representation of ERP components indicating latencies and major generators](image)

Although the figure was constructed loosely according to time course, it does not necessarily signify serial processing. Instead, each neural component indexes different aspects of information processing: encoding of sound onset in auditory cortex for P1-N1-P2, encoding of acoustic change within sound for the acoustic change complex (ACC), pre-attentive discrimination for mismatch negativity (MMN), and involuntary attention shift to stimulus change for P3a. Note also that generators of these responses differ somewhat (for review, see Steinschneider & Dunn, 2002). The examination of multiple ERP components is important, because the combination of measures may provide additional information about similarities/differences in neural processing patterns across the time-course of processing underlying perception in different language groups. Research on the application of these ERP
components to examine the influence of native language on processing of non-native speech sounds by different language groups will be reviewed in the next section.

1.3 Effects of native language on neural processing of speech sounds

The influence of native language on obligatory ERPs to speech sounds examined at midline-central electrode sites is not clear. These midline-central responses are dominated by contributions from primary auditory cortex (Ponton et al., 2002). At midline-central electrode sites, some reported no significant differences in the amplitudes, latencies or morphology of the P1-N1-P2 complex as a function of native language (Elangovan & Stuart, 2011; Sharma & Dorman, 2000). In contrast, others have reported significant effects of native language on these ERPs. For example, it has been demonstrated that native language influenced N1 latency, N1 amplitude, and P2 amplitude during an active identification task (Horev, Most, & Pratt, 2007) and N1m (magnetoencephalographic counterpart to N1) morphology as well as hemispheric laterality of N1m (Zhang et al., 2005). Relevant to this dissertation, different N1m morphological patterns were obtained for /ra/ and /la/ for Japanese listeners compared to American English listeners. Japanese listeners had a double-peaked N1m over both the right and left hemispheres whereas English listeners had only a single N1m over the left hemisphere (Zhang et al., 2005).

The T-complex over lateral-temporal sites is dominated by contributions from secondary cortex generators (Ponton et al., 2002). The T-complex appears to index some aspects of language ability (Mason & Mellor, 1984; Shafer, Schwartz, & Martin, 2011; Wagner, Shafer, Martin, & Steinschneider, 2013). For example, a majority (73%) of children with specific language disorder (SLI) showed poor Ta amplitude and T-complex waveform morphology in response to speech (Shafer et al., 2011). The effects of native language on the processing of
speech in native Polish, and native English listeners were examined using nonsense words (e.g., /pt/ vs. /pət/ and /st/ vs. /sət/) (Wagner et al., 2013). The /pt/ at syllable onset is “illegal” in the English language but not in the Polish language. There was significant difference in the T-complex evoked by /pt/ condition as a function of language group. Responses to /pt/ condition were significantly more negative for the Polish listeners than native English listeners, supporting the possibility that the lateral-temporal T-complex might potentially provide a better index of the influence of native language and some aspects of language processing than the midline-central responses.

It should be noted that much of the research on P1-N1-P2 evoked by speech utilized vowel-consonant stimuli. This dissertation utilized vowel-liquid-vowel stimuli and as a result, the response of interest is the response to the acoustic change from vowel to liquid. This change-related response is known as the acoustic change complex (ACC) (Martin and Boothroyd, 1999; 2000).

Mismatch negativity (MMN) is thought to index automatic pre-attentive speech discrimination and peaks at latencies of 100-250ms after deviant stimulus onset (Näätänen, 1992) although the precise role of MMN in processing is still under debate (May & Tiitinen, 2010; Sussman, 2007; Sussman, Winkler, & Wang, 2003). MMN is largest near frontal electrode sites (e.g. Fz, FCz) and is mainly generated in the primary auditory cortex with some involvement of the secondary auditory cortex and possibly frontal cortex (Alain et al., 1998; Doller et al., 2003; Rinne, Alho, Ilmoniemi, Virtanen, & Naatanen, 2000; Steinschneider & Dunn, 2002). There has been some inconsistency in the MMN literature regarding whether MMN primarily indexes acoustic processing or whether it also indexes linguistic processing. In other words, if MMN is sensitive to native language experience, then it is not solely driven by
acoustic aspects of sound. Much of this controversy may reflect the paradigm used to elicit MMN. Results of some MMN studies were consistent with acoustic processing because as acoustic deviance increased, the MMN was earlier and more robust (Aaltonen, Niemi, Nyrke, & Tuhkanen, 1987; Sams, Aulanko, Aaltonen, & Näätänen, 1990). Furthermore, MMN responses from within-category pairs were not significantly different from across-category pairs (Dalebout & Stack, 1999; Sharma, Kraus, McGee, Carrell, & Nicol, 1993). Interestingly, these MMN studies used speech stimuli in which the deviant differed in terms of spectral information (such as F2 or F3). Other MMN studies, however, indicated that MMN can reflect phonological processing. These studies used stimuli changing in the temporal domain (VOT). Sharma and Dorman (1999) demonstrated that MMN at Fz was larger for across-category pairs (/da/-/ta/) compared to within-category pairs (/ta/-/ta/), which was consistent with behavioral identification results. It is possible that temporal information is processed differently than spectral information when investigating categorical perception at the level of processing tapped by MMN.

Over time, it has become apparent that MMN reflects language experience. Several cross-language studies support the claim that MMN can reflect language experience by showing that MMN responses to native-phonemes were more robust than those to non-native phonemes (Dehaene-Lambertz, 1997; Näätänen et al., 1997; Winkler et al., 1999; Näätänen et al., 2001 for review). For example, Näätänen and his colleagues (1997) investigated the impact of language experience on non-native vowel perception using MMN. Subjects were Finns and Estonians, who have very similar vowel structures except for the Estonian vowel /õ/. Four vowels were selected as stimuli, /ɛ/, /ö/, /o/, and /õ/. The acoustic difference between the four vowels is only in the second-formant (F2) frequency while the other formant frequencies were kept constant. For the MMN, the vowel /ɛ/ served as the standard stimulus and the other vowels (/ö/, /o/, and
/õ/) were deviants. If the MMN is determined only by acoustics, the two language groups should have similar MMN responses to the Estonians vowel /õ/. For Estonians, the MMN amplitude increased as a function of the amount of F2 deviance from the standard /ɛ/. In contrast, the deviant /ö/ elicited a larger MMN than the deviant /õ/ in Finns even though the /õ/ had acoustically bigger deviance from the standard /ɛ/ than /ö/. The authors concluded that these results support the claim that the MMN reflects language experience and is influenced by native language phonology.

The P3a follows the MMN and is also elicited by an oddball paradigm. P3a is a frontal positivity occurring around 250-280 ms post-deviant stimulus. The P3a is different from P3b, which is associated with active attention to deviant stimuli (Snyder & Hillyard, 1976). The P3a is believed to reflect orienting to or involuntary attention shifts to stimulus changes or novel stimuli (Halgren et al., 1995; Polich, 2003). The generators of P3a seem to be relatively diffuse and include the hippocampus and pre-frontal cortex (Knight, 1984; 1996). Novitski, Tervaniemi, Huotilainen, and Naatanen (2004) reported that the P3a amplitude had a correlation with performance of working memory tasks, indicating an attentional orienting response functions as a process intervening or facilitating memory representations. Furthermore, P3a amplitude was enhanced in children after 2 months of learning a second language, suggesting that these attentional shifts increase over time, at least in children (Shestakova et al., 2003). In addition, P3a latency was correlated with working memory performance (Light, Swerdlow, & Braff, 2007). These results indicate that P3a might serve as a tool for encoding and retrieval of information related to language learning or memory storage.
1.4 Preliminary data and rationale for the current study

The effects of native language on the perception and encoding of English /r/ and /l/ by native American, Korean, and Japanese listeners were examined in a preliminary study (An, Martin, & Long, 2013). Using the same intervocalic /iri/-/ili/ continuum that will be used for the dissertation study (described later), the acoustic change complex (ACC) responses were measured at midline-central and lateral-temporal electrode sites along with behavioral identification and discrimination. For the preliminary study, the acoustic change complex was obtained by presenting five tokens from the continuum 400 times each using a 1200 ms offset-to-onset interstimulus interval (which is different from the current proposal).

Behavioral data indicated that English medial /r/ and /l/ were perceived in a categorical manner by Americans, in a categorical-like manner by Koreans and in a non-categorical manner by Japanese. ACC P1-N1-P2 responses did not differ significantly between language groups at the midline-central electrode with largest response amplitude (FCz), suggesting little effect of native language on the primary cortical encoding of these sounds. In contrast, the ACC T-complex measured over lateral-temporal electrode sites (T7 and T8) differed significantly between language groups suggesting that native language influences secondary cortical processing of these sounds. Differences between groups included Ta latency and Tb latency. Regional dipole source analysis revealed that the Japanese group had greater variability in individual source locations and more superior source locations compared to the other language groups. These results inspired the current dissertation and the goal to investigate multiple evoked potentials elicited using one electrophysiological method in order to provide a better understanding regarding the effects of native language on neurophysiologic processing of non-native speech sounds over the time course leading to perception.
1.5 Purpose of the current study and hypotheses

Different evoked potentials index different aspects of speech processing and can be used to provide useful information about the stages of neural processing leading to perception in different language groups. Most previous research has used behavioral measures only or else a single ERP component to investigate the effects of native language on the processing of speech. However, the use of multiple evoked potentials elicited using one electrophysiological paradigm combined with behavioral measures are expected to allow a more comprehensive examination of the impact of native language on the perception of speech sounds not found in the native language. In this study, a paradigm was used that elicits a range of evoked potentials that tap the obligatory encoding, pre-attentive discrimination, and orienting/attention shifting to speech along with behavioral measures.

The overall purpose of this study was to compare the effects of native language on the neurophysiological processing of English intervocalic /r/ and /l/ by native Korean, Japanese and American English listeners using the multiple ERP components (ACC, MMN, & P3a) combined with behavioral identification and discrimination. Three specific aims were investigated. The first aim was to examine the effects of native language on the perceptual identification and discrimination of English intervocalic /r/ and /l/. It was hypothesized that having a partial model (e.g., intervocalic [ɾ]-[ll]) in the native language would facilitate the ability to make the distinction and would aid in speech perception. Thus, it was hypothesized that clear categorical perception would be obtained in the American English group, categorical-like perception would be obtained in the Korean group due to their partial model, and non-categorical perception would be obtained in the Japanese group.
The second aim was to examine the effects of native language on the neurophysiologic encoding of English intervocalic /r/ and /l/. In general, it was hypothesized that ACC would be earlier and more robust for the American participants, responses would be latest and smallest in the Japanese group with the values for the Korean group falling in-between, similar to the preliminary study. More specifically, it was hypothesized that the midline-central ACC response (P1-N1-P2), which is dominated by primary auditory cortex generators, would show no significant effects of native language group, whereas the lateral-temporal ACC (Na-Ta-Tb) examined at T7/T8, which mainly reflects secondary auditory cortex generators, would show significant effects of native language. That is, the ACC P1-N1-P2 responses in the American group would be similar to those obtained in the Japanese and Korean groups. In contrast, the ACC T-complex would have largest amplitude and shortest latency for the American group, intermediate responses for the Korean group, and small/late responses for the Japanese group.

The third aim was to examine the effects of native language on the pre-attentive discrimination and involuntary attention shift of English intervocalic /r/ and /l/. It was hypothesized that the MMN would reflect effects of native language (at least for the across-category stimulus condition), with a robust, early MMN in the American group, an intermediate, later MMN in the Korean group, and a small/absent MMN in the Japanese group since previous MMN studies related to cross-language speech perception showed the benefit of native language for non-native sound perception. It is more difficult to hypothesize the effects of native language on P3a because of a lack of relevant cross-language literature. Since the P3a is dependent on discrimination, P3a would not be expected if there is no MMN. Nevertheless, if novelty increases orienting, then P3a would be larger and earlier in Japanese than in Koreans, and it
would be smaller and later in Americans. In contrast, if familiarity increases orienting, the P3a would show the opposite pattern.
CHAPTER 2: Method
2.1 Participants

Participants were adults from three different language groups: nine native speakers of American English, nine native speakers of Japanese, and nine native speakers of Korean. All participants passed a hearing screening (thresholds better than 25 dB HL from 250 to 8000 Hz bilaterally, American Speech-Language-Hearing Association, 1997) with normal tympanograms (admittance curve at 226 Hz with a single peak between +/- 50 daPa) as well as present ipsilateral acoustic reflexes (90 dB HL at 1000 Hz). They also had no history of neurological or learning problems. Personal information about Korean and Japanese participants such as age, gender, length of residence in English-speaking countries, and age of arrival in English speaking countries were obtained by the Language Experience and proficiency Questionnaire (LEAP-Q) (Marian, Blumenfeld, & Kaushanskaya, 2007) and Appendix 1 shows individual background information for Koreans and Japanese collected by LEAP-Q. Background information for participants is shown Table 2. There was no significant difference in age and proficiency variables between Korean and Japanese participants, as reflected by t-tests for independent samples (p > 0.05).
<table>
<thead>
<tr>
<th></th>
<th>American</th>
<th>Japanese</th>
<th>Korean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>9 (5 females, 4 male)</td>
<td>9 (8 females, 1 male)</td>
<td>9 (8 females, 1 male)</td>
</tr>
<tr>
<td>Mean age</td>
<td>26.6 yrs (5.8 yrs)</td>
<td>35.2 yrs (6.5 yrs)</td>
<td>30.2 yrs (4.7 yrs)</td>
</tr>
<tr>
<td>Mean length of residence</td>
<td>4.4 yrs (2 yr)</td>
<td>2.6 yrs (2.6yrs)</td>
<td></td>
</tr>
<tr>
<td>Mean age of arrival</td>
<td>30.6 yrs (7.1 yrs)</td>
<td></td>
<td>27.7 yrs (3.6 yrs)</td>
</tr>
<tr>
<td>Current exposure to English (%)</td>
<td>36 % (20.9)</td>
<td>47.7 % (25.4)</td>
<td></td>
</tr>
<tr>
<td>Mean English speaking (self-rated proficiency)</td>
<td>4.8 (3.0)</td>
<td>4.3 (2.6)</td>
<td></td>
</tr>
<tr>
<td>Mean English listening (self-rated proficiency)</td>
<td>5.3 (2.9)</td>
<td>4.9 (2.9)</td>
<td></td>
</tr>
<tr>
<td>Mean English reading (self-rated proficiency)</td>
<td>4.8 (2.4)</td>
<td>4.1 (2.4)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Background information for the American, Japanese, and Korean participants including means and standard deviations in parentheses (Note: The LEAP-Q is scored on a 10 point self-rated proficiency scale, with non-proficiency scored 0 and perfect proficiency scored as 10).

### 2.2 Stimuli

A series of 11 vowel-consonant-vowel (VCV) syllables was generated by changing the third formant frequency in acoustically equal 130 Hz step using HLsyn speech synthesis software resulting in sounds that varied from perceived /iri/ to /ili/ by native American English speakers. A schematic diagram of the stimulus is shown in Figure 2. The stimuli consisted of several portions: an initial 300 ms steady-state vowel portion, a first transition from vowel to

---

1 Before conducting the pilot experiment, the two endpoints, token1 and token11, were blindly and randomly presented to three native American English speakers five times in order to test stimulus quality. The individuals reported what was heard. All three subjects reported /iri/ or /ili/ appropriately for the endpoint stimuli. Then the subjects were asked about “goodness” of /iri/ and /ili/ and the responses from these pilot subjects included “natural”, “realistic”, and “identifiable”. Then three different native American English speakers were recruited and a 2 alternative forced choice identification test was conducted. The results showed a clear categorical boundary with the endpoints yielding high identification accuracy.
consonant at 300 ms, a steady-state consonant portion from 345 ms to 555 ms, a second transition from consonant to vowel at 555 ms, and a 300 ms steady-state vowel portion, for a total duration of 900 ms. The 11 speech tokens used for the intervocalic liquid syllable series differed from one another only in the F3 (transition and steady-state consonant portions). The transition of F3 was varied in steps of 130 Hz from 1500 Hz for Token 1 (perceived as /iri/) to 2800 Hz at Token 11 (perceived as /ili/). Most of these acoustic parameters were adapted from McGovern and Strange (1977). The transition durations were based on Lisker (1957), who provided data for intervocalic liquids.

As seen in Figure 2, the speech tokens had a constant F0 of 120 Hz for the first 600 ms of the syllable and then a gradually falling F0 contour after that from 120 to 100 Hz. Frequency values for F1, F2, and F4 were identical for all tokens. The first formant increased from 300 Hz to 400 Hz during the first 22.5 ms of the first transition portion and then declined from 400 Hz to 300 Hz during the last 22.5 ms of the second transition portion. The second formant declined linearly from an initial steady-state vowel value of 2100 Hz to an initial steady-state consonant value of 1100 Hz followed by mirror image values for the second transition and vowel portion. The forth formant frequency was kept constant at 3500 Hz throughout the entire syllable. Linear windowing was applied to the first and last 20 ms of each stimulus. The stimuli were then equalized for RMS amplitude.
2.3 Procedure

Behavioral and electrophysiological measures were both obtained using a Neuroscan system with the order counterbalanced across participants in all groups. Testing was conducted in a double-walled sound attenuated booth. The stimuli were presented binaurally at 70 dB SPL using EAR 3A insert earphones. The procedures were identical for the three groups of participants.

1. Behavioral Measures

a. Familiarization

Participants heard the ordered series from token 1 to token 11 without feedback. They were then told that the stimuli were several examples of the English syllables /iri/ and /ili/. The stimuli were then randomly presented. Finally, the two endpoints of the continuum, token 1 and...
token11 were presented five times.

b. Identification task

The participants were told to listen to each syllable and then press a button on a response pad to indicate whether the syllable contained an “r” (press “1”) or an “l” (press “4”). They were instructed to respond to every speech token with one of the two choices, and to guess if needed. The 11 speech tokens were presented in random order with a 1200 ms offset to onset ISI. Each token was presented 10 times, giving a total of 110 stimulus presentations (11 stimulus*10 times). Two different randomizations were presented, yielding a total of 20 judgments for each speech token. Participants were encouraged to press the response button with equal weight on the speed and accuracy of their button presses. No feedback was given.

c. Discrimination task

A 3-interval alternative forced choice oddity test was used. Six pairs were selected from the speech tokens such that each pair (AB) differs by three steps (i.e. token 1-4, 4-7, & 7-10). These stimuli were chosen because pilot data indicated that the boundary of /r/ and /l/ for American and Korean participants occurred between token 4 and token 7. For each pair, trials were constructed by duplicating one token in the pair: all six permutations (AAB, ABA, BAA, ABB, BAB, & BBA) for each of the three pairs were used. Thus, one oddity test comprised 18 triads (the three pairs * all the six permutations). The triads were arranged in random order with a 1-sec offset-to-onset ISI and a 3-sec offset-to-onset intertriad interval (ITI), which served as subject response time. Two different randomizations were presented twice, yielding a total of 24 stimulus presentations for each pair. They were informed to respond to every triplet and guess when necessary, by pressing a button on a response pad to indicate whether the odd stimulus on each trial was first (press “1”), second (press “2”), or third sound (press “3”). They waited until
all 3 stimuli had been presented, then made a decision. Participants were instructed to press the buttons giving equal weight and speed versus accuracy of their button presses. No feedback was provided.

2. **Electrophysiological Measures**

An oddball paradigm was used with stimuli presented as standards on 80% and stimuli presented as deviants on 20% of the trials. Speech sounds were presented using an offset-to-onset inter-stimulus intervals (ISI) of 1100 ms\(^2\). Participants watched a muted video during testing. A previous study indicated that the phonetic boundary for these stimuli occurred between token 4 and token 7 in American speakers (An, Martin, & Long, 2013). Therefore, two across-category (token pairs 4S7D & token pairs 7S4D)\(^3\) and four within-category conditions (token pairs 1S4D, 4S1D, 7S10D, &10S7D) were selected as stimulus sets. That is, six stimulus pairs, which included the reversed order of presentation, were presented to participants. The six pairs were presented in counter-balanced order. Each pair had a total number of 750 trials consisting of 600 trials for standards and 150 trials for deviants. Six blocks were presented and each block included 750 trials. Evoked potentials were recorded using a Neuroscan system and a 64-channel electrode cap. Electrode impedances were maintained below 5000 Ohms. The ground was placed halfway between FPz and Fz. The ongoing electroencephalographic activity recorded at each

---

\(^2\) The ISI of 1100 ms was selected after piloting different ISIs including 500 ms, 800 ms, 900 ms, & 1100 ms. While a rapid ISI is usually preferred for MMN to strengthen the sensory memory of the standard, the onset and ACC responses in response to the standard (/iri/ or /ili/) were not detectable in the more rapid conditions. The 1100 ms condition was the shortest ISI that generated clear onset responses to these stimuli. Also, previous research has indicated that the use of a long ISI enhances perception differences across language groups (Werker and Logan, 1985) and that language-specific experience may be needed to generate MMN for inter-onset ISIs of 2 seconds or more (Yu, 2013).

\(^3\) Tokens 4S7D indicates that token 4 functions as a standard and token 7 as a deviant; tokens 7S4D indicates that token 7 functions as a standard and token 4 as a deviant.
electrode was digitized (1000 Hz A/D rate), amplified (gain=1000), and band-pass filtered (0.15 – 100 Hz; 6 dB/octave). Data were stored for offline analysis. After recording, the data were processed offline by applying an eyeblink reduction algorithm (Semlitsch, Anderer, Schuster, & Presslich, 1986), baseline correction (-100 to 1123 ms), digital filtering (1 to 30 Hz, 12 dB/octave, zero phase-shift filter to avoid distortion), artifact rejection (± 100 µV) and averaging. The averaged waveforms were referred to an average reference. Grand mean waveforms were computed for each subject in each condition and group grand mean waveforms were created for each language group and each condition.

2.4 Data Analyses

1. Behavioral Data

a. Identification

The percentage that each of the eleven speech tokens was identified as /r/ versus /l/ was calculated for each individual and then pooled over participants in each language group. The boundary location of /r/ and /l/ was the point at which tokens were identified as /r/ with 50% accuracy (chance level). In addition, a Categorical Perception Index

$$f = \sqrt{\sum_{i=1}^{11} (\alpha_i - b_i)^2}$$

was calculated to quantify the slope of the identification function (Paul, Bott, Wienbruch, & Elbert, 2006). Here, \(\alpha_i\) represents the number of responses for /iri/ and \(b_i\) the number of responses for /ili/ to the \(i\)-th stimulus. A high CPI implied that the participants’ categorical judgments were sharp and clear.
b. Discrimination

Percent correct discrimination (%) was calculated for each stimulus pair in each individual and then pooled over participants in each language group (chance-level=33%). The discrimination peak was the stimulus pair with the highest percent correct.

2. Electrophysiological Data

a. P1-N1-P2/T-complex

The latencies and amplitudes of the midline-central (P1-N1-P2) and lateral-temporal (T-complex, Na-Ta-Tb) responses to /iri/ and /ili/ stimuli were measured at midline-central (FCz) and at lateral-temporal (T7 & T8) electrode sites where P1-N1-P2 and T-Complex, respectively are large in amplitude. This permits comparison of neural responses dominated by both primary and secondary cortex generators. Since this study was concerned with the response elicited by the medial /r/ or /l/, all latencies reported were referenced to the onset of /r/ or /l/. Amplitude measures were taken at the point of largest amplitude and latency measures were obtained at the center of the peak. Amplitudes and latencies of each peak were analyzed using mixed model analyses of variance (ANOVA) as a function of language group and token (or token in each hemisphere). Main effects and interactions were considered significant if $p < 0.05$. Least significant difference (LSD) post-hoc measures were conducted only when significant main effects or interactions were obtained.

b. MMN & P3a as indices of sound discrimination

The midline-frontal electrode site at which MMN and P3a were maximal (Fz) was used for analysis. Difference waveforms were created by subtracting the response to standards from the response to deviants for each condition and for each participant. Paired t-tests were used to
determine if the MMN in the difference waveforms was statistically significant (significantly different from zero). Pilot data examined in American participants indicated the presence of two different time-frames for the MMN (early and late MMN). The early MMN occurred between 400 and 650 ms and the late MMN occurred between 655 and 905 ms. Therefore, MMN amplitudes were taken as the largest point in each response window. MMN latencies were taken at the peak or in the center in the case of multiple peaks in each response window. The MMN was considered to be present when it was largest in amplitude at fronto-central electrode sites and inverted in polarity at the mastoids. The MMN was considered as not detectable when topography was incorrect and also if the response to deviants was more positive than the response to standards. If MMN responses were present in fewer than 4 participants in a language group, the condition was omitted from the analyses. Amplitudes and latencies of each peak were analyzed using mixed model analyses of variance (ANOVA) as a function of language group and time and stimulus condition. Main effects and interactions from the mixed model analyses of variance (ANOVA) were considered significant if \( p < 0.05 \). LSD post hoc measures were conducted only when significant main effects or interactions were obtained.

Similar to MMN, there was an early and a late P3a. The early P3a occurred between 600 and 700 ms and the late P3a occurred between 900 and 1000 ms. P3a amplitudes were taken at the point of largest amplitude and P3a latency measures were obtained in the center of the peak in each response window. If P3a responses were present in less than 4 participants in each language group, the condition was regarded as “absent” for that language group and omitted from the analyses. Amplitudes and latencies of each peak were analyzed using mixed model analyses of variance (ANOVA) as a function of language group and time and stimulus condition. Main effects and interactions from the mixed model analyses of variance (ANOVA) were
considered significant if $p < 0.05$. LSD post hoc measures were conducted only when significant main effects or interactions were obtained.
CHAPTER 3: Results
3.1 Behavioral Results

Grand mean behavioral identification and discrimination performance for the three language groups is shown in Figure 3.

Figure 3. Grand mean behavioral identification and discrimination performance for syllable-medial tokens for the three language groups (note: dashed lines at 50% on the identification graph and at 33% on the discrimination graph indicate chance levels respectively, error bars show ±1 SE)
American participants consistently divided the stimuli into two distinct phoneme categories revealing clear categorical perception. The boundary between phonemes /r/ and /l/ fell between token 4 (F3=1890 Hz) and token 6 (F3=2150 Hz). American participants discriminated stimuli drawn from across different phoneme categories most accurately whereas they could not reliably distinguish stimuli identified as the same phoneme. The discrimination peak for American participants reached 91% correct and the discrimination peak was located at comparison pair token 4 and 7, which straddles the phoneme boundary the identification data.

Korean participants were able to identify the stimuli and showed categorical-like perception. The boundary between categories was shifted slightly toward tokens 5 (F3=2020 Hz) and 7 (F3=2280 Hz), toward the /l/ category. Perception was shifted by an F3 of 130 Hz toward the /l/ category compared to Americans. The discrimination peak for Korean participants reached 92% correct and was located between token 4 and token 7. Japanese participants showed two identification categories, albeit with no clear, sharp category boundary. Discrimination performance of the Japanese, however, suggested non-categorical perception (near chance of 33%) for the Japanese. Two Japanese participants with a relatively early age of arrival to an English speaking country (i.e., S2 and S4 in Japanese participants, see Appendix 1 for details) showed better identification and discrimination, although their performance is still poorer than the average performance of Koreans.

The categorical perception index (CPI) for each participant and each language group is shown in Figure 4. There was greater variability as well as lower CPI scores for Japanese participants compared to the other language groups. As seen in Figure 4, the group averaged CPI was 60.0 for the Americans 55.2 for the Koreans, and only 38.8 for the Japanese. Supporting these observations, a one-way analyses of variance (ANOVA) on CPI revealed a significant main
effect of language group \([F(2, 24)=15.1580, p < 0.001]\). LSD post hoc testing revealed that the CPI in Americans was significantly steeper than that in Japanese \((p < 0.001)\). Furthermore, CPI was also significantly steeper for Koreans than for Japanese \((p < 0.001)\), but there was no significant difference between Americans and Koreans \((p = 0.25)\). In other words, the sharpest identification slope was obtained in the Americans and Koreans and the least steep slope was obtained from the Japanese. Indicating that American and Korean listeners had the most categorical performance and Japanese listeners had the least categorical performance.

![Individual CPI and Grand mean CPI](image)

Figure 4. Individual subjects’ categorical perception index (left) and grand mean categorical index (right) for the three language groups (Note: Error bars show ±1 SE and *** indicates < .001)
3.2 Acoustic Change Complex (ACC) N1-P2

The grand mean waveforms elicited by the four different speech tokens at the midline-central site (FCz) for each language group are presented in Figure 5. The ACC was clearly evoked in all three language groups, but was smaller in amplitude compared to the onset response, as expected. Subtle differences in the response across tokens were apparent.

Figure 5. Grand mean ACC waveforms for each language group from electrode site FCz. Responses to four tokens (token1, 4, 7, & 10) from the /iri/ to /ili/ continuum are overlaid. Onset P1-N1-P2 responses are indicated by the box and ACC responses by the circles. The ACC occurred at approximately 50-200 ms after the onset of /r/ or /l/ (~350-500 ms re: stimulus onset).
Grand mean ACC N1 and P2 latencies and amplitudes for each token and each language group are shown in Figure 6. There are no apparent differences in N1 or P2 latency as a function of language group or token. While some differences in N1 and P2 amplitude are apparent in Figure 6, the large variability meant that the mixed model analyses of variance as a function of group and token revealed no significant main effects of language group or token, and no significant interactions between group and token.

Figure 6. Grand mean ACC (N1 & P2) latencies and amplitudes at FCz (Note: Error bars show ±1 SE)
3.3 Acoustic Change Complex (ACC) T-complex

Grand mean waveforms elicited by the four speech tokens at lateral-temporal sites (T7 & T8) are shown in Figure 7. The T-complex (Ta-Tb) from each of the three language groups was clearly elicited but small in amplitude compared to the onset T-complex. It occurred at approximately 80-190 ms after the onset of /r/ or /l/ (~380-490 ms re: stimulus onset). Note that the T-complex is most distinct in Japanese listeners (see the bottom row in Figure 7).
Figure 7. Grand mean ACC T-complex waveforms taken from electrode site, T7 and T8 overlaid for the four tokens (token 1, 4, 7, & 10) from the /iri/ to /ili/ continuum are shown in each language group. Onset T-complex (Na-Ta-Tb) responses are indicated by the box and ACC T-complex responses are indicated by the circle.
1. Latencies

Grand mean ACC T-complex Ta and Tb latencies obtained using electrodes, T7 and T8, are shown in Figure 8. As seen in Figure 8, Ta latencies for Japanese were longer on average at both T7 and T8 than those of Americans and Koreans for all the tokens. A mixed model analysis of variance on Ta latency revealed that there was a significant main effect of language group [F(2, 24) = 21.05, p < 0.001], but there were no significant main effects of hemisphere [F(7, 168) = 0.40, p = 0.90] and no interaction of hemisphere by language group [F(14, 168) = 0.74, p = 0.73]. LSD post-hoc analysis indicated that Ta latencies of Japanese (mean = 424 ms) were significantly longer in both hemispheres compared to Americans (mean = 399 ms) (p < 0.001) and Koreans (mean = 400 ms) (p< 0.001), whose latencies were not significantly different (p = 0.96).

There were no apparent differences are present as a function of language group or token for Tb latencies at T7. However, Tb latencies at T8, were on average longest for Koreans while Japanese had the shortest latencies (Figure 8). A mixed model analysis of variance on Tb latency data revealed that there were significant main effects of language group [F(2, 24) = 11.3, p < 0.001], hemisphere [F(7, 168) = 13.9, p < 0.001], and a significant interaction between group and hemisphere [F(14, 168) = 7.0, p < 0.001]. LSD post-hoc analysis indicated two clear patterns. First, Tb latencies at T7 (left hemisphere) did not differ significantly between language groups (p = 0.94), but Tb latencies at T8 showed significant differences between language groups (p < 0.001). At T8, the Tb latencies were longest for Koreans (mean = 498 ms), whereas Japanese had the shortest latencies (mean = 470 ms) (p < 0.001), with intermediate Tb latencies for Americans (mean = 482 ms) (p < 0.001). Second, Tb latencies were significantly shorter for
Americans and Koreans at T7 than at T8 (p < 0.05 for Americans, p < 0.001 for Koreans), whereas Japanese did not show a significant effect of hemisphere (p > 0.05) for Tb latency.

2. Amplitudes

Grand mean ACC T-complex Ta and Tb amplitudes at T7 and T8 are shown in Figure 9. As seen in Figure 9, both Ta and Tb amplitudes showed large variability across tokens and Tb amplitudes at T7 were smaller than at T8 in general. Mixed model analysis of variance on Ta amplitudes revealed that there were no significant main effects of language group [F(2, 24)=1.78, p=0.19] or hemisphere [F(7, 168)=1.25, p=0.28], and no significant interaction between language group and hemisphere [F(14, 168)=1.07, p=0.39]. Mixed model analysis of variance on Tb amplitudes revealed that there was a significant main effect of hemisphere [F(7, 168)=5.55,
p<0.001, but not language group [F(2, 24)=0.41, p=0.67], and there was no significant interaction between language group and hemisphere [F(14, 168)=1.03, p=0.43]. LSD post-hoc analysis on hemisphere indicated that Tb amplitudes at T8 (right hemisphere) were more negative than those at T7 (left hemisphere) across the tokens and language groups.

Figure 9. Grand mean ACC T-Complex Ta and Tb amplitudes at T7 and T8 (Note: Error bars show ±1 SE)
3.4 Mismatch Negativity (MMN)

Mismatch negativity was measured using the difference waveforms obtained by subtracting responses to stimuli presented as deviants from the responses to stimuli presented as standards. The responses to standards and deviants, along with the difference waveforms are shown for each stimulus pair in Appendices 2-7. Inversion of MMN at the mastoids (M1 & M2) is also shown in Appendices 8 and 9.

1. Across-category pairs
   
a. Token pair 4S7D

   The grand mean difference waveforms elicited by token pair 4S7D (token 4 as a standard & token 7 as a deviant) at Fz are shown for each language group in Figure 10. Since stimuli were a synthetic vowel-consonant-vowel (VCV) format, there were two transitions in each stimulus. The first is the first transition from vowel (/i/) to liquid (/r/ or /l/) at 300 ms and the other is the transition from liquid (/r/ or /l/) to vowel (/i/) at 555 ms. Both early (400-650 ms) and late (655-905 ms) MMN responses were obtained. A paired t-test was used to determine if the MMN was statistically significant in the grand mean waveform (p < 0.05). Statistically significant MMN regions are indicated by the red-colored boxes on the grand mean waveform in each language group. The early MMN (400-650 ms) was statistically significant only for Americans and Koreans; however the late MMN (655-905 ms) was statistically significant for all the three language groups. Grand mean MMN latencies and amplitudes at Fz for token pair 4S7D are shown in Figure 11. Early MMN latencies and amplitudes did not differ significantly between Americans and Koreans (p > 0.05), but the early MMN was not detected for Japanese. Late MMN latencies and amplitudes did not show significant differences between the three language groups (p > 0.05).
Figure 10. Grand mean difference waveforms for across-category pairs (token pairs 4S7D and 7S4D) at Fz are shown for each language group. The vertical dotted line at 300 ms indicates the first transition from vowel /i/ to liquid (/r/ or /l/) and the vertical dotted line at 555 ms indicates the second transition from liquid (/r/ or /l/) to vowel /i/ in the stimuli.
As seen in Figure 10, for token pair 7S4D, both early (400-650 ms) and late (655-905 ms) MMN responses were obtained. Paired t-tests revealed that the early MMN (400-650 ms) was statistically significant only for Americans and Koreans; however the late MMN (655-905 ms) was statistically significant for all the three language groups, similar to the response to tokens 4S7D. Statistically significant MMN regions were indicated by the blue-colored ovals on the grand mean waveforms. As seen in Figure 10, the latency regions of the early MMN for tokens
7S4D was earlier than for tokens 4S7D in Americans and Koreans. Americans showed the most negativity for the late MMN for tokens 7S4D compared to other groups. Grand mean MMN latencies and amplitudes at Fz for tokens 7S4D are shown in Figure 11.

Early MMN latencies did not differ significantly between Americans and Koreans, and late MMN latencies did not show a significant difference between the three language groups. However, late MMN amplitudes for token pair 7S4D indicated a significant effect of language group \( F(2, 22)=7.095, p < 0.01 \). LSD post-hoc test revealed that Americans showed significantly larger amplitudes than Koreans \( p < 0.01 \) and Japanese \( p < 0.05 \). Mixed model analysis of variance on MMN token pair 7S4D amplitudes in Americans and Koreans revealed that there were no significant main effects of language group or timing (early vs. late); however, there was a significant interaction of timing (early vs. late) and language group \( F(1, 14)=13.96, p<0.01 \). LSD post-hoc analysis indicated that the amplitude of the late MMN was significantly larger for Americans than for the other groups \( p < 0.05 \) for all comparisons.
2. **Within-category pairs perceived as /r/**

   a. **Token pair 1S4D**

      The grand mean difference waveforms elicited by token pair 1S4D (token 1 as a standard & token 4 as a deviant) at Fz are shown for each language group in Figure 12. As seen in Figure 12, weak early (400-650 ms) or late (655-905 ms) MMN responses were obtained. A paired t-test was used to determine if MMN was statistically significant in the grand mean waveforms (p < 0.05). Statistically significant MMN regions were indicated by the red-colored boxes on the grand mean waveform. The early MMN (400-650 ms) was statistically significant only for Americans; however, the late MMN (655-905 ms) was statistically significant only for Koreans and Japanese. Grand mean MMN latencies and amplitudes at Fz for tokens 1S4D are shown in Figure 13 (note that conditions omitted in which 4 or fewer participants in a group had present responses). There were no significant effects of language group on MMN latencies and amplitudes (p > 0.05).
Figure 12. Grand mean difference waveforms for within-category pairs perceived as /r/ (token pairs 1S4D and 4S1D) at Fz are shown for each language group. The vertical dotted line at 300 ms indicates the first transition from vowel /i/ to liquid /r/ or /l/ and the vertical dotted line at 555 ms indicates the second transition from liquid /r/ or /l/ to vowel /i/ in the stimuli.
Figure 13. Grand mean MMN latencies (top) and amplitudes (bottom) for within-category pairs perceived as /t/ at Fz (Note: Error bars show ±1 SE)

b. Token pair 4S1D

The grand mean difference waveforms elicited by token pair 4S1D (token 4 as a standard & token 1 as a deviant) at Fz are shown for each language group by blue color in Figure 12. As seen in Figure 12, weak early (400-650 ms) or late (655-905 ms) MMN responses were obtained. A paired t-test was used to determine if the MMNs were statistically significant in the grand
mean waveforms (p < 0.05). Statistically significant MMN regions were indicated by the blue-colored ovals on the grand mean waveforms. The early MMN (400-650 ms) was statistically significant only for Americans and Koreans; however, the late MMN (655-905 ms) was statistically significant for all the three language groups. Grand mean MMN latencies and amplitudes at Fz for token pair 4S1D are shown in Figure 13. There were no significant effects if language group on MMN latencies or amplitudes (p > 0.05).

3. **Within-category pairs perceived as /l/**

a. **Token pair 7S10D**

The grand mean difference waveforms elicited by token pair 7S10D (token 7 as a standard & token 10 as a deviant) at Fz are shown for each language group in Figure 14. The early MMN was not detected, but the late MMN was present. A paired t-test was used to determine if MMN were statistically significant in the grand mean waveforms (p < 0.05). Statistically significant MMN regions are indicated by the red-colored boxes on the grand mean waveforms. Only the late MMN (655-905 ms) was statistically significant for all the three language groups. Grand mean MMN latencies and amplitudes at Fz for token pair 7S10D are shown in Figure 15. There were no significant effects of language group on MMN latencies and amplitudes (p > 0.05).
Figure 14. Grand mean MMN waveforms for within-category pairs perceived as /l/ (token 7S10D and token 10S7D) at Fz for each language group. The vertical dotted line at 300 ms indicates the first transition from vowel /i/ to liquid /r/ or /l/ and one at 555 ms indicates the second transition from liquid /r/ or /l/ to vowel /i/ in the stimuli.
b. Token pair 10S7D

The grand mean difference waveforms elicited by token pair 10S7D (token 10 as a standard &
token 7 as a deviant) at Fz are shown for each language group in Figure 14. MMN was
somewhat difficult to identify in these waveforms and he only statistically significant MMN
regions was the late MMN, which was statistically significant only for Japanese.

Figure 15. Grand mean MMN latencies (top) and amplitudes (bottom) for within-category pairs perceived
as /l/ at Fz (Note: Error bars show ±1 SE)
4. Summary of MMN results

Six different stimulus sets elicited MMN either the early (400-650 ms), the late (655-905 ms) response time frame, or both. Each language group showed different patterns of MMN results. Results of the paired t-tests carried out on the grand mean waveforms are summarized in Table 3. When stimuli crossed the phonetic boundary, the early MMN (400ms-650ms) was statistically significant only for Americans and Koreans; however, the late MMN (655ms-905ms) was statistically significant for most conditions for all three language groups.

<table>
<thead>
<tr>
<th>Category</th>
<th>Early MMN (400-650ms)</th>
<th>Late MMN (655-905ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>American</td>
<td>Korean</td>
</tr>
<tr>
<td>across-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>category</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Token</td>
<td>Token</td>
<td>Token</td>
</tr>
<tr>
<td>4S7D</td>
<td>1S4D</td>
<td>7S4D</td>
</tr>
<tr>
<td>√</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
| Table 3. Significant paired t-test results on the grand mean MMN waveforms. The symbol, √, indicates that the MMNs were significantly different from zero (p < 0.05) and the symbol, -, means that the t-values were not significant.

Figure 16 summarizes grand mean latencies (top) and amplitudes (bottom) for the six stimulus sets for the three language groups. Data from across- vs. within-category pairs were collapsed to determine whether the early (or late) MMN responses from across-categories are significantly different from those from within-categories.
Figure 16. Grand mean early and late MMN latencies (top) and amplitudes (bottom) at Fz (Note: Error bars show ±1 SE)

**a. Comparisons between early MMN across- and within-category**

Americans and Koreans showed early MMN responses for across-category pairs (token pairs 4S7D & 7S4D) and for within-category pairs (token pairs 1S4D & 4S1D). Data from these pairs were collapsed to see whether the early MMN responses from across-categories are earlier and larger than those from within-categories. Mixed model analyses of variance (ANOVA) on the early MMN latencies revealed that there were no significant main effects of language group (American vs. Korean) or category (across- vs. within-), and no interaction between group and category. However, mixed model analyses of variance (ANOVA) on the early MMN amplitudes revealed a main effect of category \[F(1, 16)=4.603, p < 0.05\]. There was no significant main
effect of language group and no significant interaction between group and category (p > 0.05).
LSD post-hoc analysis indicated that early MMN amplitudes across-category were significantly
more negative (larger) than within-category.

b. Comparisons between late MMN across- within-category

As seen in Figure 16, all the three language groups showed late responses for across-
category pairs and most for within-category pairs. Data from the two across-category pairs were
collapsed and data from /r/ and /l/ within-category pairs were also collapsed to see whether the
late MMN responses from across-categories are earlier and larger than those from within-
categories. Mixed model analyses of variance (ANOVA) on the late MMN latencies revealed
that there were no significant main effects of language group (American vs. Korean vs. Japanese)
or category (across- vs. within-), and there was no significant interaction between group and
category. However, mixed model analyses of variance (ANOVA) on the late MMN amplitudes
revealed that although there was no significant main effect of language group, there were main
effects of category [F(2, 40)=10.12, p< 0.001], and a significant interaction between group and
category [F(4, 40)=4.5, p < 0.01]. LSD post-hoc analysis indicated that late MMN amplitudes
from across-category were significantly more negative than those from the within-category
conditions (p < 0.01). Further, late MMN amplitude from across-category in Americans was
significantly more robust than all the other conditions (p < 0.05 for all comparisons).

3.5 P3a

The grand mean P3a waveforms taken from the difference waveforms elicited by the
across-category pairs (token pairs 4S7D & 7S4S), within-category pairs perceived as /r/ (tokens
pairs1S4D & 4S1D) and within-category pairs perceived as /l/ (tokens pairs 7S10D & 10S7D) at Fz are shown for each language group in Figure 17. Like MMN, there were also early (600-700 ms) and late (900-1000ms) P3a responses. As seen in Figure 17, Americans and Koreans showed early P3a responses for across-category pairs and all the three language groups showed late P3a responses. For within-category pairs, three language groups had late P3a responses in general. Grand mean early and late P3a latencies and amplitudes are shown as a function of token in each language group in Figure 18. Early P3a latencies and late P3a latencies did not differ significantly between the three language groups except for the latencies of early P3a for token pair 7S4D. P3 latency for token pair 7S4D in Koreans was significantly longer than that of Americans (p < 0.05). Early P3a amplitudes and late P3a amplitudes did not differ significantly between the three language groups (p > 0.05).
Figure 17. Grand mean waveforms for across- and within-category pairs at Fz in each language group. The vertical dotted line at 300 ms indicates the first transition from vowel /i/ to liquid /ɾ/ or /l/ and one at 555 ms indicates the second transition from liquid /ɾ/ or /l/ to vowel /i/ in the stimuli. The first column indicates P3a waveforms elicited by the across-category pairs. The red arrow indicates P3a corresponding with tokens 4S7D and the blue arrows corresponding with token pair 7S4D. The second column indicates P3a waveforms elicited by within-category pairs perceived as /ɾ/. The red arrow indicates P3a corresponding with token pair 1S4D and the blue arrow for token pair 4S1D. The third column indicates P3a waveforms elicited by within-category pairs perceived as /l/. The red arrows indicates P3a corresponding with token 7S10D and the blue arrows corresponding with tokens 10S7D.
Figure 18. Grand mean early and late P3a latencies (top) and amplitudes (bottom) at Fz (Note: Error bars show ±1 SE).

Table 4 summarizes results from multiple neurophysiologic and behavioral responses to English intervocalic /iri-ili/ stimuli from the three language groups and also shows the effects of native language on perception and neurophysiologic processing of English intervocalic /ʃ/ and /l/.
Table 4. Summary of significant results as a function of language group.

<table>
<thead>
<tr>
<th>Effects of Native language</th>
<th>Significant results</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC N1-P2 at FCz</td>
<td>NO</td>
</tr>
<tr>
<td></td>
<td>No significant differences between language groups</td>
</tr>
<tr>
<td>ACC T-complex at T7 &amp; T8</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>Ta latency</td>
</tr>
<tr>
<td></td>
<td>- similar Ta latencies for Americans and Koreans at T7 and T8, but significantly longer Ta latencies for Japanese at T7 and T8 compared to other language groups.</td>
</tr>
<tr>
<td></td>
<td>Tb latency</td>
</tr>
<tr>
<td></td>
<td>- significantly shorter Tb latencies for Americans and Koreans at T7 than at T8, but Japanese did not show a significant effect of hemisphere.</td>
</tr>
<tr>
<td></td>
<td>Tb latency at T8</td>
</tr>
<tr>
<td></td>
<td>- Japanese (shortest) &lt; Americans &lt; Koreans (longest)</td>
</tr>
<tr>
<td></td>
<td>Tb amplitude</td>
</tr>
<tr>
<td></td>
<td>- significantly more negative at T8 than at T7 for all three language groups</td>
</tr>
<tr>
<td>MMN at Fz</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>Early MMN (400-650 ms)</td>
</tr>
<tr>
<td></td>
<td>- obtained for Americans and Koreans, but not for Japanese</td>
</tr>
<tr>
<td></td>
<td>- significantly larger for across-category pairs than for within-category pairs in these language groups</td>
</tr>
<tr>
<td></td>
<td>Late MMN (655-905 ms)</td>
</tr>
<tr>
<td></td>
<td>- obtained for all three language groups</td>
</tr>
<tr>
<td></td>
<td>- significantly larger for across-category pairs in Americans compared to other language groups</td>
</tr>
<tr>
<td></td>
<td>- significantly larger for across-category pairs than for within-category pairs only in American</td>
</tr>
<tr>
<td>P3a at Fz</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>Early P3a (600-700 ms)</td>
</tr>
<tr>
<td></td>
<td>- obtained for Americans and Koreans, but not for Japanese</td>
</tr>
<tr>
<td></td>
<td>- P3a latency for tokens 7S4D was significantly shorter in Americans than for Koreans</td>
</tr>
<tr>
<td></td>
<td>Late P3a (900-1000 ms)</td>
</tr>
<tr>
<td></td>
<td>- obtained for all three language groups</td>
</tr>
<tr>
<td></td>
<td>- no significant difference across language groups</td>
</tr>
<tr>
<td>Behavior (ID &amp; Dis)</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>Categorical perception by Americans, categorical-like perception with the shifted identification boundary by Koreans, and non-categorical perception by Japanese</td>
</tr>
</tbody>
</table>
CHAPTER 4: Discussion
The overall goal of this dissertation was to compare the effects of native language on the processing of medial /r/ and /l/ in American, Korean, and Japanese listeners using ACC, MMN, P3a and behavioral measures. Significant effects of language group were obtained for the ACC T-complex, MMN, P3a and behavioral identification and discrimination. These findings are interpreted and discussed below.

4.1 Perceptual identification and discrimination

English intervocalic /r/ and /l/ were perceived in a categorical manner by Americans, in a categorical-like manner by Koreans and in a non-categorical manner by Japanese. These findings are consistent with the hypotheses and preliminary data (An et al., 2013), and are consistent with previous behavioral studies (Miyawaki et al., 1975; Henly & Sheldon, 1986; Ingram & Park, 1998; Sheldon & Strange, 1982). Since the Japanese and Korean participants in this study did not differ significantly in terms of their language proficiency and all arrived in an English speaking country after the age of 19 years, these results suggest that native language sound repertoire differentially influences the perception of English /r/ and /l/, which are non-native sounds for the Koreans and Japanese listeners. Native language can both facilitate and interfere with speech sound perception. In this study, the native repertoire facilitated the perception of /r/ and /l/ for American and Korean listeners, but it interfered with, or at least did not facilitate, perception for Japanese listeners. As reviewed in the introduction, English medial /r/ and /l/ are phonemic for American participants, so Americans showed a clear categorical perception with a boundary between token 4 (F3=1890 Hz) and token 6 (F3=2150 Hz). In the Japanese language system, English /r/ and /l/ are not phonemic, so Japanese participants exhibited non-categorical perception consistent with previous studies (e.g., Miyawaki et al., 1975; Mackain et al., 1981).
Therefore, native language interferes with the perception of English medial /r/ and /l/ for native Japanese participants. Categorical-like perception of medial /r/ and /l/ was obtained by native Korean participants. The categorical boundary for Korean participants was shifted toward /l/ between token 5 (F3=2020 Hz) and token 7 (F3=2280 Hz). That is, the boundary of Korean participants was shifted by 130 Hz of F3 toward the /l/ category compared to Americans. The categorical-like perception obtained for Korean participants reflects the effects of their native language; specifically, the partial phonetic model available in the Korean language (intervocalic [ɾ]-[l]) facilitated the perception of English /r/ and /l/. Even though English /r/ and /l/ are not phonemic in the Korean language system, Koreans performed better than Japanese. For this reason, it seems that Koreans had categorical-like perception. This was supported by the finding of a steep categorical perception index (CPI) that did not differ significantly from Americans, but was a steeper CPI compared to Japanese. It is not clear from this study, whether the model would facilitate perception in Koreans if /r/ and /l/ were not in medial position. A possible reason for the shift toward /l/ in Koreans might be due to the Korean single liquid /l/. That is, Koreans have a good exemplar of /l/ whereas they have a less-good exemplar of /r/, potentially leading to a boundary shift toward their good exemplar (/l/). This result might be explained by category goodness assimilations in the perceptual assimilation model (PAM)-L2 (Best & Tyler, 2007; Best, 1995). Category goodness assimilations occur when two non-native categories are equally perceived as the same native category but one is perceived as being better than the other. This postulation predicts that second language learners can discriminate these non-native categories well, even though not as well as native speakers who have separate categories.

In addition, these behavioral results are consistent with the predictions of the native language neural commitment model (Kuhl, 2004; Kuhl et al., 2008). Since all the Japanese and
Korean participants came to the U.S. after the age of 19 years, their neural networks are committed to their native language patterns and would be highly established. For example, since Japanese neural networks would be committed to Japanese /r/ as a poor prototype of English /r/ or /l/, Japanese participants must assign new speech sounds (English /r/ and /l/) to their existing native Japanese /r/ category. As a result, Japanese showed poor performance for both identification and discrimination. However, since Korean neural networks would be committed to both Korean /l/ as a good prototype of English /l/ and Korean [ɾ] as a partial prototype of English /r/, Korean participants could assign new speech sounds (English /r/ and /l/) to their two existing categories (/l/ and [ɾ]). Accordingly, Koreans showed good identification and discrimination performance comparable with those of Americans. In sum, the behavioral findings indicated clear effects of native language on the perception of English intervocalic /r/ and /l/ between American and Korean and Japanese participants.

4.2 Encoding tapped by the Acoustic Change Complex (ACC)

1. The ACC P1-N1-P2 complex

The ACC P1-N1-P2 latencies and amplitudes did not show significant differences between language groups, consistent with the hypothesis and with the findings of the preliminary study (An et al., 2013). In contrast to the preliminary study, the ACC in this study was evoked using a more complex paradigm, in which MMN was simultaneously elicited. Nonetheless, although the responses are smaller using this complex paradigm, they are generally consistent with the previous study. These results also align with previous studies suggesting that the midline-central P1-N1-P2 responses to a speech continuum (e.g. /ba/-/pa/, /da/-/ta/, /ga/-/ka/ continua) are sensitive to acoustic changes, and not phonetic changes when a passive listening
condition is used (Elangovan & Stuart, 2011; Sharma & Dorman, 1999, 2000; Sharma et al., 2000). However, while these previous studies investigated the onset P1-N1-P2 in response to temporal information (VOT), the current study examined the P1-N1-P2 complex elicited by spectral information (F3). Previous studies reported that P1-N1-P2 latencies, P1-N1 amplitude, and N1-P2 amplitude increased significantly as VOT change increased, whereas the present results did not show systematic amplitude or latency increases as F3 increased. It is possible that the P1-N1-P2 responses are more sensitive to VOT changes than spectral formant frequency changes.

Attention might be a factor that facilitates phonetic processing because there is evidence that the amplitude of N1-P2 can reflect categorical perception for voice onset time when it is obtained using an active paradigm. Dorman (1974) presented a within-category shift sequence (20-0 ms VOT) and an across-category shift sequence (20-40 ms VOT) to participants using an oddball paradigm in which subjects attended to the stimuli being presented. The N1-P2 amplitude was significantly larger in response to the across-category than the within-category stimuli. Dorman, therefore, concluded that amplitude of N1-P2 can be used as an index of categorical phonetic perception. However, since the current study was conducted while participants ignored the stimuli being presented and the speech stimuli differed in terms of spectral information, these parametric differences may explain the lack of consistency with Dorman (1974). Amplitudes and latencies did not show consistent shifts with increases in F3 frequency in the current study. In sum, ACC P1-N1-P2 responses did not differ significantly between language groups in the present study, suggesting little effect of native language on the primary cortical encoding of these speech sounds. In other words, the neural networks
established by native language did not differentially alter the primary cortical encoding of these sounds regardless of language.

2. The ACC T-complex

In contrast to the midline-central response, the lateral-temporal ACC T-complex differed significantly between language groups suggesting that native language influences secondary cortical processing of English /r/ and /l/. Differences between language groups included Ta latency, Tb latency, and Tb amplitude. These findings are consistent with the hypothesis and are consistent with the results of the preliminary study (An et al., 2013).

Japanese participants showed significantly prolonged Ta latencies (424 ms) over both hemispheres (T7 and T8) compared to Americans (399 ms) and Koreans (400 ms). The prolonged Ta latencies in Japanese seem to reflect difficulty encoding the transition from vowel (/i/) to liquid (/r/ or /l/). Tb latencies in the left hemisphere did not differ significantly between language groups, but Tb latencies in the right hemisphere were significantly different between language groups. Japanese showed the shortest Tb latencies (470 ms), Koreans had the longest Tb latencies (498 ms), and Americans demonstrated intermediate Tb latencies (482 ms) in the right hemisphere. The very short Tb latencies in Japanese listeners possibly reflects acoustic processing of English medial /r/ and /l/. If /r/ and /l/ are encoded as simple acoustic signals, latencies would be shorter. Further supporting this assertion is that the waveform morphology was sharper/more clear in the Japanese compared to the other language groups, which could also be explained by acoustic processing. The longer Tb latencies observed in Americans and Koreans likely reflects further phonologic/linguistic processing of these sounds. Eulitz et al. (1995) reported that the Tb latencies in adults were prolonged in response to a vowel, compared
to a simple tone. Therefore, it might be possible that Japanese may treat the stimuli, /iri/ or /ili/ as simple stimuli, not as linguistic events, giving relatively short Tb latencies. In contrast, Americans and Koreans probably treat the stimuli, /iri/ or /ili/ as linguistic events, leading to prolonged Tb latencies. Several studies have reported that the Tb peak can be a marker of linguistic processing because smaller amplitude or delayed or deviant latencies are found in children with autism, SLI, or Down syndrome (Tonnquist-Uhlen, 1996; Bruneau et al., 2003; Groen et al., 2008; Shafer et al., 2011).

Japanese did not show hemispheric laterality for Tb latencies. However, Americans and Koreans had significantly shorter Tb latencies in the left hemisphere than those in the right hemisphere, as would be expected for stimuli processed as linguistic. The longer latency and increased difference for Tb latencies across hemispheres in Koreans may reflect greater difficulty processing these sounds compared to Americans. Previous studies using dipole source modeling suggested that the Tb response is attributed to radial dipoles located in the secondary auditory cortex (Scherg and Von Cramon, 1985, 1986; Ponton et al., 2002). The shorter Tb latencies in the left hemisphere in our study might be related to the properties of the left auditory cortex allowing for efficient auditory processing for speech or language. Furthermore, the Tb peak from all four tokens showed negative amplitudes at T8 compared to those at T7 in all three language groups. Zatorre and Belin (2001), using positron emission tomography, determined that processing of spectral aspects of speech is weighted toward the right hemisphere, whereas temporal aspects are processed in the left hemisphere. The present study manipulated third formant frequency transition, a spectral/spectro-temporal cue. More negative Tb amplitudes at T8 appear may reflect a predominance of the right hemisphere processing for spectral features regardless of native language. In sum, differences in Ta and Tb peaks across language groups
supports that potential for the T-complex to be used as a marker of linguistic processing to reflect the effects of native language. That is, it appears that the neural networks established by native language differentially influence the secondary cortical processing.

4.3 Pre-attentive discrimination tapped by the Mismatch Negativity (MMN)

The MMN differed significantly between language groups in this study suggesting that native language influences the automatic, pre-attentive discrimination of English medial /r/ and /l/. An early and a late MMN were obtained.4

When stimuli crossed the phonetic boundary (token pairs 4S7D and 7S4D), an early MMN (400-650ms) was detected only for Americans and Koreans. The early MMN was not detected (not statistically significant) for the Japanese. In contrast, however, the late MMN (655-905ms) was detected in all three language groups. When considering the timing of the early MMN and the late MMN, it appears that the early MMN might reflect discrimination of the first transition in the stimuli from vowel (/i/) to liquid (/r/ or /l/) and the late MMN might reflect the discrimination of the second transition in the stimuli from liquid to vowel. The behavioral results indicated that Americans and Koreans showed significantly better discrimination of /r/ and /l/ than Japanese. Previous studies have supported that better behavioral discrimination is closely associated with shorter MMN latencies and/or larger MMN amplitudes (Näätänen et al., 1993; Kujala et al., 2001; Novitski et al., 2004).

The fact that the early MMN was detected only for Americans and Koreans, and not for Japanese and that the late MMN was detected for all three language groups raises the possibility

4 Presence of both the early and late MMN was supported by clear inversion of the responses at mastoid electrodes along with statistical significance of the two peaks.
that each language group might differentially utilize acoustic information for phonetic processing. In this study, the first transition from vowel to liquid was most difficult to discriminate or was impossible to discriminate for the Japanese because the discrimination “target” was not present in their language. In contrast, the second transition from liquid to vowel was probably somewhat less difficult, because a MMN was detected. The transition cues are critical for perception of English /r/ and /l/ compared to cues from the steady-state portion in initial position (Iverson & Kuhl, 1996). Our MMN results support not only the importance of the transition cues for English /r/ and /l/ discrimination, but also that the first transition cue is essential to support behavioral perception of /r/ and /l/. Absent/undetectable early MMN and present late MMN findings might reflect slower processing of stimuli that are difficult to process (Datta et al., 2010; Shafer et al., 2005; Uwer et al., 2002). This might indicate that weighting of or automatic selection of relevant speech cues appears altered for Japanese and Koreans. Another possibility is that each language group might automatically acoustic information for phonetic processing in a differential manner.

Early MMN amplitudes from across-category pairs were significantly larger than amplitudes from within-category pairs both in Americans and Koreans. Japanese did not show any early MMN responses. Dehaene-Lambertz (1997) reported that young French adults showed significantly larger MMN for French across-category stimuli (/ba/-/da/) than for French within-category compared to Hindi across-category or within-category stimuli (/da/-/Da/). If the four speech contrasts are processed depending on acoustic information only, MMN responses would not have differed because the acoustic differences across the speech contrasts were equivalent in terms of F2 and F3. However, the French listeners showed larger MMN only for the French across-category stimuli, so it was concluded that MMN amplitude was influenced by native language and MMN can signal phonetic processing in addition to acoustic processing. Sharma
and Dorman (1999) also demonstrated that MMN was larger for across-category stimulus pairs (/da/-/ta/) compared to within-category pairs (/ta/-/ta/). Our results support the claim that MMN amplitudes can reflect effects of native language since MMN responses were more robust for across-category than within-category conditions. One would predict that MMN would be present both within- and across-category because of the acoustic differences in present in the token pairs. When MMN was absent/undetectable within-category, this might indicate that participants were processing the sounds linguistically. It is possible that the pattern of results across language groups might differ if a shorter ISI had been used (Werker and Logan, 1985; Yu, 2013).

Americans had significantly larger late MMN amplitudes for across-category pairs than for within-category pairs. Further, their MMN amplitudes were larger compared to the other language groups. Even though the processes generating late MMN do not seem to dominantly influence discrimination of intervocalic /r/ and /l/ compared to early MMN, enhanced late MMN amplitudes in Americans may indicate easier or more efficient phonetic processing, especially across-phonetic category, reflecting effects of native language. Therefore, the established neural networks by native language seem to differently influence pre-attentive discriminative processing.

### 4.4 Orienting and involuntary attention shifts tapped by the P3a

The P3a were also modified by native language. For across-category pairs, Japanese participants had no early P3a but had a present late P3a. In contrast, Americans and Koreans had both early and late P3a responses. Therefore, results support the presence of a clear attention shift to phonetic category change in Americans and Koreans but the difficulty with attention shifting to phonetic-category change in Japanese, reflecting the effects of native language. It is
possible that Japanese may have difficulty creating memory representations of /r/ and /l/ (Novitski, Tervaniemi, Huotilainen, and Naatanen, 2004) or difficulty maintaining the sounds in working memory (Light, Swerdlow, & Braff, 2007).

4.5 Summary, limitations and implications

This study examined the mechanisms underlying the processing of /r/ and /l/ over the overlapping, complex time-course of processing leading to perception. The ACC examined the encoding of /r/ and /l/ at the levels of primary (P1-N1-P2) and secondary (T-complex) auditory cortex. The MMN examined pre-attentive discrimination of /r/ and /l/, predominantly at the level of auditory cortex. P3a examined orienting/attention-switching to these sounds. Finally, behavioral measures were used to examine the perception of these sounds and more specifically, their identification and discrimination. The results of this study were generally consistent with the hypotheses presented. Further, the patterns of performance were consistent with predictions based on the native language neural commitment model.

In sum, all of the measures used in this study except for the midline-central N1-P2 complex showed significant effects of native language on the processing of /r/ and /l/.
Behavioral identification and discrimination performance reflected the presence of differential effects of native language. Americans and Koreans showed categorical perception, but Japanese did not. Further, the category boundary was shifted toward the /l/ category in Koreans. The midline-central ACC did not show significant differences as a function of language groups, suggesting little effect of native language on the primary cortical encoding of English /r/ and /l/. In contrast, the lateral-temporal ACC differed significantly between language groups, suggesting that native language influences secondary cortical processing of these sounds. The MMN
differed significantly between language groups, suggesting that native language influences the automatic, pre-attentive discrimination of English medial /r/ and /l/. The early MMN was present only for Americans and Koreans, but the late MMN was obtained in all language groups. The P3a showed different presence/absence patterns between language groups, consistent with the effects of native language as well.

It should be noted that these results apply to the synthetic American English stimuli in intervocalic position and to parameters used for the current study. Natural tokens and other syllable positions might lead to a different pattern of results. The relatively long ISI used may have forced participants toward a linguistic processing mode and it is possible that a very different pattern of results would be obtained using a short ISI.

There are several implications. The T-complex reflected the influence of native language, permitting future use of the ACC to evaluate linguistic processing. Korean’s partial phonetic model might be used to facilitate training of /r/ and /l/ perception for English learning. For those learning English from other language groups, one might look at the language to see if there are potential phonetic tools that could be used to facilitate the learning of difficult speech contrasts. This study shows that a complex paradigm can be used to simultaneously elicit ACC, MMN and P3a. The trade-off, however, appears to be relatively small amplitudes.

Future work will include more detailed topographic analysis of the effects of language group on the spatio-temporal distribution of brain responses to /r/ and /l/ in these language groups. In addition, future studies will examine the effects of auditory and language training on the neural mechanisms supporting the perception of /r/ and /l/ in native Korean and native Japanese listeners.
CHAPTER 5: Conclusions
In conclusion, native language influences the processing and perception of non-native sounds. English /r/ and /l/ were processed efficiently in the American listeners at each level of processing examined. The partial phonetic model available to Koreans facilitated processing of English /r/ and /l/, perhaps by leading to a neural network that is somewhat similar to Americans. The lack of a useable model for /r/ and /l/ processing in Japanese interfered with processing of English /r/ and /l/ at each level of processing examined. Japanese may lack neural networks committed to the properties of English /r/ and /l/, interfering with phonetic processing. Taken together, native language influenced the neural encoding (T-complex), pre-attentive discrimination (MMN), orienting/attention-switching (P3a) and perception (identification/discrimination) of English /r/ and /l/ in the American, Korean, and Japanese listeners in this study.
Appendices
Appendix 1. Individual background information for Koreans and Japanese collected by LEAP-Q

<table>
<thead>
<tr>
<th></th>
<th>CA Years</th>
<th>AOA Years</th>
<th>LOR Years</th>
<th>Exposure to English (%)</th>
<th>Proficiency</th>
<th>Speak</th>
<th>Listen</th>
<th>Read</th>
</tr>
</thead>
<tbody>
<tr>
<td>Koreans</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1</td>
<td>25.5</td>
<td>24.3</td>
<td>1.3</td>
<td>50</td>
<td>8</td>
<td>6</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>S2</td>
<td>31</td>
<td>22.5</td>
<td>7</td>
<td>10</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>S3</td>
<td>26</td>
<td>25.2</td>
<td>3</td>
<td>30</td>
<td>3</td>
<td>7</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>S4</td>
<td>40.5</td>
<td>33</td>
<td>7</td>
<td>90</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>S5</td>
<td>30.1</td>
<td>30</td>
<td>0.6</td>
<td>60</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>S6</td>
<td>26.6</td>
<td>25.4</td>
<td>1.2</td>
<td>60</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>S7</td>
<td>32.6</td>
<td>30.6</td>
<td>2</td>
<td>20</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>S8</td>
<td>27.5</td>
<td>27</td>
<td>0.6</td>
<td>70</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>S9</td>
<td>31.8</td>
<td>31.4</td>
<td>0.5</td>
<td>40</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Japanese</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1</td>
<td>40</td>
<td>34.2</td>
<td>7</td>
<td>50</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>S2</td>
<td>25</td>
<td>21</td>
<td>3</td>
<td>30</td>
<td>8</td>
<td>8</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td>25.5</td>
<td>18.6</td>
<td>6</td>
<td>70</td>
<td>8</td>
<td>9</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>S4</td>
<td>32.5</td>
<td>25.5</td>
<td>6.4</td>
<td>40</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>S5</td>
<td>36.1</td>
<td>34.1</td>
<td>2</td>
<td>30</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>S6</td>
<td>40.5</td>
<td>37.9</td>
<td>2.3</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>S7</td>
<td>42.8</td>
<td>36</td>
<td>6.5</td>
<td>20</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>S8</td>
<td>39.5</td>
<td>36.3</td>
<td>3.2</td>
<td>20</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>S9</td>
<td>34.9</td>
<td>31.7</td>
<td>3</td>
<td>60</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>
Appendix 2. The grand mean waveforms obtained in response to standards and deviants along with MMN are shown for tokens 4S7D.
Appendix 3. The grand mean waveforms obtained in response to standards and deviants along with MMN are shown for tokens 7S4D.
Appendix 4. The grand mean waveforms obtained in response to standards and deviants along with MMN are shown for tokens 1S4D.
Appendix 5. The grand mean waveforms obtained in response to standards and deviants along with MMN are shown for tokens 4S1D.
Appendix 6. The grand mean waveforms obtained in response to standards and deviants along with MMN are shown for tokens 7S10D.

- First transition at 300ms from vowel /i/ to liquid (/r/ or /l/)
- Second transition at 555ms from liquid (/r/ or /l/) to vowel /i/

Late MMN

<table>
<thead>
<tr>
<th>Latency (ms)</th>
<th>Standard token 7</th>
<th>Deviant token 10</th>
<th>Difference waveform</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>600</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>800</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 7. The grand mean waveforms obtained in response to standards and deviants along with MMN are shown for tokens 10S7D.
Appendix 8. The grand mean waveforms for tokens 1S4D, 4S7D, and 7S1D by red and their inversion waveforms at left mastoid (M1) by blue and right mastoid (M2) by black.
Appendix 9. The grand mean waveforms for tokens 4S1D, 7S4D, and 10S7D by red and their inversion waveforms at left mastoid (M1) by blue and right mastoid (M2) by black.
References
the perception of a phonetic continuum. *Biological Psychology, 24*(3), 197-207.

University of Illinois).


[guidelines]. www.asha.org/policy

of english /t/ and /l/ by korean and japanese listeners. *Proceedings of Meeting on
Acoustics, Montreal, Canada. , 19. (2-7 June)

W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-

Best, C. T., & Tyler, M. D. (2007). Nonnative and second language speech perception:
Commonalities and complementaries. In In M.J. Munro, & O.-S. Bohn (Eds.) (Eds.),
Second language speech learning: The role of language experience in speech
perception and production (pp. 13-34). Amsterdam: John Benjamins.


1-15.


*Electroencephalography and Clinical Neurophysiology, 59*(1), 9-20.


Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology, 38*(1), 1-21.


