TESTING PHONOLOGICAL REPRESENTATIONS THROUGH BEHAVIORAL AND ELECTROPHYSIOLOGICAL METHODS

By

Andrew Trotter

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Linguistics—Doctor of Philosophy

2019

ABSTRACT

TESTING PHONOLOGICAL REPRESENTATIONS THROUGH BEHAVIORAL AND ELECTROPHYSIOLOGICAL METHODS

By

Andrew Trotter

Languages differ in the sounds that make up their phonemic inventories. These sounds, or phonemes, are represented abstractly in the mind of a speaker, making up the underlying representations of lexical items. Additional sounds may exist in a given language, surfacing via a derivative process of phonological rule application, where they are called allophones. There are also many speech sounds that are unrepresented entirely, as separate categories, within a particular language. In either case, learners of a second language have difficulty perceiving and producing the sounds that do not exist or have differing phonological status in their native language. English and Spanish present us with an ideal case study of these differences. Both languages contain the [d], [ð], and [ɾ] sounds, but they differ in how they organize them. While /d/ is a phoneme in both languages, in Spanish [ð] is an allophone of /d/, while /ɾ/ is a separate phoneme. Conversely, [ɾ] is an allophone of /d/ in English while /ð/ is a separate phoneme. There is a large amount of literature showing that sounds that contrast in one's language are more perceptible than those that do not. Boomershine et al., (2008) showed this to be true of the [d], [ð], and [ɾ] sounds in native Spanish and English speakers, where Spanish speakers more easily perceived the differences between [d]/[ɾ] than [d]/[ð].

English speakers, on the other hand, had difficulty distinguishing [d] from [ɾ], but no issue with [d] and [ð]. The studies in this dissertation extend this work by first replicating the results of Boomershine et al., (2008) with a group of monolingual English speakers as well as a group of native Spanish speakers on a forced choice perception task. Additionally, I add a group of native English-speaking advanced learners of Spanish and show that their perception of the relevant sounds is more like that of the native Spanish speakers on a number of behavioral metrics. In a second study, I test the same three speaker groups in an Electroencephalography study (EEG) using the mismatch negativity protocol (MMN) which has been previously shown to probe auditory categorical perception (Näätänen et al., 1978). If the MMN is sensitive to phonemic contrasts, as has been claimed, the expectation is that, similarly to the behavioral perception results, speakers should show larger mismatch responses to phonemic contrasts than to allophonic contrasts. I also explore the possibility of using the MMN to probe category formation in the learner group. However, while there are subtle differences in the EEG data for each speaker group, the results are contra predictions. Instead, all three language groups pattern similarly on each sound in the EEG study, with larger MMNs being elicited by the [ɾ]/[d] comparison than the [ð]/[d] comparison. This mismatch response suggests that the MMN is not probing phonological status but is sensitive to phonetic category. A discussion of methodology and the validity of using the MMN in phonology research is include

For Ozzie

ACKNOWLEDGEMENTS

It has been a wonderful journey. I've been lucky enough to pursue my own research interests, I've had access to wonderful academic resources, I've sat in lectures by the preeminent linguists and cognitive scientists of the day, and more often than not, there was free pizza afterward. But graduate school takes something out of you. It vaporizes time, crushes personal economics, and stunts social growth. I cannot say for sure it was worth it. Ask me in twenty years. I don't know anyone who's been able to grit their way to a PhD. At least, whenever I read one of these things, it's the people surrounding the candidate that apparently did all the heavy lifting. It has been no different for me. With that in mind I'd like to acknowledge those who kept me going and ultimately pushed me over the finish line. First and foremost, my family; my Mother Vicki Trotter, Father Jim Trotter, and Stepmother Janet Trotter. My cousin Nicholas Oas and his wife Carrie Waggoner, both fellow Spartans who kept me well fed and socialized throughout the years. My family's love and encouragement were always felt. My best friends Stephen Wright and Nate Graham, who let me keep one foot outside of the academy. My lovely partner Alba, who turned out to be the most surprising plot twist of all. She kept me tethered to planet Earth in the final weeks of frantic writing. I'd like to especially recognize the members of my cohort, Joe Jalbert and Ho-Hsin Huang, who due in no small part to our collective misery became invaluable friends. Ours, I think, was a special cohort; small in number and tightly knit.

Joe and Ho-Hsin got hitched halfway through this journey, and I was honored to stand up with them on that day. They now have a beautiful little girl, and while all of our lives have changed dramatically from the start of graduate school, I'm thrilled to have had such great friends accompany me to the end. My committee members, Yen-Hwei Lin whose thoughtful comments on all of my work were so appreciated, Suzanne Wagner who should know her words of encouragement throughout my graduate career were always timely, Bill VanPatten whose humor always lightened the air in the room, especially when reverence was expected. Finally, my co-advisors Karthik Durvasula and Alan Beretta. Karthik, who against his own better judgement became a friend first and advisor second. He always knew exactly when to give praise and encouragement, and when to lay on the pressure. I owe him much for this. Alan has been a mentor and friend since I was an undergraduate. He endured missed deadlines and endless stress all while pulling every string he could find to support me. There is simply no way this dissertation could have been completed without his unwavering support, relentless positivity, and advocation on my behalf. I'm most proud about what we accomplished in the neurolinguistics lab. The lab started out as Alan and I presiding over a bunch of boxes filled with equipment of foreign origin. Our collective experience with EEG close to zero, we decided to have a go at it anyway. Now, three generations of students have been brought up in the lab, measuring craniums, ruining perfectly coiffed hair, and producing some great science along the way. We've been part of a wonderful

international collaboration with the University of Iceland. Most importantly, the lab has been a welcoming place for students from all backgrounds. We've had students in syntax, semantics, phonetics, phonology, and sociolinguistics participate and produce by incorporating EEG into their work. All are invited to join and free to tinker as much or as little as they see fit. Alan has masterfully engineered the lab meetings to be a supportive and collaborative space where curious undergrads up to veteran PhDs engage one another fully. I am humbled to have been a part of it.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Introduction

Sounds that differ from one another in a language can be categorized as either contrastive, sounds that native speakers of said language can readily distinguish from one another and easily produce, or non-contrastive, sounds that native speakers have difficulty distinguishing from one another as well as producing. These contrastive sounds, or phonemes, constitute the phonemic inventory of the language and are represented in the mind of the speaker. The world's languages can differ in the number and type of sounds they make use of in their phonemic inventory such that one language may contain phonemes that another language may lack. This property of language is a major contributor to the difficulty learners encounter when acquiring a second language (Flege, 1995) . The non-contrastive sounds can either surface within a particular language via a process of phonological rule application, where they are called allophones, or they may be unrepresented entirely within the language. In either case, learners have difficulty perceiving and producing the sounds that do not exist or have differing phonemic status in their native language. These phonemes and allophones are found at different levels of representation. In the generative view, there exist at least two levels, the underlying level and surface level. There may be multiple derivative iterations before arriving at the surface form, but we can think of phonemes being at the underlying level and allophones being surface representations (Kenstowicz, 1994;

McCarthy, 2010). The differences between contrastive and non-contrastive sounds in different languages is readily illustrated when looking at English and Spanish. There are slight differences in place of articulation of [d] between the languages. The [d] is alveolar in English and surfaces as dental in Spanish (Martínez-Celdrán et al., 2003). Throughout the dissertation, I'll be referring to both simply as [d]. In English both the voiced alveolar stop sound /d/ as in *doze*, and voiced interdental fricative /ð/ as in *those* are contrastive. However, [ð] is an allophone of /d/ in Spanish, where it surfaces only after a spirantization rule applies (Harris, 1969). Conversely, in Spanish, the sounds /ɾ/ and /d/ are separate phonemes (though there are no minimal pairs, or lexical items whose pronunciations are distinguished solely by the two sounds), while in English [ɾ] is an allophone of both /d/ and /t/ (Ladefoged & Johnson, 2011). An English-speaking learner of Spanish must learn to actively contrast /ɾ/ and /d/, while a Spanish-speaking learner of English must do the same with /d/ and /ð/, a task both students often do with great difficulty. In this first chapter I lay out the facts concerning the sounds of interest in both English and Spanish and how those facts motivate the present study. I give an overview of the project and the experiments within, and I outline the content of the remaining chapters for the reader. I'll begin with the basics of the phoneme and allophone distinction on which the dissertation relies.

1.1 Description

This dissertation is concerned with how a learner's perception of both native and non-native sounds changes as experience in a second language grows. The studies within specifically ask whether a native English-speaking learner of Spanish is able to form a phonemic representation for [ɾ] like that of a native Spanish speaker. They also ask how learning the allophonic rule for [ð] in Spanish alters the perception of [ð] in by the native English speaker. There is some debate on how a second language is manifest in the brain, particularly whether the sounds from the native language (L1) and the sounds from the second language (L2) are discrete systems or if they share some or all storage and retrieval properties (Paradis, 1985). If the sound systems in L1 and L2 interact, then we can expect to find bi-directional influence from both languages within a speaker (Cook, 1992; Selinker, 1974). Furthermore, this influence should change based on the proficiency level in the L2. A key question in second language learning is whether or not learners create new phonemic categories for L2 contrasts, or if they expand the phonetic space of their L1 via some memory mechanism to include the relevant L2 sounds, while mapping them phonologically to some formerly acquired L1 phoneme. The experiments in this dissertation are designed to target the phonological level of representation in the mind of the learner, and to look for differences in categorical perception between L1 and L2 contrasts. In short, the goal is to observe the formation of new categories of sounds in a learner and how this acquisition may affect the L1 representations.

The experiments herein utilize the mismatch negativity response (MMN) which is an automatic response to categorical distinctions measured using electroencephalography (EEG). The MMN is a negative deflection in the difference wave when subtracting the average ERP waveform of a "standard" stimuli from the average ERP waveform of a "deviant" stimuli. A typical MMN paradigm consists of repeated "standard" stimulus and an oddball "deviant" stimulus in a 7:1 ratio (Kilner et al., 2009; Näätänen & Alho, 1995, 1997; Näätänen, et al.,2004). These stimuli are usually categorically different in some way (e.g. different phonemes). When the "standard" is repeated, it builds a trace for that category in memory. Once this memory has been built up in the auditory cortex, it is possible to interrupt the stream of "standards" with a deviant, resulting in the elicitation of an MMN. The "deviant" sound is compared to the memory trace of the "standard" and if found to be different, will elicit a characteristic ERP waveform on the fronto-central electrodes about 150ms post stimulus. The ability to reliably detect categorical distinction in speech perception allows us to track changes in categorical perception of speech sounds in a learner. A detailed account of the MMN component, and its history in speech perception work is provided in Chapter 2. With that in mind, the big picture questions that this dissertation attempts to answer are:

- Do we see evidence of L2 category formation in second language learners? I argue that we do see this in the behavioral study outlined in Chapter 3, but that the electrophysiological results are less informative here.
- Do we see evidence that perception is modulated by level of representation? Again, the behavioral evidence points to this being true.
- Are we able to use the Mismatch Negativity to probe phonemic (underlying) representations? The results of the electrophysiological study in Chapter 4 suggest that the MMN probes phonetic category but is not sensitive to phonemic category.

1.2 Outline of Dissertation

Chapter 2 motivates the need for the studies in the dissertation, introduces the relevant sounds in each language group, and details the past work on which my studies are built upon. I'll start with a descriptive account of the [d, ð, and ɾ] sounds in both English and Spanish. Following that, review the literature on acquisition of the Spanish /ɾ/ phoneme and the allophonic rule for [ð] by native English speakers, and discuss theories of second language acquisition that may guide our predictions. Both behavioral and electrophysiological data are reviewed. Both studies in this dissertation use three groups of speakers. There is a group of monolingual English speakers, a group of native Spanish speakers, and a group of native English-speaking advanced learners of

Spanish. Chapter 3 is dedicated to a behavioral study where these groups are subject to a timed discrimination task, which while not particularly novel sets a baseline of performance to help interpret the results of the second study, detailed in Chapter 4. The electroencephalography study in Chapter 4 uses the same speakers as the behavioral study in a passive auditory mismatch task, allowing us to see any automatic or pre-attentive responses to the sounds of interest. The tasks, their methodologies, and results and statistical analyses are detailed in their respective chapters. Finally, Chapter 5 provides the conclusion, discussion, and directions for future work. It contains an interpretation of the behavioral results that clearly show the advanced Spanish learners becoming more like the native Spanish speakers in their perception. I also discuss the EEG results which do not necessarily align with the behavioral results. I find an unexpected mismatch response to the allophonic [d] - [ɾ] comparison among the monolingual English speakers, along with a stronger mismatch response on both the allophonic and phonemic comparisons in the advanced Spanish learners. In fact, both comparisons pattern similarly for all three language groups in the EEG experiment, while exhibiting clear differences in the behavioral study. Comparisons to similar EEG work and possible explanations are discussed.

Chapter 2: Motivation and Prior Work

In this chapter, I will first discuss the status of the three sounds focused on in this dissertation in both English and Spanish ([d], [ɾ], and [ð]). There is asymmetry in how these languages classify the sounds in terms of phonemes or allophones, allowing us to study the effect allophony has on perception. I discuss the difficulties English speakers encounter when learning to correctly utilize both the flap and fricative in Spanish, and I then turn to discussing two current models of second language acquisition and what predictions each makes for the current study. Finally, I discuss how the mismatch negativity experimental paradigm can potentially bring further clarity to this research by manipulating the level of representation that a listener taps into during an experiment and highlight previous electrophysiological work on the acquisition of L2 sounds.

2.1 Introduction

For native speakers of English, learning Spanish involves altering their representations of two sounds in their L1 (these are not the only sounds that differ between English and Spanish, but are the ones studied in this dissertation). The studies herein are focused on the discriminability of the [d], [ð], and [ɾ] sounds in English and Spanish. There are slight differences in place of articulation of /d/ between the languages. The /d/ is alveolar in English and dental in Spanish (Martínez-Celdrán et al., 2003). Spanish contrasts the voiced alveolar stop /d/ and the alveolar flap /ɾ/ phonemically while in English, [ɾ] is an allophone of both /d/ and /t/ (Harris, 1969; Ladefoged & Johnson, 2011). Figure 1 illustrates the underlying and surface levels of representation for the [d], [ð], and [ɾ] sounds in English and Spanish.

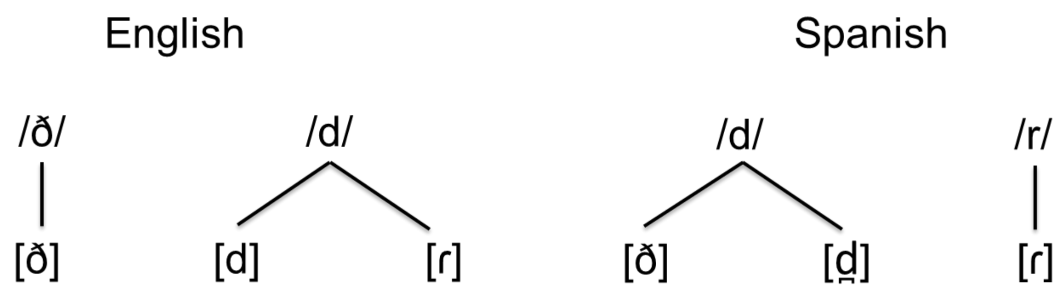

Figure 1: Underlying and surface forms for [d], [ð], and [ɾ] sounds in English and Spanish.

English speakers learning L2 Spanish must learn to contrast /ɾ/ with /d/. A key question that this dissertation attempts to address is whether a learner can form a new category for /ɾ/ or if proper /ɾ/ perception by an L2 Spanish speaker does not rely on the creation of a phonological representation of /ɾ/. English speakers also must learn the spirantization rule in Spanish, whereby stops become fricatives after sonorants. In Spanish, /d/ surfaces as the voiced dental fricative [ð] in the same context that /d/ surfaces as [ɾ] in English further complicating acquiring this allophonic rule for Spanish learners. /ð/ already exists as a phoneme in English so learners of L2 Spanish must learn to use the allophonic representation simultaneously with the phonemic representation.

2.1.1 The Status of d, ð, and ɾ in English

In English, /d/ and /ð/ are different phonemes while [ɾ] is an allophone of /d/ and /t/. The voiced dental fricative contrasts with the voiced alveolar stop as in *those* [ðoz] and *doze* [doz]. However, the alveolar flap [ɾ], only surfaces as an allophone of /d/ and /t/ post sonorant and before unstressed vowels as in *writer* or *reader* (Kahn, 1976; Zue & Laferriere, 2005).This flapping rule is extremely productive in American English, being realized over 90% of the time the environment allows it (Herd et al., 2010; Patterson & Connine, 2001).

2.1.2 The Status of d, ð, and ɾ in Spanish

In Spanish, unlike in English, /d/ and /ɾ/ contrast phonemically. There are no minimal pairs because /d/ is spirantized to [ð] after sonorants. As with flapping in English, Spanish spirantization, where voiced obstruents /b,d,g/ are spirantized to [ß,ð,ɣ], is a highly productive phonological rule in Spanish, with spirantization of /d/ occurring 99% of the time (Waltmunson, 2005). This dissertation is concerned with the contrast of the [d] and [ð] sounds, which are allophonic in Spanish. /d/ undergoes spirantization to [ð] unless it follows a liquid or nasal as in (1a)-(1b).

(1)

a. /hada/ -> [aða] 'fairy'
b. /falda/ -> [falda] 'skirt'

Since /ð/ contrasts with /d/ in English and since Spanish spirantization occurs in a similar environment as English flapping, this difference in how [ð, d, ɾ] are phonemically categorized in the two languages may cause difficulties for English learners of Spanish. The following portion of the chapter details the difficulties reported for English speakers acquiring the /d-ɾ/ contrast in Spanish.

2.2 Spanish Acquisition in English Learners

This section details English-speaking Spanish learners' acquisition of /ɾ/ as a meaningful contrast as well as the spirantization rule for post-sonorant Spanish voiced stops.

2.2.1 Acquisition of Spanish /ɾ/ for English Speakers

English speakers produce [ɾ] as an allophone of /t/ and /d/ after sonorants and before unstressed syllables, but must learn to contrast it with /d/ in Spanish. Face (2006) showed that at least in terms of production, this contrast is very difficult for beginner and even intermediate native English-speaking learners of Spanish. In a reading task, 41 native speakers of English were placed into two groups: one group of intermediate Spanish learners who had taken four semesters of college-level Spanish, and one advanced learner group who were Spanish majors and enrolled in upper-level Spanish courses. Their production was compared to a group of native Spanish speakers. The native Spanish speakers produced the [ɾ] correctly 92% of the time, while the intermediate Spanish learner group did so only 49% of the time. The advanced Spanish learners approached the native Spanish group with 79% correct production of [ɾ]. When errors happened, both learner groups tended to replace the [ɾ] with the American English retroflex [ɹ].

2.2.2 English-speakers acquisition of Spanish spirantization

English-speaking L2 learners of Spanish have to acquire a number of phonetic and phonemic differences in order to sound like a native Spanish speaker. One of these differences is Spanish spirantization of the voiced obstruents /b, d, g/. Previous studies on English-speaking novice and intermediate learners have shown that production of the spirtantized forms is rarely achieved. Zampini (1994) found that voiced obstruents were spirantized only 32% of the time. Of the three /b, d, g/, the /d/ to [ð] spirantization was the least likely to be achieved at less than 10% of the expected instances. In her study, when failing to produce the fricative, the voiced stop was always produced instead. Zampini claimed that the existence of the phoneme /ð/ in English is what causes the /d/ to [ð] spirantization to lag behind /b/ to [ß] and /g/ to [ɣ] in learner's production. Herd (2011), however, offered that it is possible that the learners were producing the alveolar flap [ɾ] intervocalically instead of [ð] since that is the environment in which it is already produced in English. While Zampini's studies are informative for new learners of Spanish, they do not include advanced learners or bilinguals in their experiments. A study by Face & Menke (2009) looked at the Spanish production of 53 English speaking L2 Spanish learners broken into three groups based on Spanish instruction level. While they found similar results to Zampini among the less experienced groups, their advanced learner group showed significant development toward native-like rates of spirantization. The most advanced group had an average of thirteen years of Spanish instruction and fifteen months of time in a Spanish-speaking country. These learners produced spirants in 80% of the expected contexts. This suggests that this spirantization rule can be learned in adulthood with enough training.

2.3 Theories of Second Language Acquisition

This difficulty in acquiring L2 contrasts and correctly applying phonological rules is of course not limited to English speakers learning Spanish. It is well established that listeners perceive differences in sounds that are present in their native language more easily than those that are not (Best, et al., 1988; Dupoux et al., 1997; Strange, 1995; Strange & Jenkins, 1978, amongst others). Errors in pronunciation of non-native words by non-native speakers of a language are the hallmark of this difficulty. For many L2 learners, these errors persist throughout life, even when they are conscious of them, and continuously corrected. One of the key debates in the second language literature, and an important question asked in this dissertation, is whether or not a L2 learner can form non-native phonemic categories late in life. The Critical Period Hypothesis (Birdsong &

Molis, 2001; Johnson & Newport, 1989; Lenneberg, 1967) states that a learner must be exposed to language stimuli before a critical age (sometime before the completion of puberty) in order to attain native-like performance in that language. The strong version of the hypothesis would claim that L2 learners, beyond the critical period, have no hope of becoming native-like in their second language, at best approximating a native speaker. Advocates of this view point to the strong correlation between foreign-accented (FA) speech and either age of arrival (AOA) or first exposure to the L2, something Flege et al., (1995) showed can account for 94% of the variance in FA rating among L2 learners. However, in their view, AOA correlates with a plethora of other variables on L2 input and usage that confound the data and the claim that failure for the majority of L2 learners to achieve native-like speech production and perception is due to neurological development. Most theories of second language acquisition reject a strong critical period hypothesis, and there is evidence that a decline in language learning ability may be tied to factors other than the onset of puberty. In a massive study of native and L2 English speakers, Hartshorne et al. (2018), had over half a million subjects take an English grammar quiz online along with a demographic questionnaire. Test results along with demographic information and linguistic background information were used to create means for each subject's native language and dialect. This data was used to train a computational model for second language learning. The results suggest that L2 acquisition ability remains intact at least into early adulthood before beginning to decline around age 18. They claim that the existence of a critical period for second language acquisition, if it exists, cannot be due to changes in neuronal development or hormones in childhood or adolescents. Instead, it may be due to changing social pressures in adulthood, neuronal changes later in life, or increased interference from the L1.

2.3.1 Speech Learning Model

Flege's Speech Learning Model (SLM) (Flege, 1991, 1995, 2007, 2011) leaves open the possibility for learners to create new categories. The ease of forming L2 categories rests on dissimilarity of the L2 phoneme from existing L1 phonemes. Flege's Speech Learning Model rejects the critical period hypothesis. While "earlier is better" is uncontroversial in any model of L2 learning, the SLM claims that there is no neuro-biological mechanism that closes off language learning, and that the same acquisition mechanism children use to learn their L1 is preserved across the lifespan (Flege, 2007). For the SLM the type, frequency, and quality of input can also explain the bulk of the variance in FA speech outcomes. When it comes to the acquisition of nonnative contrasts, the SLM makes predictions about the ease of learning those contrasts based on how similar (rather dissimilar) the L2 sounds is from an existing L1 sounds. The SLM predicts that initially, an L2 learner will substitute the unfamiliar L2 sound with the closest (in terms of phonetic similarity) L1 sound. As experience in the L2 grows, these sounds will gradually dissimilate from each other until a new category is formed. SLM additionally predicts that there may be change in the original L1 category as the creation of the new L2 category crowds the phonetic space, causing a dispersive shift that might perceptibly alter the L1 category, making it markedly different from that of a monolingual. As it relates to the sounds studied in this dissertation, the SLM would predict that learners would initially replace [ɾ] with a perceived similar L1 sound (/ɹ/), but that advanced learners would eventually be able to create a category for /ɾ/.

2.3.2 Perceptual Assimilation Model

Like the Speech Learning Model, the Perceptual Assimilation Model (Best, 1994, 1995; Best et al., 2001; Best & Tyler, 2007) relates the ease with which a learner will acquire an L2 sound by how readily the learner is able to assimilate that sound to an existing L1 sound. Under this framework, non-native sounds are perceived in one of three ways: They can be assimilated to an existing L1 category as a good exemplar, perceived as a speech sound but not a good exemplar of an existing L1 sound, or perceived as non-speech. If the L2 sound is perceived as a good exemplar of an existing L1 sound, the learner will have difficulty creating a new category for that sound and the likelihood of learning the L2 sound will be small. However, if the L2 sound is not readily perceived as an exemplar of an existing category, the learner should be able to perceive the difference and potentially form a new category for that sound. PAM is a

Direct Realist theory (Fowler, 1986, 1989, 2005), which assume that phonological primitives are articulatory gestures that have physicality (Browman & Goldstein, 1986, 1989). This differs from the SLM, which assumes that representations of made up of acoustic cues. PAM also makes explicit distinctions between phonetic and phonological categories, whereas the SLM does not. For PAM, category creation would require a shift in the L1 exemplar in order for the target L2 sound to not assimilate to that L1 category.

Under both models, the possibility for adults to acquire new categories exists, though the underlying mechanism for doing so differs. While the experiments in this dissertation are not designed to test specific models of second language acquisition, I mention these because the ability for adults to form new L2 categories and their relation to the L1 are central to the predictions tested.

2.4 Mismatch Negativity

The Mismatch Negativity (MMN) is a component of the auditory event-related brain potential (ERP) that responds to change (Näätänen, et al., 1978; Näätänen & Michie, 1979). Though originally discovered using encephalography (EEG), equivalent components have been found using other brain recording methods such as magnetoencephalography (MEG) (Csépe et al., 1992; Hari et al., 1984), functional magnetic resonance imaging (fMRI) (Celsis et al., 1999; Molholm et al., 2005), and positron emission tomography (PET) (Tervaniemi et al., 2000). The "classic" MMN manifests as a negative deflection in the difference wave obtained by subtracting the ERP generated by a frequent “standard” stimulus from that of an odd-ball “deviant” stimulus. It is generally seen on the frontocentral electrodes, peaking at approximately 150ms post onset of the deviant stimulus. The MMN component overlaps with the so-called N1-P2 complex, which is a negative deflection in the ERP occurring 100ms post stimulus followed by a positive-going deflection around 200ms post stimulus (Näätänen et al., 1989; Sams et al., 1985; Tiitinen et al., 1994). These components are so-called “obligatories”, regularly occurring with any stimulus, but can be differentiated from the MMN by looking for reversed polarity of the MMN on the mastoid or temporal electrodes (Deacon et al., 2000). Generation of the MMN depends upon the creation of a “standard” in short-term memory. This memory trace is ephemeral, so if the “standard” stimuli are presented in too low a number or with too long an interstimulus interval (ISI), it cannot be maintained, and no MMN will be generated upon presentation of the “deviant” stimulus (Mäntysalo & Näätänen, 1987). The presentation of the “standard” stimuli can also affect the MMN response. The more consecutive “standard” stimuli that are presented, the larger the response to the “deviant” will be. Likewise, the shorter the ISI between standards, the larger the MMN amplitude (Sabri & Campbell, 2001). Thus, the ratio of standards to deviants in the experiment presentation matters. A typical MMN paradigm consists of repeated standard stimulus and an oddball deviant stimulus in a 7:1 ratio (Kilner et al., 2009;

Näätänen et al., 2007, for reviews). While the MMN is strongest on frontocentral electrodes in EEG, the cortical source of neuronal activity cannot be straightforwardly inferred from scalp-distributions due to interference from the skull. Other recording methods have been informative here, with intracranial, fMRI, and PET recordings suggesting that the MMN has its origin in the auditory cortex, specifically Broadman areas 22, 41, and 42. (Baudena et al., 1995; Halgren et al., 1995; Halgren et al., 1998). There are a couple of ways to set up an MMN experiment but in general, the simplest way to do this is with two tokens. A single standard and deviant. A variation that has been used increasingly in language related studies is to use a variable standard (Eulitz & Lahiri, 2004; Phillips et al., 2000). That is to use multiple similar tokens (so a few different versions of [sa]) and then a few different versions of [za]. You still present the standard and the deviants in the same 7:1 ratio, but because the tokens aren't identical, the standard is less reliant on acoustic similarity and hopefully more on phonetic category. There are also two ways you measure the response. I use both in my dissertation. First is the classic MMN, where you simply subtract the voltage of the deviant from the voltage of the standard. The second way, sometimes called the identity MMN, is based on presentation context. You subtract the voltage of on stimulus when it is presented as a deviant, from the voltage of that same stimulus when presented as a standard. The idea is that the iMMN controls for the mismatch response generated simply from the differences between the two stimuli (Pulvermüller & Shtyrov, 2006;

Tavabi et al., 2009). Figure 2 shows the single and variable standard paradigms as well as illustrating the difference between the classic and identity MMN. In the first two blocks of stimuli, there is only a single deviant along with a single standard token presented seven times the number of deviant presentations. In the second block, the variable standard, multiple different tokens of both the standard and deviant sounds are presented in the same 7:1 ratio of standard to deviant presentations. In the “Classic” MMN measurement, you simply subtract the mean voltage of the deviants in a block from the mean voltages of the standards in a block. Alternatively, with the “iMMN” you subtract the mean voltage of a sound when it is presented as a deviant in one block, from the mean voltage of that same sound when it is presented as a standard in another block.

[sa] [sa] [sa] [sa] [sa] [sa] [sa] [za]…
[za] [za] [za] [za] [za] [za] [za] [sa]…
Single standard and deviant

[sa_1] [sa_2] [sa_3] [za_1] [sa_2] [sa_1] [sa_1] [sa_3]…
[za_1] [za_1] [za_3] [sa_2] [za_2] [za_2] [za_3] [za_1]…
Variable standards and deviants

Two ways to measure response:
“Classic” MMN – [za] subtracted from [sa]
“Identity” MMN – [za] as a deviant from [za] as a standard

Figure 2: Single token and variable token MMN paradigms. The variable standard varies the tokens within the same phonetic category, minimizing the effect that acoustics plays in building the standard.

2.5 Previous MMN Work

2.5.1 Cross-linguistic

Näätänen et al., (1997) showed that responses to linguistic stimuli in a mismatch paradigm were both distinct from non-linguistic acoustic stimuli and language specific. In an experiment run on both Finnish and Estonian subjects, they found that in the frequent presence of a prototypical Finnish phoneme /e/), an infrequent stimulus (MMN deviant) of either another prototypical Finnish phoneme /ö/ or the Estonian phoneme not found in Finnish /õ/, had different MMN responses. In the Finns, the MMN response to the Finnish deviant /ö/ was enhanced in amplitude when compared to the response to the Estonian phoneme. Crucially, when Estonian speakers were subjected to the same experiment, they did not show a difference in amplitude for deviant /ö/ when compared to deviant /õ/. They claim that this language-dependent MMN effect shows the existence of language-specific neural traces to phoneme representations. They also found that, the MMN elicited in the Finns by the vowel existing in Finnish, was localized to the left posterior auditory cortex while the MMN generated by the Estonian vowel were bilaterally generated. This suggests subcomponents of the MMN that are sensitive to the differences between purely acoustic and phonological information.

Phillips et al. (2000) in a magneto-encephelography (MEG) study, showed that the MMN can be used to show evidence for abstract phonological representations. The MMN has a magnetic analog that can be elicited using MEG as opposed to EEG. While this is sometimes still called the MMN in the literature, the MEG specific component is commonly called the MMF for Mismatch Field (Hari et al., 1984). Where previous mismatch studies tended to use a single exemplar token for their standard and deviant sound, (Phillips et al., 2000) varied their standard tokens such that the construction of a standard would be dependent on recognizing the phonological features of the segment. In this case, they presented English speaking subjects with multiple tokens of /d/ and /t/ where the VOT was varied within each category. Because the standard is no longer acoustically the same stimulus in this design, any memory trace must be based on the phonological category /d/ or /t/ rather than simple acoustics (Aulanko et al., 1993). They observed an MMN in this varied standards presentation. In a second phonetic experiment, they increased the VOT of all their stimuli by 20ms, causing half of the previous standards to fall just above the categorical perception threshold for voicing and into the /t/ category. So, while the stimuli were still presented in a 'many to one' ratio in terms of relative VOT length, the number of /t/ and /d/ stimuli were equal. No MMN was observed in the second experiment, suggesting that no standard was able to be created in the listener based on VOT length, and that the MMN is sensitive to phonological category. This varied standard technique has since been utilized to probe even more granular phonological information by others. Eulitz & Lahiri, (2004) and Hestvik & Durvasula, (2016) have used this paradigm to test predictions for underspecification theory.

Kazanina et al., (2006) expanded on this work to show that not only are there language-specific responses to speech sounds, but that those responses can differ based on the functionality of the sounds and contrasts within the language. Russian speakers contrast [d] and [t] phonemically across a voice-onset time (VOT) continuum. While Korean speakers also produce [t] and [d] across a similar VOT continuum, they do not contrast these sounds phonemically. In Korean, [d] and [t] map to a single phoneme realized as [d] between voiced segments and [t] elsewhere. While it had been well established in behavioral studies that discrimination of allophonic contrasts is poorer than in phonemic ones, and thus functional status matters in speech perception, these behavioral differences could have arisen via higher-level conscious processes. Establishing electrophysical results via the MMN would show that these functional differences in speech sounds with a language were available preattentively.

Miglietta et al., (2013) ran an MMN study on phonemic perception versus allophonic perception in Italian speakers. In their study, they found distinct MMNs were elicited for both phonemic and allophonic contrasts. They looked at the allophonic variation [ɛ-e] and phonemic contrast [e-i]. They found no amplitude differences between the MMN responses to these contrasts, but did find significant differences in peak latency. The phonemic contrasts had earlier peak latencies than did the allophonic alternations, suggesting easier processing of the phonemic contrasts.

2.5.2 Work on L2 Learners

Comparatively little work has been done utilizing the MMN in second language acquisition research. Of the MMN work that has been done, the majority focuses on bilingual populations or change detection in children within different language environments (Cheour et al., 1998; Cheour, et al., 2001; Cheour, et al, 2002; Peltola et al., 2005; Shestakova et al., 2003). Relatively little has been done in the way of adult second language learning, though there is evidence that the MMN can probe category formation in adult second language learners.

Winkler et al., (1999) was one of the earliest studies to provide evidence for category formation in second language learners. In their study, native speakers of Finnish, native speakers of Hungarian, and native Hungarians who were second language learners of Finnish were tested on the /e/-/ä/ contrast. These sounds are allophones in Hungarian but distinct in Finnish. The naïve Hungarian speakers showed no MMN for this contrast and similarly could not distinguish the sounds behaviorally. Native Finns showed a distinct MMN for this contrast and easily distinguished the sounds in a behavioral task. Importantly, the fluent L2 Finnish speakers showed both MMN and behavioral results similar to the Native Finns.

Peltola et al. (2003) achieved slightly different results, showing that the MMN amplitude does vary within second language learners as fluency changes. Their study looked at vowel discrimination in native Finnish learners of English. Native English speakers, Native Finnish speakers, and Native Finns with advanced English training were all played blocks of Finnish /i/ - /e/, English /ɪ/ - /e/, English /i/ - /ɪ/, and English /e/ - /ɪ/ in an MMN paradigm. There is no phonemic /ɪ/ vowel (as in "bit" in English), in Finnish. The results as shown in the difference waves suggest two things. First, it does not appear that the "advanced" learners of English group perform like native English speakers on English contrasts. The authors assume this to mean that phonological category formation has not yet taken place in these learners. Second, these learners display a significantly lower amplitude MMN for the Finnish vowel contrast than do the native Finnish speakers with no English training. The claim here is that perhaps the learning process has somehow degraded their perception of L1 constants. Others have reported bidirectional influence of L2 learning on the L1 in production experiments (Mackay & Flege, 2004; Yeni-Komshian et al., 2000) but to my knowledge this is the first instance of neurophysiological evidence of that influence on perception.

Herd (2011) investigated whether English speaking learners of Spanish could be trained to both perceive and produce the [d], [ɾ], and [r] contrasts in Spanish. She tested participants on each contrast before and after a number of training protocols. While primarily concerned with understanding which types of training produce the most effective results, she used both traditional behavioral identification and production tasks along with electroencephalography. This ERP data can show whether there is automatic processing of the contrasts taking place, and if the training protocols are enough to show evidence for phonemic category formation in the learners. She ran an MMN task on two groups of speakers. One native Spanish speaking group and another group of English speakers who had undergone perception training on the Spanish contrasts in question. For the English trainees, the MMN task was administered twice, both pre and post training. As one would predict, the native Spanish speakers showed a clear MMN response to the [d]-[ɾ] comparison. Surprisingly, the English trainees also showed a significant MMN to the [d]-[ɾ] comparison in both the pre and post training environment. While the mismatch response to [ɾ] in the pre-training English speakers is not predicted, it is possible that the learners had already created that category. My EEG study finds similar results, with a large MMN response to /ɾ/ among Spanish learners. Additionally, and more surprisingly, the monolingual English speakers in my study (not tested by Herd), also showed a large MMN response to the [d]-[ɾ] comparison. This makes it harder to interpret the results of Herd and my learners as having created a new phonemic category for /ɾ/.

Barrios (2013) and (Barrios, et al., 2016) ran a very similar study to the second study in this dissertation. In their study, they looked at L1-Spanish speakers who were advanced L2-English learners as opposed to the advanced L2-Spanish learners here.

Their study also utilized magnetoencephalography (MEG) rather than electroencephalography (EEG) as in the current study. Due to these similarities, the predictions in their study and mine are more or less aligned and thus comparing results is informative. They found no statistical significance in the differences in groups, but noted that the Mismatch Field (MMF, the MMN correlate generated by MEG, Hari et al., 1984) amplitudes elicited were contra expectations. In their study, English speakers both showed similar MMF responses for both the [d]-[ɾ] comparison and the [d]-[ð] comparison, while Spanish speakers had similar MMF responses to English speakers on the [d]-[ɾ] as well as an unexpected contrastive MMF for [d]-[ð]. These MEG responses largely mirror the results of my EEG study detailed in Chapter 4.

In this chapter I discussed the sounds of interest and talked about how they are represented differently in both English and Spanish. I summarized the acquisition work on English-speaking learners of Spanish for the flap and spirantization rule and discussed the major theoretical framework in second language acquisition. I introduced the mismatch negativity paradigm and how it may be used to probe categorical perception, specifically phonetic and phonemic categories. I reviewed the literature on the MMN as it pertains to linguistic contrasts as well as the MMN literature on second language learners and bilinguals. In summary, there is strong evidence that within a language, the MMN is sensitive to phonological status in speech sounds. In some cases, MMN responses tend to be stronger for sounds that are contrastive in the language than for those that are not. Allophony seems to play a part in the mismatch response. In some cases, it results in attenuation of the amplitude of the MMN, where in others it seems to affect the peak latency of the response. This dissertation contributes to the growing body of work on advanced L2 learners and bilinguals and their neurophysiological responses to native and non-native contrasts. Here, the literature has been considerably less convergent, highlighting the need for additional work in this area. In the next chapter, I will discuss a psycholinguistic study testing the discrimination of [d], [ɾ], and [ð] among three groups of speakers. Naïve English speakers, Native Spanish speakers, and English-speaking advanced learners of Spanish. In Chapter 4, I will present a mismatch negativity study on those same subjects to see if any insights on their speech perception can be gained during the pre-attentive processes that the behavioral study in Chapter 3 is unable to tease out.

Chapter 3: Behavioral Task

3.1 Introduction

The following chapter discusses in detail the behavioral task given to each of the subjects in my three language groups. This chapter will establish the perceptual facts as they fall out from the AXB discrimination task. This will create a set of comparison results to check against the electrophysiological results presented in Chapter 4. One of the reasons to pursue neurolinguistic methods such as the MMN with tools like EEG or MEG lies in the belief that these tools and methods can reveal insights about language processing that the relatively high latency of behavioral studies cannot. Second, replicating prior results on the perceptual sensitivities between the relevant sounds for the three language groups is important as both a sanity check and to ensure that our participant populations are reasonably categorized. There are robust results on the perceptibility of contrastive versus non-contrastive sounds within speakers (Beddor & Strange, 2005; Boomershine, et al., 2008), but here I claim that the simple presence or absence of a sound from a speaker's phonemic inventory is not sufficient to predict discriminability. Like Boomershine et al. (2008), I claim that the more nuanced phonological status of the sounds is what determines perceptibility. That is, both allophony and contrast need to be taken into account. To that end, replication of the

Boomershine results (with a slightly different discrimination task) will give needed weight to this claim. I present the within-language results for each language group (monolingual English speakers, native Spanish speakers, and advanced Spanish learners) in turn. I then look at the cross-language comparisons. I will present the behavioral results for each group in three different ways. First, a simple accuracy rating. This is straightforwardly the percentage of trials in which each subject correctly identified the target sound in the AXB task. Second, a sensitivity measurement, in this case A′, which will control for any response bias in the subjects. I use A′ instead of the more widely used d′ (Macmillan & Creelman, 2004) as an estimate of sensitivity because the A′ calculation, while not truly non-parametric, is less reliant on assumptions about the distribution of the signal and noise curves (Donaldson, 1993; Zhang & Mueller, 2005; Stanislaw & Todorov, 1999). Finally, I will also present an analysis of the reaction times from the offset of the last stimulus in the AXB task.

3.2 Methods

Every participant took part in a behavioral task immediately prior to the EEG experiment. This task was designed to get a baseline perceptual discriminability of the three crucial sounds in the dissertation for each language group. The goal here was to both replicate the findings of past perception studies to ensure the 'goodness' of my stimuli as well as to collect accuracy and reaction time data for the three groups to use as comparison with the EEG results. An AXB task was used to force attention to the sounds. An AXB task is a common discrimination task in which three stimuli are played sequentially. Subjects are instructed to attend to the second stimuli, X, and then indicate which of the first (A) or third (B) stimulus is most similar to the target X. There is some evidence to suggest that AXB tasks produce more reliable discrimination results when compared to the more standard ABX paradigms, as ABX tests tend to show strong baises toward "B = X" (Gerrits & Schouten, 2004; Schouten, et al., 2003). Subjects were told they were being timed and to make their choices quickly. All behavioral experimentation took place in the Michigan State University Psycholinguistics lab in East Lansing, MI. An image of the AXB task given to participants in Praat is below in Figure 3.

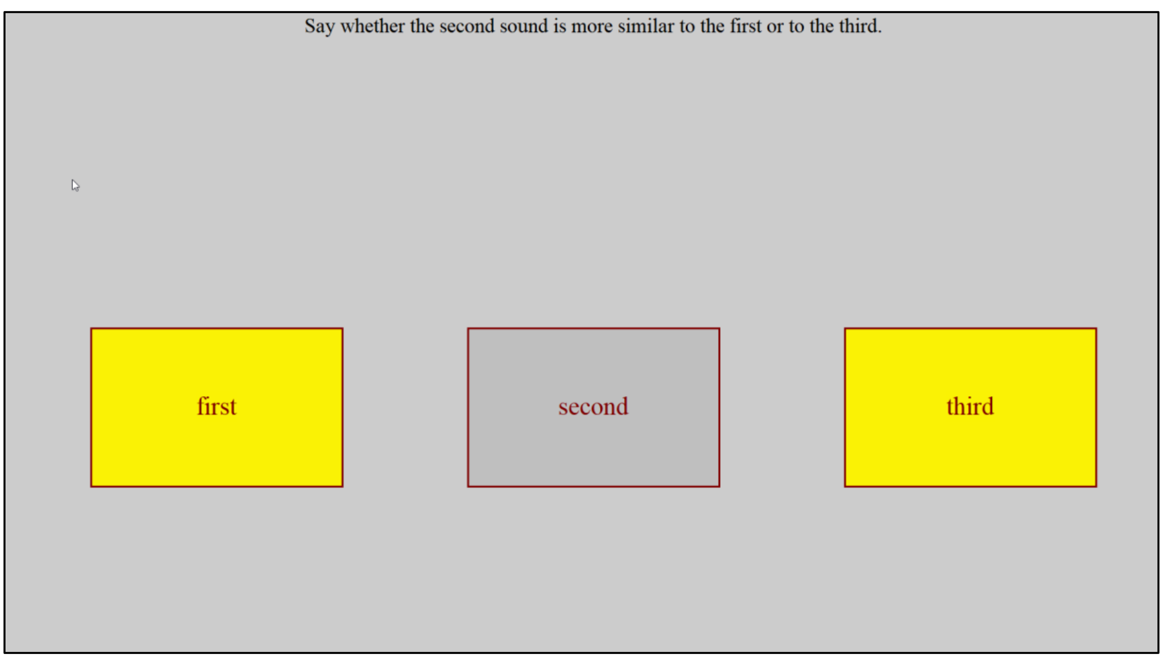

Figure 3: Screenshot of AXB Experiment in Praat

3.2.1 Participants

All subjects were Michigan State University students and were placed into one of three groups. One group of monolingual English speakers who had no formal instruction in Spanish beyond high school (N=13, 6 men, 7 women, average age=23 years) and had never spent time in a Spanish speaking country. One of the 13 monolingual English subjects is omitted from the behavioral results due to investigator error. A second group of native Spanish speakers (N=11, 4 men, 7 women, average age=26 years) who were L2 speakers of English. One speaker was from Spain, one was Argentinian, and the rest spoke Mexican or Caribbean Spanish. All speakers' native dialects follow the patterns described in Chapter 1 for all sounds of interest. The language background survey given to participants is available in APPENDIX B. A final group of native English speakers with advanced training in Spanish (N=14, 6 men, 8 women, average age=24 years) were included. These speakers had a minimum of 6 consecutive semesters of college Spanish instruction including courses on Spanish phonology. Subjects in this group also had spent a minimum of 3 months in a Spanish speaking country. Additionally, subjects in this group all stated intentions of using Spanish in their daily lives either at work or abroad. Most were training to become Spanish teachers. These speakers self-reported their average Spanish ability as 8 on a 0-10 scale, with 0 meaning no ability and 10 being 'native-like'. In all groups, no subject had any history of hearing or language disorders. All participants were right handed. All participants had normal or corrected-to-normal vision.

3.2.2 Stimuli

Materials consisted of three tokens each of the following VCV sequences: [ada], [aɾa], [aða]. All tokens were recorded by native Greek speakers, two female tokens and one male token were used. These tokens were graciously supplied by Amanda Boomershine and her colleagues and are a subset of the stimuli used in their 2008 paper which compared discriminability in English and Spanish speakers in two different tasks designed to emphasize either phonetic or phonemic contrasts (Boomershine et al., 2008). The stimuli used here come from their 3rd and 4th experiments, where they tested subjects using tokens produced by native Greek speakers. The Greek language contrasts all three sounds of interest phonemically and, most importantly, has a phonemic contrast between all three sounds intervocalically. Using the Greek tokens guarantees that no influence from English or Spanish exists in the stimuli that could potentially bias listeners. It also makes for natural sounding stimuli as English speakers would have difficulty producing [aɾa] naturally in isolation, while Spanish speakers would have similar problems producing natural tokens of [ada] because of the spirantization process in the language. The nine total tokens (three each of [ada], [aɾa], and [aða] were normalized in Praat to be a comfortable and consistent volume of 70dBA. To control the amplitude across tokens and speakers, the peak amplitude was equated for each of the tokens. To further control for any coarticulatory or cueing effects, the stimuli were further modified in Praat to have equal vowel length (150 +/- 10ms) on either side of the target consonant. To maintain a natural sounding token without pops or distortion, vowel fragments were spliced out peak-to-peak at zero crossings until the vowel lengths of all stimuli were roughly equal. The consonants under study are naturally of varying length and so consonantal fragments were spliced out to create stimuli of the same total length throughout the study (350 +/- 10ms). For detailed spectrograms of each stimulus, see the Figures 4-12 below.

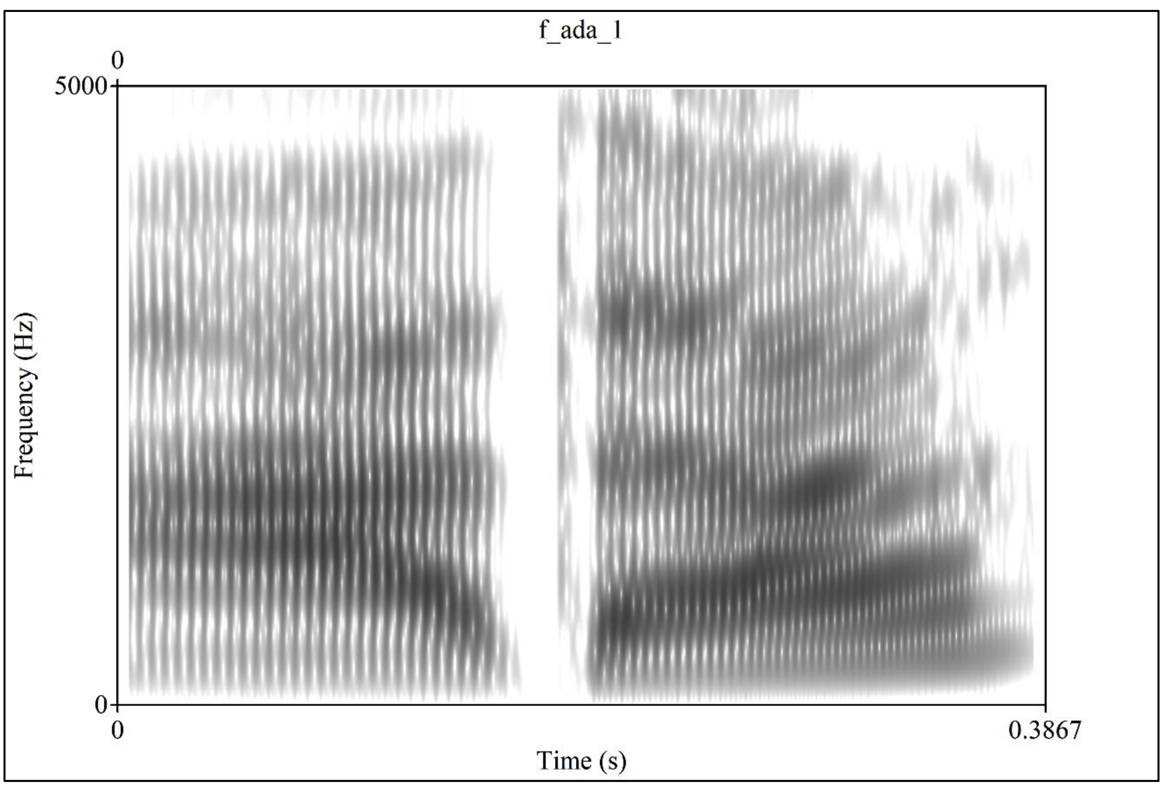

Figure 4: Spectrogram of [d] Stimulus, Version 1, Female Speaker

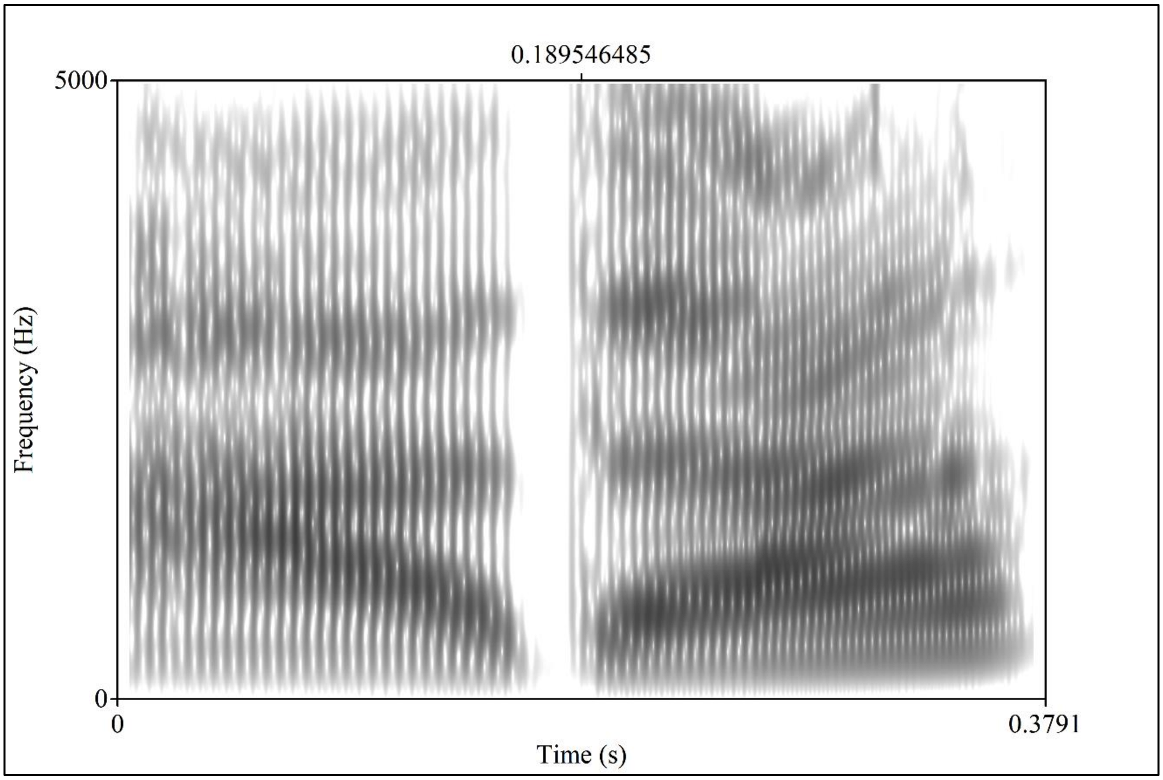

Figure 5: Spectrogram of [d] Stimulus, Version 2, Female Speaker

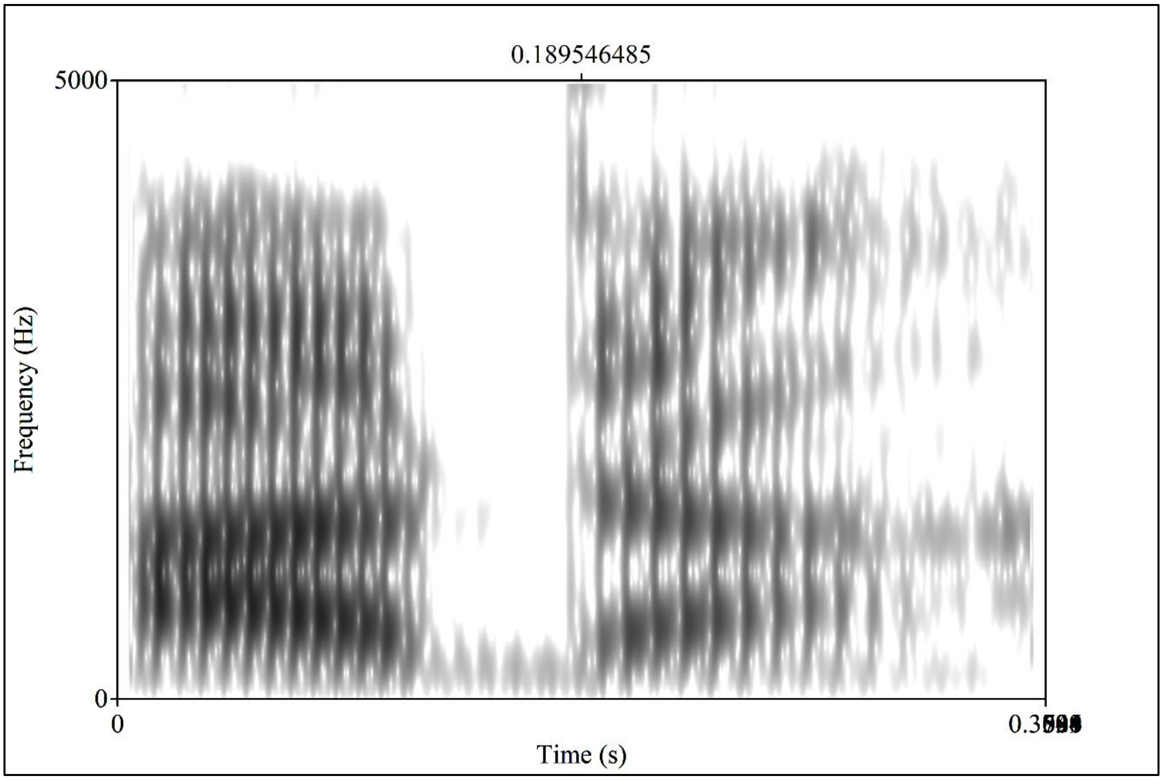

Figure 6:Spectrogram of [d] Stimulus, Version 3, Male Speaker

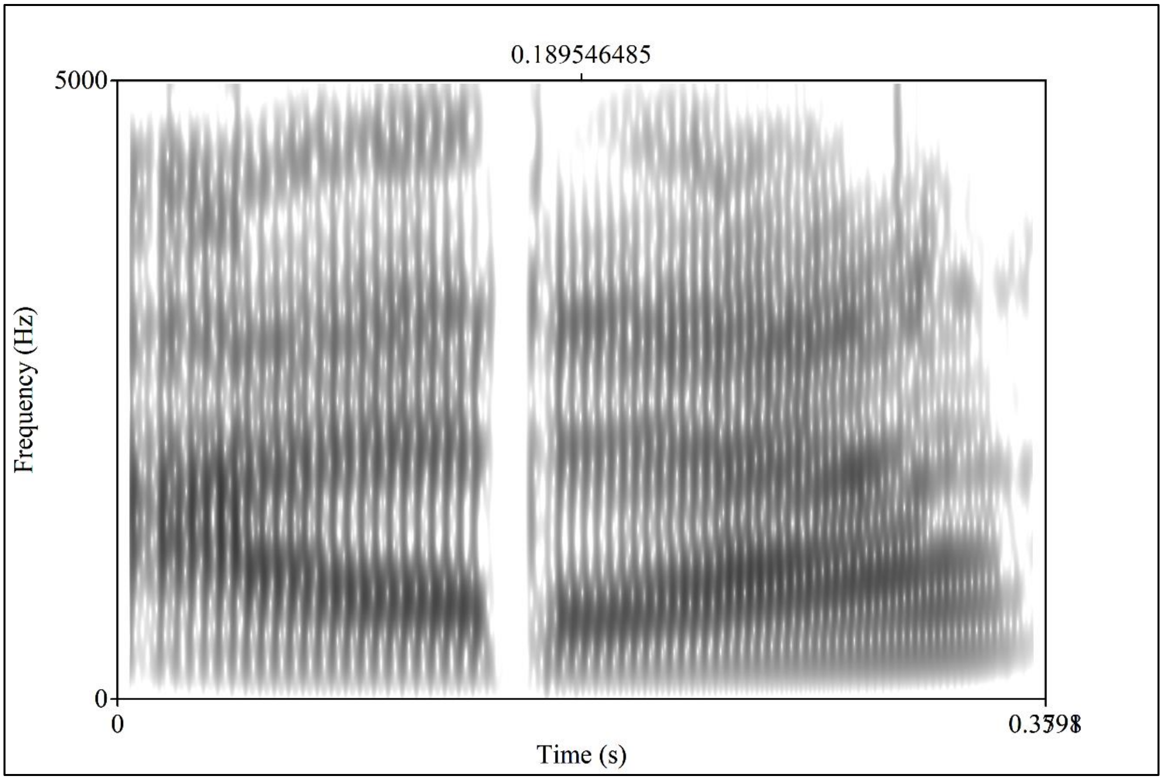

Figure 7: Spectrogram of [r] Stimulus, Version 1, Female Speaker

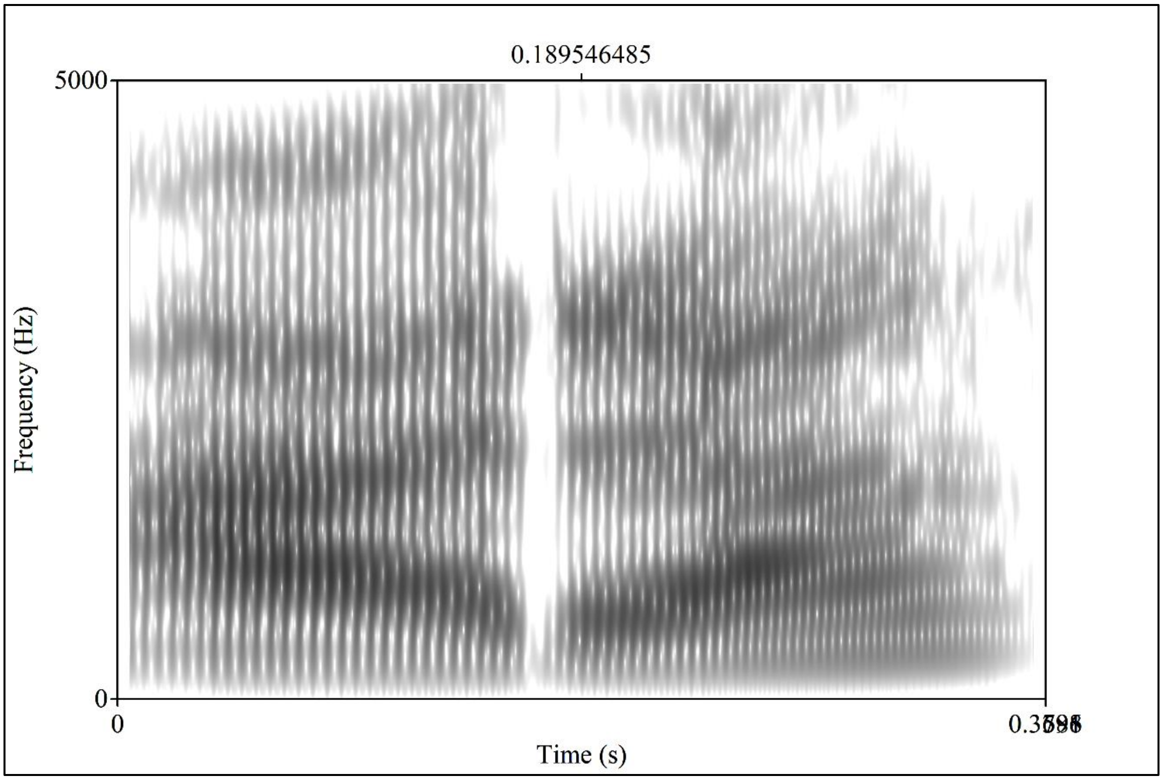

Figure 8: Spectrogram of [r] Stimulus, Version 2, Female Speaker

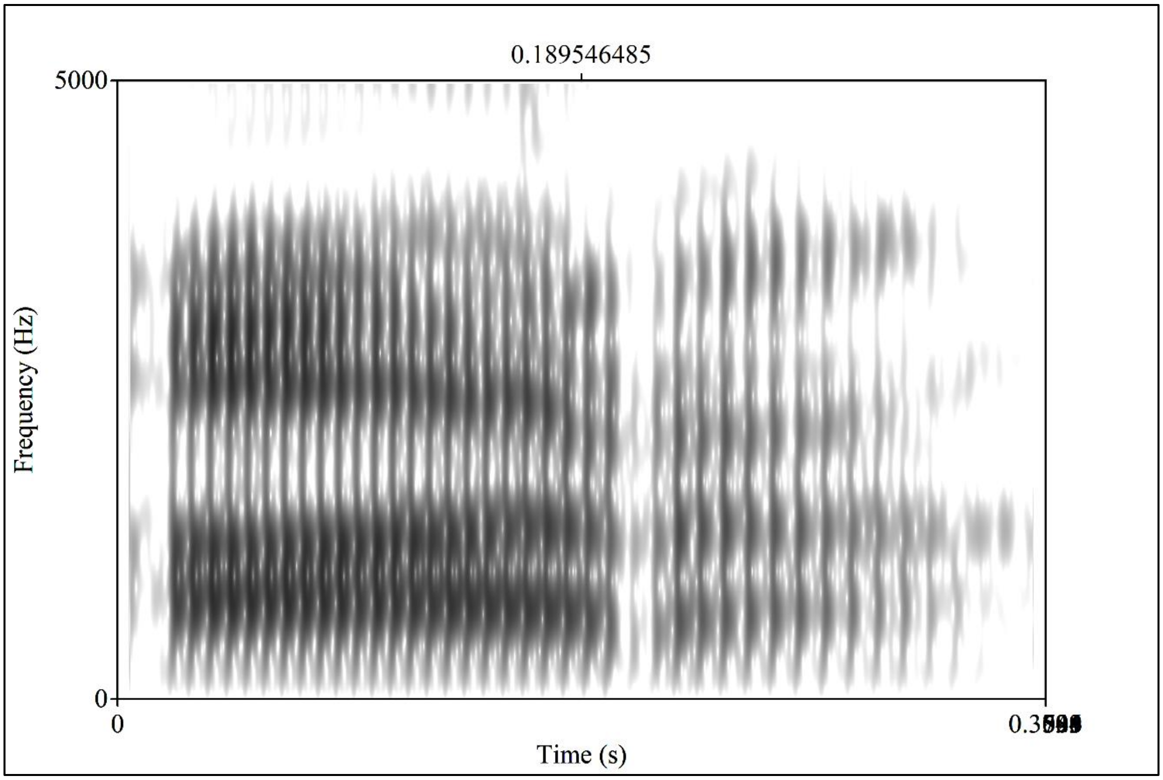

Figure 9: Spectrogram of [r] Stimulus, Version 3, Male Speaker

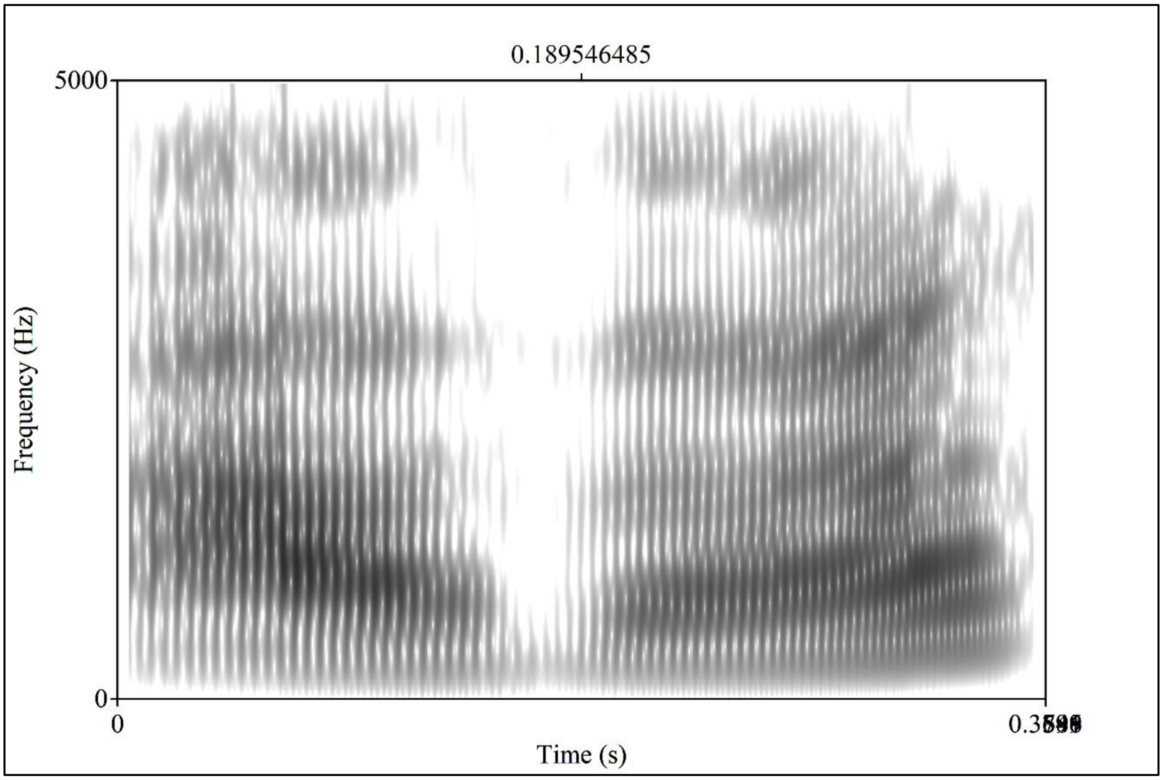

Figure 10: Spectrogram of [th] Stimulus, Version 1, Female Speaker

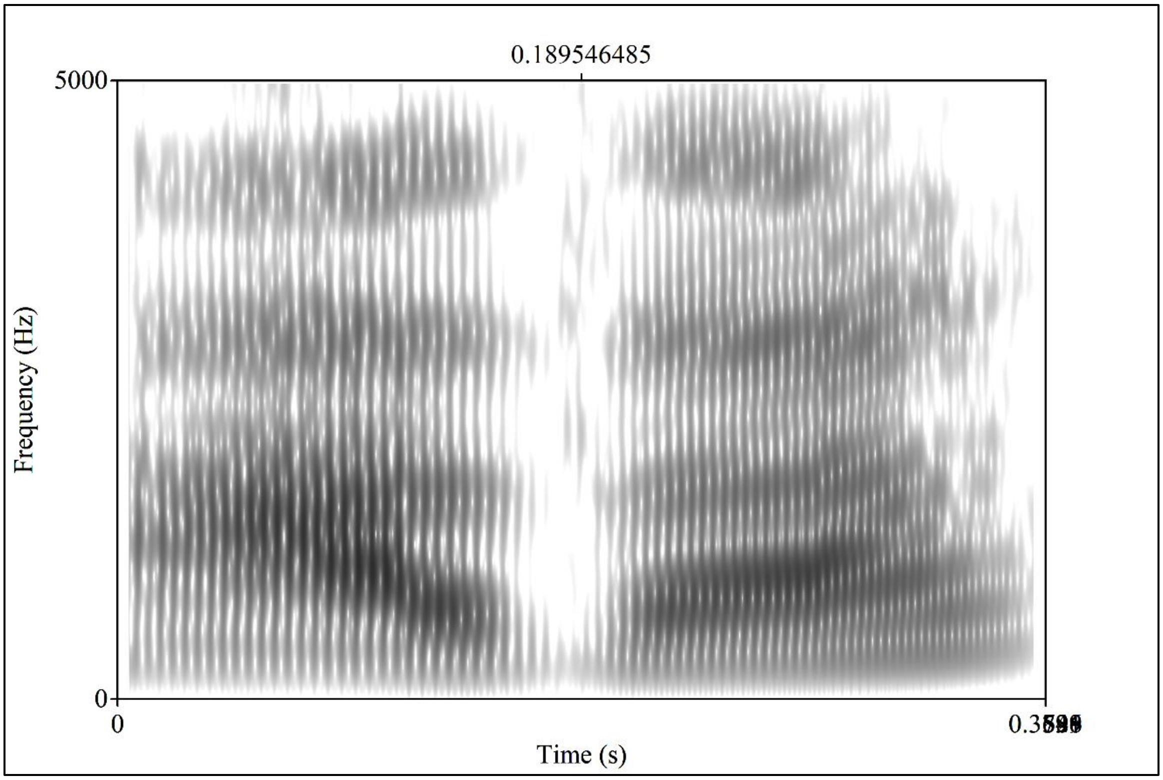

Figure 11: Spectrogram of [th] Stimulus, Version 2, Female Speaker

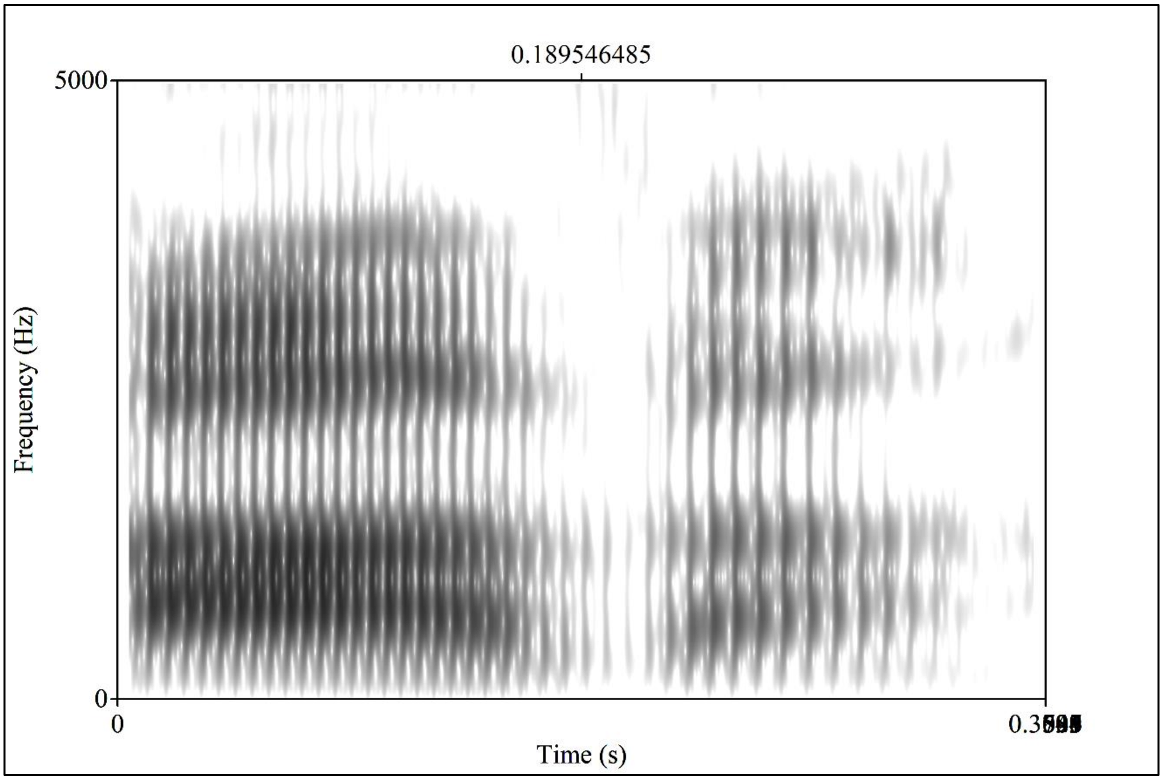

Figure 12: Spectrogram of [th] Stimulus, Version 3, Male Speaker

3.2.3 Procedure

Subjects performed an AXB discrimination task. Each subject heard three tokens in a row and were asked to choose whether the first sound they heard or third sound they heard was most similar to the second sound in the sequence. Stimuli were real speech tokens of [d] [ð] and [ɾ] produced by Greek speakers and detailed in the above section. The middle (2^{nd}) sound was always the male version of the token while the peripheral sounds were always a version of the female tokens. For example, a subject might hear [ada] followed by [ada] followed by [aɾa]. They would then choose (correctly) that the first sound is more similar to the second. "First" and "Third" were represented as large text boxes on the screen and participants would indicate their choice with a mouse. For each trial, the inter-stimulus interval (ISI) between sounds was 200ms. Between trials, the inter-trial interval was an additional 200ms. There was no time limit for the subject to choose an answer (though reaction times greater than 3500ms were excluded from the final analysis). The experiment software used was Praat (Boersma & Weenink, 2013). Stimuli were provided via closed-back over-ear headphones (Sennheiser HD202). Both accuracy and reaction times were recorded. All possible combinations of sounds and orders were given randomly to each subject for a total of 72 trials each. This took each participant on average about 10 minutes to complete.

3.3 Results & Analysis

This section provides details about the results of the AXB discrimination task for each of the three language groups as well as the three different measures taken as indices of discriminability. Percentage accuracy, A′ estimates, and mean reaction times were calculated for each participant and then compared both within language groups and across language groups. A note on the statistical tests used throughout Chapters 3 and 4; T-tests are one-tailed tests in comparisons that had a priori predictions about the direction of the outcome, while two-tailed tests were used otherwise. In practice, this means that for the monolingual English group and the native Spanish group, the expectation is that the phonemic contrast would be more easily discriminated than the allophonic contrast in the respective native language. Within language groups, the tests are paired. Across languages groups, because the groups vary in size, they are unpaired.

3.3.1 Within-language group results

This section gives the results and statistical analyses for the AXB task comparing discriminability of different pairs of sounds within a single language group. That is, how readily do monolingual English speakers discriminate [d] from [ɾ] compared to their ability to distinguish [d] from [ð] We begin with said group.

3.3.1.1 Behavioral results for monolingual English speakers

It is known that pairs of sounds that contrast phonemically in a speaker's language are perceptually more distinct than sounds which are not phonemically contrastive. Given prior results on the differences in discriminability of phonemic and allophonic contrasts, we predicted that the monolingual English speakers would perform poorly on this task when the target pair was [d]/[ɾ], but would be significantly more adept at discriminating the sounds when the target was [d]/[ð].

The behavioral results for English speakers are robust. As predicted, English speakers easily discriminate between [d] and [ð] which contrast phonemically in English. In contrast, the allophonic pair [d] and [ɾ] are much more difficult for the English speakers to discriminate between. This difference is borne out on all three dimensions: percentage accuracy, A′ values, and reaction times. The mean percentage accuracy for [d]/[ð] was 94.1% which was significantly higher than the mean percentage accuracy for [d]/[ɾ] of 68.1% ($t(11)=-3.84$, $p=0.002$). Mean A′ scores for the [d]/[ð] contrast were significantly higher ($M=0.969$) than those for the [d]/[ɾ] contrast ($M=.710$) ($t(11)=-3.43$, $p=0.003$). In addition to being more sensitive to the [d]/[ð] contrast, English speakers also responded significantly faster during the AXB task on the phonemic pair ($M=2.92s$) than the allophonic [d]/[ɾ] pair ($M=3.23s$) ($t(11)=-3.02$, $p=0.006$).

To summarize, the monolingual English speakers performed significantly better when discriminating the phonemic [d]/[ð] pair than the allophonic [d]/[ɾ] pair along every dimension measured. Figures 13-15 below show the behavioral results for all of the language groups and should be referred to in order to visualize the results discussed in the following sections.

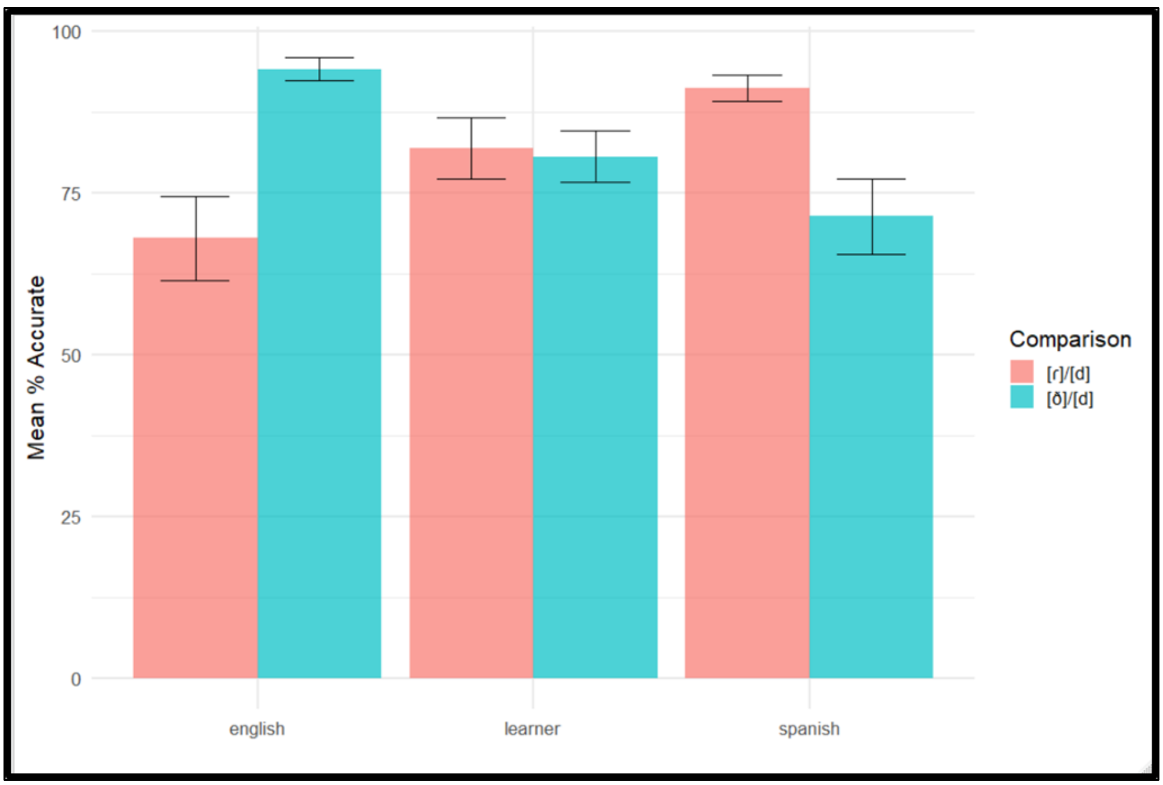

Figure 13: Mean % Accurate by Speaker Group and Sound Comparison

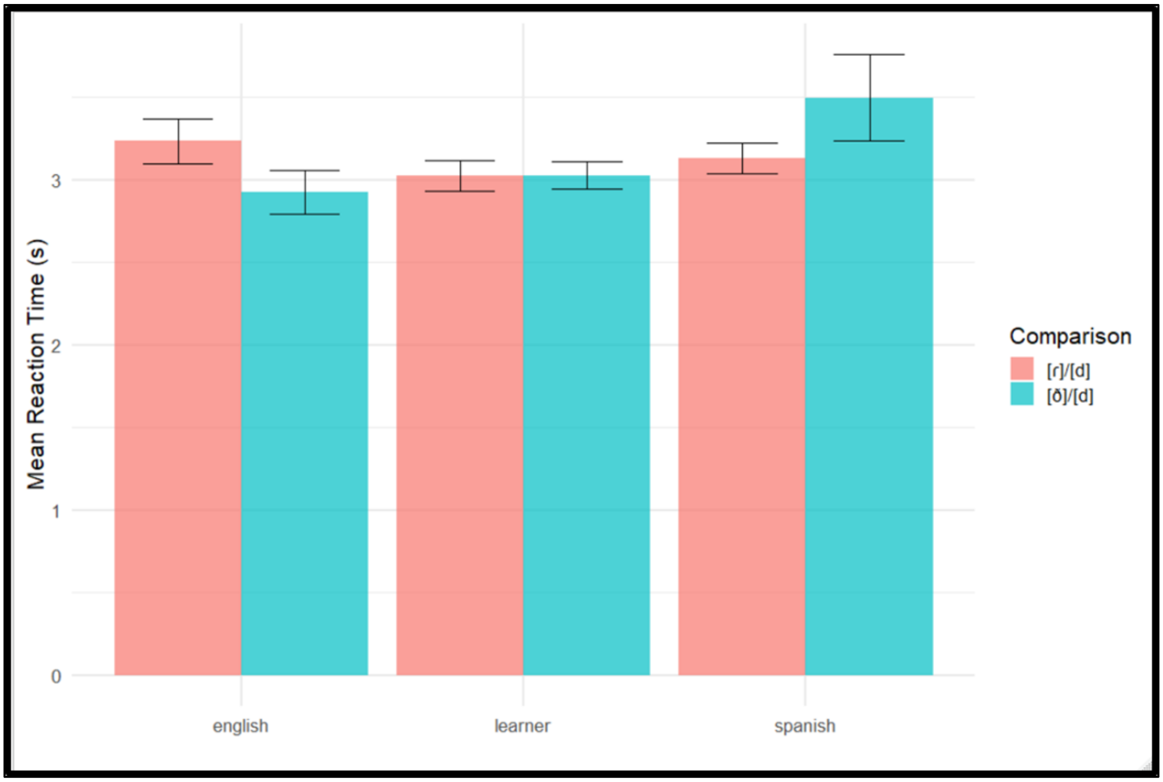

Figure 14: Mean Reaction Times by Speaker Group and Sound Comparison

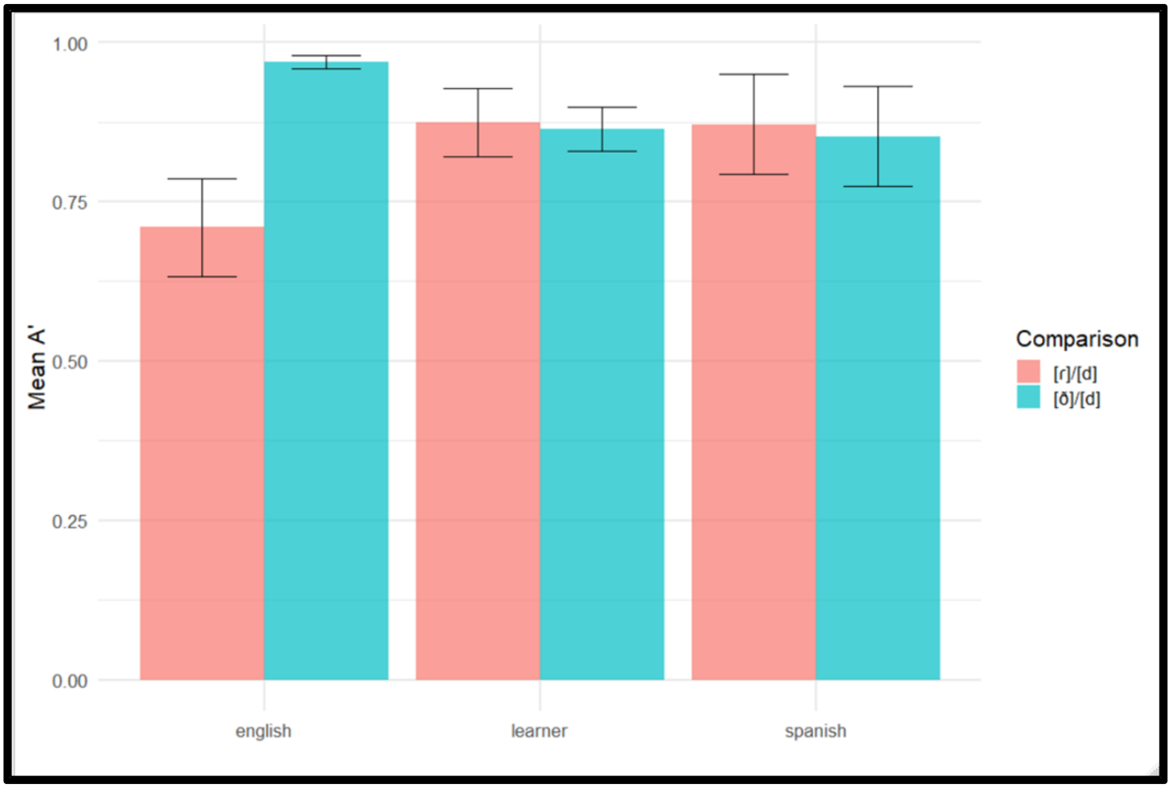

Figure 15: Percent A′ by Speaker Group and Sound Comparison

3.3.1.2 Behavioral results for native Spanish speakers

For the Spanish speakers, the prediction was the mirror of the English results. That is, the phonemic pair of [d]/[ɾ] would be much easier for the Spanish speakers to discriminate than the allophonic pair [d]/[ð]. These predictions were largely borne out. For the behavioral AXB task the native Spanish speakers had a mean accuracy of 91.2% when discriminating between [d] and [ɾ]. This high accuracy was expected given that [d] and [ɾ] are phonemically contrastive in Spanish. Although the group performed well on this contrast, it is noteworthy that they were slightly less accurate at this phonemic contrast than the monolingual English speakers were on their phonemic contrast of [d]/[ð]. This could be simply due to differences in acoustic similarity between the contrasts but could also be due to the fact that unlike [d] and [ð] in English, [d] and [ɾ] while phonemes in Spanish, do not appear in the same environment. [d] is spirantized between vowels and [ɾ] is never word initial in Spanish. As expected, the native Spanish speakers were much less accurate on the allophonic pair [d] and [ð]. On this contrast they had a mean accuracy of 71.4% ($t(10)=2.58$, $p=0.02$). Despite this large difference in accuracy, sensitivity scores were surprisingly not significantly different between the contrasts for Spanish speakers. Mean A′ for the phonemic [d]/[ɾ] pair were 0.87, while the allophonic pair had a mean A′ score of 0.85 ($t(10)=0.22$, $p=0.415$). This group also showed reaction time differences between the contrasts. Reaction times to the phonemic [d]/[ɾ] pair were faster (M=3.13s) than the allophonic [d]/[ð] pair (M=3.50s). This result was marginally significant ($t(10)=-1.57$, $p=0.07$).

These results suggest, like those from the monolingual English speakers, that phonemically contrastive sounds are more readily discriminated between than allophonic sounds. While both the contrastive [d]/[ɾ] pair and the allophonic [d]/[ð] had similar A′ sensitivity estimates, both percentage accuracy and reaction times were improved on the contrastive pair. It is important to note that unlike the monolingual English speakers, these native Spanish speakers have extensive experience in a second language, namely English. So, it is perhaps unsurprising that their discriminability measurements pattern slightly different from expectations when compared to the monolingual English speakers who have no L2 experience at all, let alone any in Spanish. Still, the trend toward the contrastive sounds having an advantage in discriminability is clear.

3.3.1.3 Behavioral results for advanced Spanish Learners

Predictions for the advanced Spanish learners are less straightforward. There are a several possibilities to discuss. If the learners have been able to acquire both the phonemic representation of [ɾ] and the allophonic variant of [ð] through their experience in L2 Spanish, then one would expect their results to look similarly to the native Spanish speaking group. However, if they have been unable to acquire 'native-like' representations of the Spanish sounds, then they should pattern more like the monolingual English speakers. A third possibility is some type of intermediate performance between the two groups, pointing toward acquisition-in-progress (Peltola et al., 2005). For accuracy rates, the learners mean for [d] - [ɾ] was 81.2% compared to a mean of 80.6% for [d]/[ð] ($t(13)=0.2$, $p=0.84$). In terms of sensitivity, the mean A' score for [d]/[ɾ] was 0.874. The mean A' score for the [d] - [ð] comparison was 0.864. These were not significantly different ($t(13)=0.148$, $p=0.442$). Reaction times for the two contrasts were also not significantly different. The mean reaction time for the [d] - [ɾ] comparison was 3.02s while the mean reaction time for the [d] - [ð] comparison was 3.03s ($t(13)=-0.028$, $p=0.489$).

3.3.1.4 Summary of within language behavioral results

The previous behavioral results suggest that for both the monolingual English speaking group and the native Spanish speaking group, the straightforward predictions that L1 phonological status prevails in perception. That is, that contrasts that are phonemic in L1 are more readily perceived than contrasts that are allophonic in the L1. This is shown by statistical differences in three different measures of discriminability; percentage correct, A' sensitivity measures, and reaction time measurements. For the advanced learners of Spanish however, behaviorally these contrasts are not significantly different on all three of these measures, perhaps indicating some intermediate level of acquisition. We will now turn to the between language comparisons which will give us more insight into this learning trajectory by the advanced Spanish learners. Table 1 summarizes these results. An asterisk indicates that the difference between the [d] - [ð] and [d] - [ɾ] comparisons was statistically significant.

	Mean Reaction Time (s)		% Correct		A′	
	[d]/[ð]	[d]/[ɾ]	[d]/[ð]	[d]/[ɾ]	[d]/[ð]	[d]/[ɾ]
Monolingual English Speakers	2.92*	3.23*	94.1*	68.1*	.969*	.710*
Native Spanish Speakers	3.50*	3.13*	71.4*	91.2*	.850	.870
Advanced Spanish Learners	3.03	3.02	81.2	80.6	.864	.874

Table 1: Behavioral Results for all Language Groups and Sound Comparisons, * denotes significant differences between sound comparisons.

3.3.2 Across-language group comparisons

In the previous sections, I detailed the results from within each speaker group. The following sections will look at how the different language groups compare to one another on the measures of mean amplitude, peak amplitude, and peak latency. Refer to Figures 11-13 above to visualize these results.

3.3.2.1 Monolingual English Speakers and Native Spanish Speakers

I will first compare the monolingual English speakers to the native Spanish speakers. There are clear predictions about how these groups should differ behaviorally, that is the phonemically contrastive pair in the respective L1 should be advantageous for perception. The within language behavioral results discussed above give us no reason to believe otherwise. For the [d]/[ɾ] contrast, which is allophonic in English but phonemic in Spanish, the English speakers had a mean A′ score of 0.71 while the Spanish speakers were predictably more sensitive to this contrast with A′=0.87. This result was marginally significant ($t(21)=-1.450$, $p=0.08$). The monolingual

English speakers were also much less accurate on this contrast (68%) than the native Spanish speakers (91.2%) ($t(21)=-2.81$, $p=.006$). The monolingual English speakers had a mean reaction time to this contrast of 3.23s, slower than the native Spanish speakers at 3.13s ($t(21)=0.58$, $p=0.28$). Turning to the opposite case, the [d]/[ð] contrast, English speakers had a mean A′ score of 0.969 while the Spanish speakers had an A′ score of 0.852. This result was significant ($t(21)=1.711$, $p=0.05$). The monolingual English speakers had significantly higher accuracy when discriminating this contrast at 94.1% compared to 71.4% for the native Spanish speakers ($t(21)=3.96$, $p=0.0005$). For reaction times, the monolingual English speakers had a mean reaction time of 3.29s compared to 3.50s for the native Spanish speakers. ($t(21)=-2.13$, $p=.02$).

3.3.2.2 Advanced Spanish Learners and Monolingual English Speakers

The more interesting case is of course the comparisons of our English and Spanish groups with the advanced learner group. We've seen that the advanced Spanish learners showed no statistical differences between the [d]/[ɾ] or [d]/[ð] comparisons. At the outset of the chapter, I laid out three possible outcomes for the learner group. They could behave like the monolingual English speakers, or they could behave like the native Spanish speakers, or something else entirely, presumably an intermediate level of perception between the other groups. I will first compare them to the monolingual English speakers. For the [d] - [ɾ] comparison, the learners have significantly higher A′ scores ($M=0.874$) than do the English speakers ($M=0.710$) ($t(24)=-$

1.810, p=0.04). The advanced Spanish learners are also more accurate on this contrast (81.2%) than the monolingual English speakers (68%) (t(24)=-1.76, p=004). The advanced Spanish learners had a mean reaction time of 3.02s compared to 3.23s for the monolingual English speakers. On the [d]/[ð] comparison, for which the learners need to acquire an allophonic variant for a sound they already contrast phonemically, A′ scores average 0.864 compared to the English speakers 0.969. This is significantly lower than their monolingual peers (t(24)=2.722,p=0.006). In terms of accuracy, the learners correctly discriminated [d] from [ð] 80.6% of the time compared with 94.1% for the monolingual English speakers. This difference was statistically significant (t(24)=2.96, p=0.007). The advanced Spanish speakers had a mean reaction time of 3.03s while the monolingual English speakers had a mean reaction time of 3.29s.

Based on A′ sensitivity estimates and the accuracy hit rates, it looks as though the advanced learners of Spanish have improved significantly at discriminating [d] from [ɾ] over their monolingual counterparts. This could signal acquisition of a phonemic representation of [ɾ] which would make these sounds clearly contrastive for them. More puzzling is the advanced learners of Spanish significant decline in discriminability between [d] and [ð] when compared to the monolingual English speakers. This suggests a level of L2 interference perhaps caused by having to maintain multiple representations of the same sound. While unexpected, work on Finnish bilinguals have shown similar degradation in discriminability of L1 contrasts for vowels (Peltola et al., 2003)

3.3.2.3 Advanced Spanish Learners and Native Spanish Speakers

When compared to the native Spanish speakers, the results look quite different. On the [d]/[ɾ] contrast, which is phonemically contrastive in the target language, the learners have a mean A′ score of 0.874, while the native Spanish speakers score A′=0.872. These scores are not statistically different (t(23)=0.025, p=0.49). The native Spanish speakers were correct 84.7% of the time while the advanced Spanish learners had an 81.9% accuracy rate (t(23)=0.36, p=0.36). Mean reaction time for the advanced Spanish learners was 3.02s compared to 3.13s for the native Spanish speakers. Similarly, the [d]/[ð] comparison, which is allophonic in the target language, the learners have a mean A′ score of 0.864. The native Spanish speakers A′ score of 0.852 is not significantly different than that of the learners (t(23)=-0.151, p=0.44). The advanced Spanish learners had an accuracy of 80.6%, while the native Spanish speakers had an accuracy of 71.4%. This difference was not significant (t(23)=1.28, p=0.21). The advanced Spanish learners had a mean reaction time of 3.03s compared to 3.50s for the native Spanish speakers.

Based on A′ scores and accuracy rates, the performance of the advanced learners of Spanish is not clearly distinguishable from that of the native Spanish speakers on both contrasts. As mentioned in section 3.3.1.3, part of this could be due to the fact that the native Spanish speaking group is not monolingual. Their experience in English could have increased their ability to discriminate [d] from [ð]. What is clear is that the advance Spanish learners are able to reliably discriminate the Spanish phonemic contrast between [d] and [ɾ].

3.3.2.4 Summary of cross language results

The cross-language group results from the behavioral discrimination task seem to suggest that as expected, the advanced Spanish learners are becoming more 'native Spanish-like' in the perception of both contrasts. Although I used a slightly different discrimination task with my subjects (AXB vs ABX), the results here replicate those of the English and Spanish speakers from Boomershine et al., (2008), while adding the results of the advance Spanish learners.

Chapter 4: Electrophysiological task

4.1 Introduction

Immediately following the AXB task all subjects went to prepare for the EEG experiment. All EEG experiments took place in the MSU neurolinguistics lab.

4.2 Methods

4.2.1 Participants

All subjects from the behavioral experiment in Chapter 3 participated in the MMN task. To recall, there was one group of monolingual English speakers who had no formal instruction in Spanish beyond high school (N=13, 6 men , 7 women, average age=23 years) and had never spent time in a Spanish speaking country. A second group of native Spanish speakers (N=11, 4 men, 7 women, average age=26 years) who were L2 speakers of English. A final group of native English speakers with advanced training in Spanish (N=14, 6 men, 8 women, average age=24 years) were included. These speakers had a minimum of 6 consecutive semesters of college Spanish instruction including courses on Spanish phonology. Subjects in this group also had spent a minimum of 3 months in a Spanish speaking country.

4.2.2 Stimuli

Stimuli in the MMN task were identical to those in the behavioral task described above in Chapter 3 save for the exclusion of the male voiced tokens. The MMN procedure used only the female versions of the token (2 versions of each sound) to control for acoustic differences between male and female speakers.

4.2.3 Procedure

Subjects are briefly introduced to the EEG equipment and the preparation and procedure is explained to them. Subjects begin by having a head measurement taken to ensure a properly sized electrode cap is selected. This is so the electrode placement is consistent between all subjects regardless of morphological differences. Once a cap is fitted, preparation of the individual electrodes begins. A small amount of electrolyte gel is used between the electrode and the subjects scalp (onestep clear gel). The skin under each electrode is then gently abraded with a wooden applicator to create good conductivity between the scalp and electrode via the gel. To ensure the best signal, all impedances on each electrode are kept below 5kΩ. Once the subject is prepared for recording, they are instructed to remain still for the duration of the experiment. Because the MMN is a so-called preattentive process, subjects do not need to directly attend to the stimuli being presented (Näätänen et al., 1993; Sussman et al., 2003). To distract subjects (and to keep them awake), participants were made to watch a movie while the

MMN series of sounds were presented to them. The movie was on silent with no subtitles. Subjects were each played two separate series of sounds through headphones (Sennheiser HD203) in the MMN paradigm, where a series of similar stimuli are presented in order to build a 'standard' representation in the listener's mind, followed by the presentation of a 'deviant' stimulus that differs in at least one dimension from the standards. One series consisted of standard [ada] and deviant [aða], followed by the reverse presentation. The second series consisted of standard [ada] and deviant [aɾa], followed by the reverse presentation. The order of these series was alternated for every other subject. The electroencephalogram (EEG) was recorded from 32 sintered Ag/AgCl electrodes held in place on the scalp by an elastic cap (GND WaveGuard 64 Electrode cap; Advanced Neuro Technology BV., Enschede, The Netherlands). It was amplified using a Full-band EEG DC Amplifier (Advanced Neuro Technology), with a 256 Hz sampling rate (and a bandpass filter of 0.01 to 30 Hz which was applied offline). Electrode impedances were kept below 5 kΩ. Averaged mastoids were used as reference (Dien, 1998; Luck, 2005). All signals were recorded continuously and captured within Advanced Neuro Technology's Advanced Source Analysis software (Zanow & Knösche, 2004). A 100ms pre-stimulus baseline was subtracted from all epochs. Epochs were defined as -100ms – 700ms in relation to the stimulus onset. A voltage rejection criterion of +/- 75 μV was applied to the data such that any epoch with samples beyond this threshold were removed from analysis. This resulted in a loss of approximately 6%

of total trials. Individual subject epochs were then averaged using an 800ms averaging window. Grand average difference waves were then constructed from the resulting subject averages. All post-processing of the recording data along with measurements and ERP plots were done with ERPLAB, an ERP plug-in for the MATLAB toolbox ERPLAB (Delorme & Makeig, 2004; Lopez-Calderon & Luck, 2014; The Mathworks Inc, 2018).

4.3 Results & Analysis

4.3.1 Within-language group results

This section gives the results and statistical analyses for the EEG task comparing discriminability of different pairs of sounds within a single language group. For the MMN task, I will present the results in two ways. I will first compare the average ERP amplitude over the 300-400ms time window of each of the crucial sound pairs against one another in the classic MMN configuration. This shows the amplitude of the deviant response when compared to the amplitude of the response to the standard stimulus. Following that, I will show each of the crucial sounds in turn in the identity or iMMN configuration, with the average amplitude of the deviant stimulus being compared to the average amplitude of the standard stimulus of the same sound. It is important to show both comparisons given that there have been significant differences in how MMNs have been interpreted in the speech perception literature depending on the configuration used in analysis. In the next section, I will present the cross-language comparisons, but here I will discuss each language in turn.

4.3.1.1 EEG results for monolingual English speakers

In this section I will describe several analyses of the ERP obtained from the monolingual English speakers on the MMN task. The MMN literature presents a number of ways one may determine the electrodes of interest, but the mismatch negativity is well known to present in the fronto-central region of the scalp. For this reason, it is unnecessary to perform whole-head analysis by region-of-interest. Indeed, it is not uncommon to perform single electrode analysis when doing MMN studies in EEG. Here, I choose instead to do a cluster analysis, averaging values from four fronto-central electrodes on the basis of visual inspection of mastoid inversions, which distinguishes the MMN from other fronto-central components (Schrager, 1998) . The contiguous electrodes Fz, FC1, FC2, and Cz form on the vertex area of the scalp and were the most visually similar in the grand-averaged waveform to the inverted mastoids M1 and M2.

The first analysis I present is an averaged amplitude between two ERP latencies. The window of analysis was chosen as 300-400ms post stimulus onset because the VCV stimuli used in my experiment contain initial vowels of approximately 150ms in length, making 300-400ms post onset 150-250ms post consonant onset. This is both the traditional location of the MMN (part of the N1-P2 complex), and where clear mismatches occurred in my data by visual inspection (Kilner et al., 2009; Näätänen et al., 2007). Before looking at the results of individual mismatch contrasts in the EEG task, I present the results of the omnibus ANOVA. In the 300-400ms time window there was a main effect of condition, standard vs. deviant (F(1,72)=9.54, p=.003). There was also an interaction effect of segment ([d], [ɾ], [ð]) and condition (F(2,72)=4.28, p=.02). I will now turn to the post-hoc tests for the various contrasts.

4.3.1.1.1 Results for [ɾ]

I will first look at the mismatch between the allophonic flap consonant [ɾ] and [d]. Based on previous work, I predicted that the mismatch for [ɾ] in English would be attenuated due to its non-contractiveness with the phoneme /d/. It was also shown that these same subjects displayed poor discriminability for the [d]/[ɾ] contrast in the previous behavioral task. This predicted result was not borne out. The monolingual English speakers displayed a prominent classic MMN response for [ɾ] upon visual inspection (Figure 16). When compared to a standard [d], the deviant [ɾ] was significantly more negative in the 300-400ms time period (t(12)=5.23, p=1.1e-4). The iMMN analysis yielded similar results (Figure 17). [ɾ] as a deviant was significantly more negative in the 300-400ms time window than [ɾ] as a standard (t(12)=5.31, p=9.1e-5).

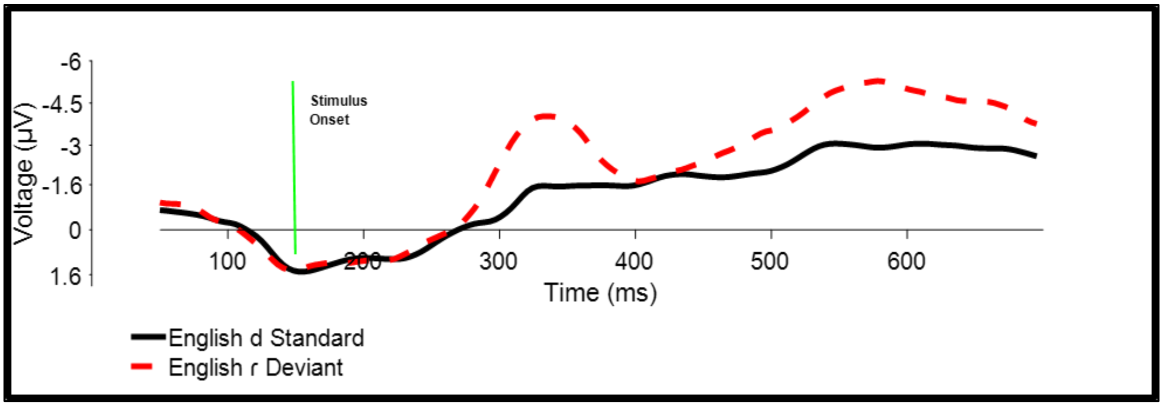

Figure 16: Averaged deviant [ɾ] and standard [d] ERP for the monolingual English speakers. Showing a large MMN from 300-400ms

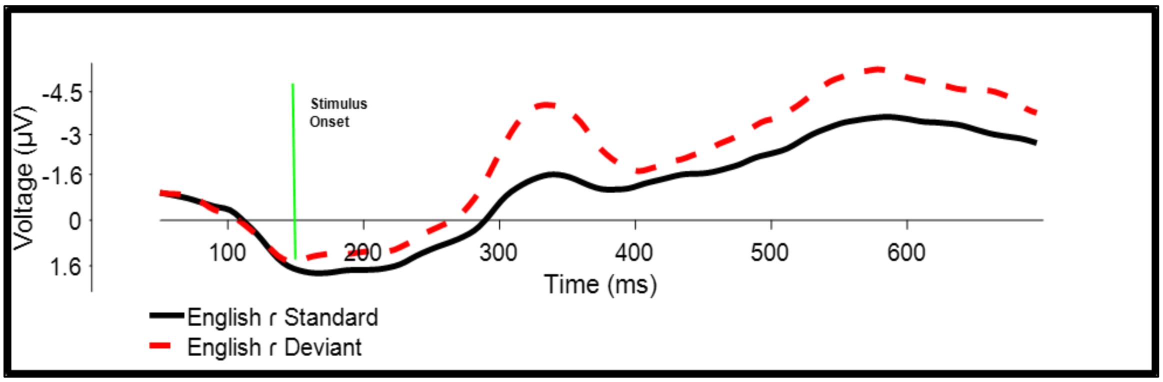

Figure 17: Averaged deviant [ɾ] and standard [ɾ] ERP for the monolingual English speakers. Showing a large iMMN from 300-400ms

4.3.1.1.2 Results for [ð]

Turning now to the mismatch responses for [ð]. Because [ð] contrasts phonemically with [d] in English, and behavioral results showed excellent discriminability between the two sounds, I predicted a prominent MMN response from the monolingual English subjects. This prediction was borne out. In the 300-400ms time window, there was a significant classic MMN for deviant [ð] when compared to standard [d] (t(12)=4.43, p=.0004) (Figure 18). The iMMN response was even stronger when comparing deviant [ð] to standard [ð] (t(12)=3.56, p=.002) (Figure 19). One thing to note when comparing the mismatch responses for [ð] and [ɾ] is that the [ɾ] components while greater in amplitude are not nearly as sustained as those generated by [ð].

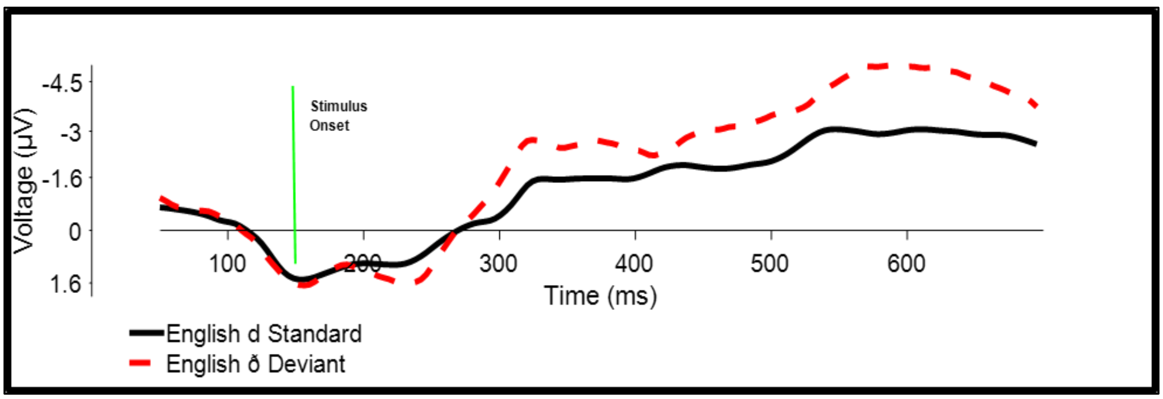

Figure 18: Averaged deviant [ð] and standard [d] ERP for the monolingual English speakers. Showing a MMN from 300-400ms

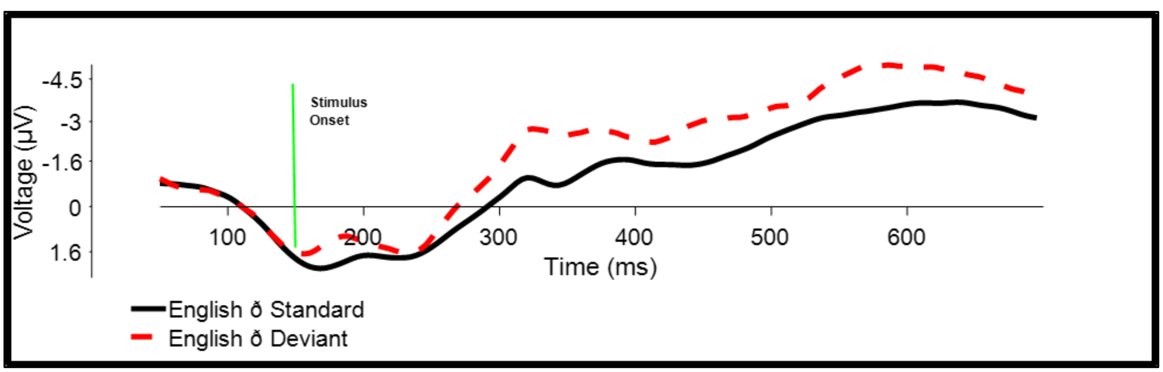

Figure 19: Averaged deviant [ð] and standard [ð] ERP for the monolingual English speakers. Showing a iMMN from 300-400ms

4.3.1.2 EEG results for native Spanish speakers

In this section, I will detail the electrophysiological results of the native Spanish speaking group on the mismatch task. This task was identical to the one given to the monolingual English speakers, and as in the previous section, I will discuss the classic MMN and iMNN results in the 300-400ms time windows for both [ɾ] and [ð].

4.3.1.2.1 Results for [ɾ]

In Spanish, unlike in English, [ɾ] contrasts phonemically with [d] and so a clear mismatch response for [ɾ] is expected. In the 300-400ms time window, ANOVA revealed a main effect for condition, standard vs. deviant (F(1,60)=10.65, p=.001). Post hoc tests revealed a significant classic MMN for deviant [ɾ] when compared to standard [d] (t(10)=3.42, p=.003) (Figure 20). There was also a significant iMMN for deviant [ɾ] when compared to standard [ɾ] (t(10)=3.87, p=.002) (Figure 21).

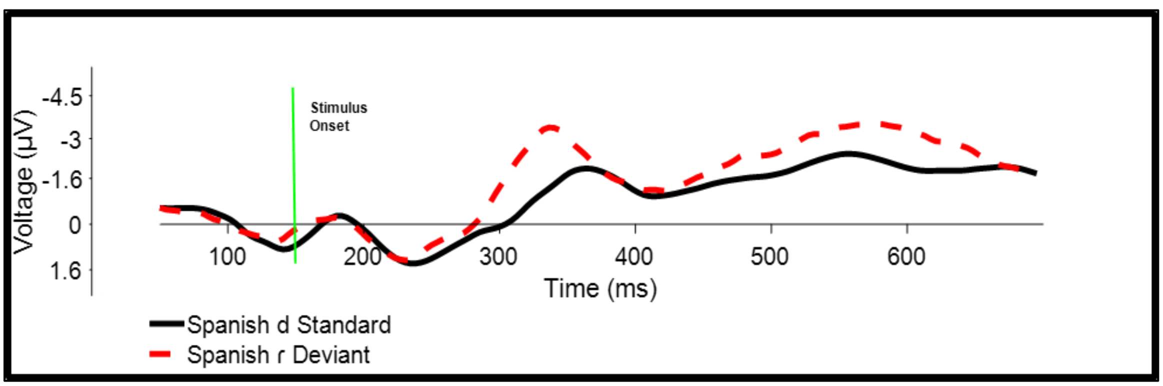

Figure 20: Averaged deviant [ɾ] and standard [d] ERP for the native Spanish speakers. Showing a large MMN from 300-400ms

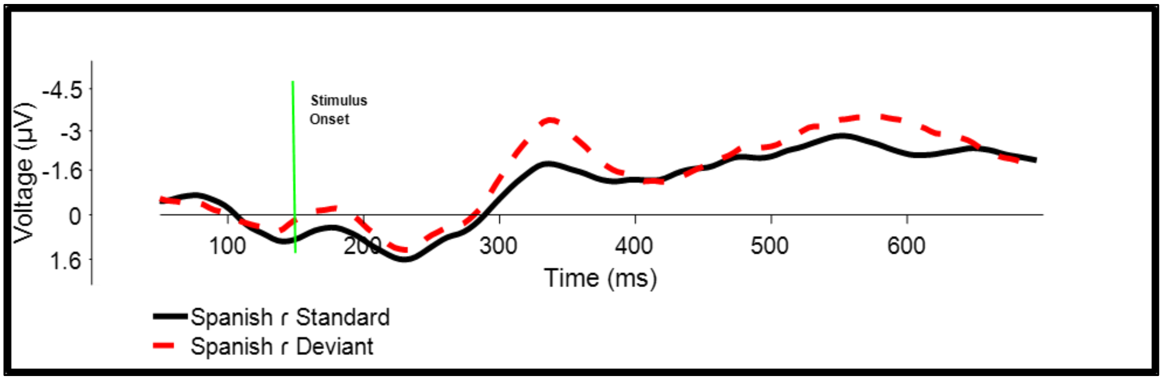

Figure 21: Averaged deviant [ɾ] and standard [ɾ] ERP for the native Spanish speakers. Showing a large iMMN from 300-400ms

4.3.1.2.2 Results for [ð]

For [ð], which is an allophone of [d] in Spanish and behavioral results showed that native Spanish speakers find the distinction between [ð] and [d] significantly harder to perceive than the phonemic [ɾ]/[d] contrast. For these reasons, an attenuated mismatch response is expected for [ð] in this group. This result is borne out in the MMN analysis. In the 300-400ms time window, the classic MMN for deviant [ð] when compared to standard [d] showed a significant mismatch (t(10)=3.78, p=.002) (Figure 22). However, in the iMMN analysis where deviant [ð] is compared to standard [ð], there was only a marginal difference (t(10)= 1.53, p=.08) (Figure 23).

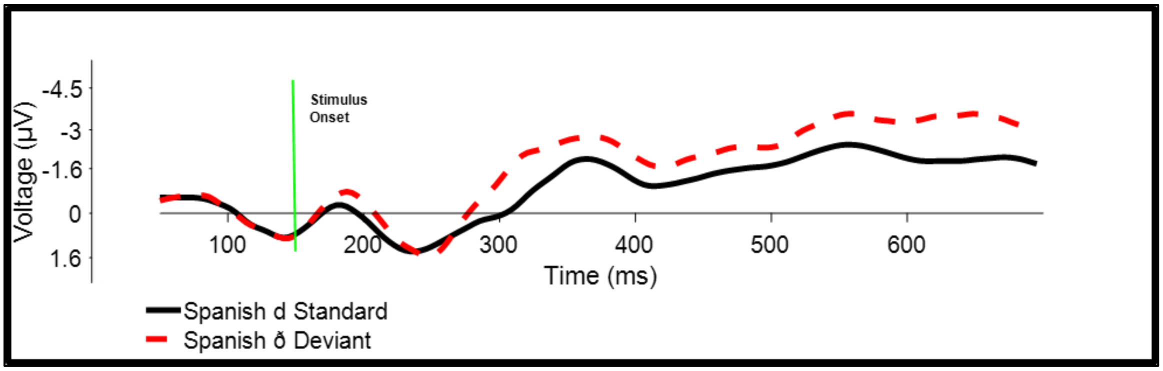

Figure 22: Averaged deviant [ð] and standard [d] ERP for the native Spanish speakers. Showing a MMN from 300-400ms

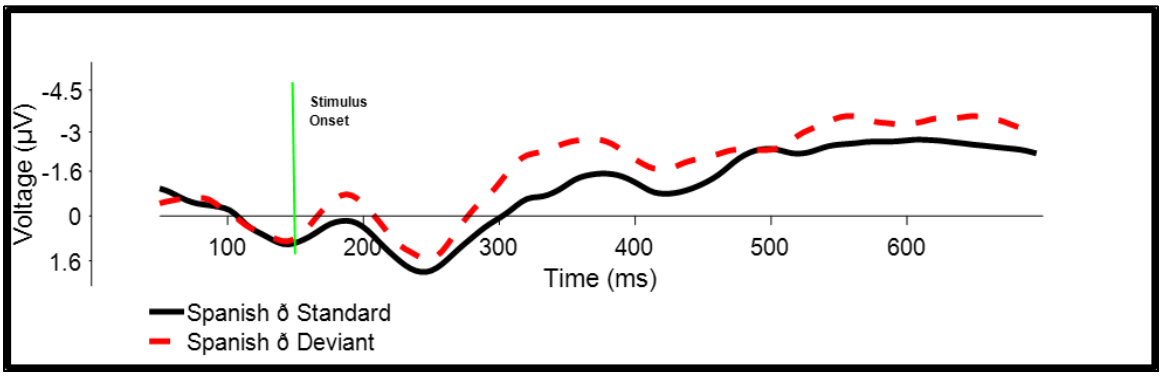

Figure 23: Averaged deviant [ð] and standard [ð] ERP for the native Spanish speakers. Showing an iMMN from 300-400ms

4.3.1.3 EEG results for advanced Spanish Learners

In the 300-400ms window the advanced Spanish learners had a main effect for condition ($F(1,78)=11.03$, $p=.001$) and marginally significant interaction effect of segment by condition ($F(2,78)=2.41$, $p=.09$).

4.3.1.3.1 Results for [ɾ]

There was a significant classic MMN for [ɾ] (t(13)=3.73, p=.003) (Figure 24). There was also a significant iMMN (t(13)=3.67, p=.003) (Figure 25).

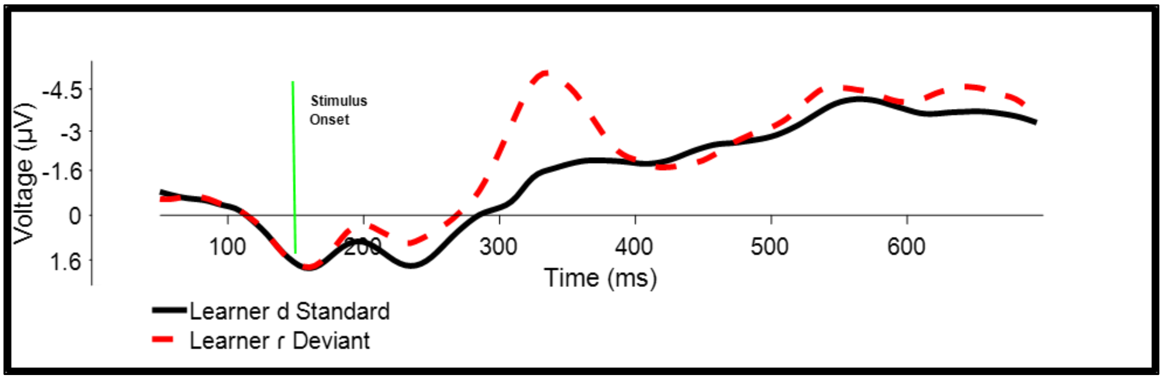

Figure 24: Averaged deviant [ɾ] and standard [d] ERP for the advanced Spanish learners. Showing a large MMN from 300-400ms

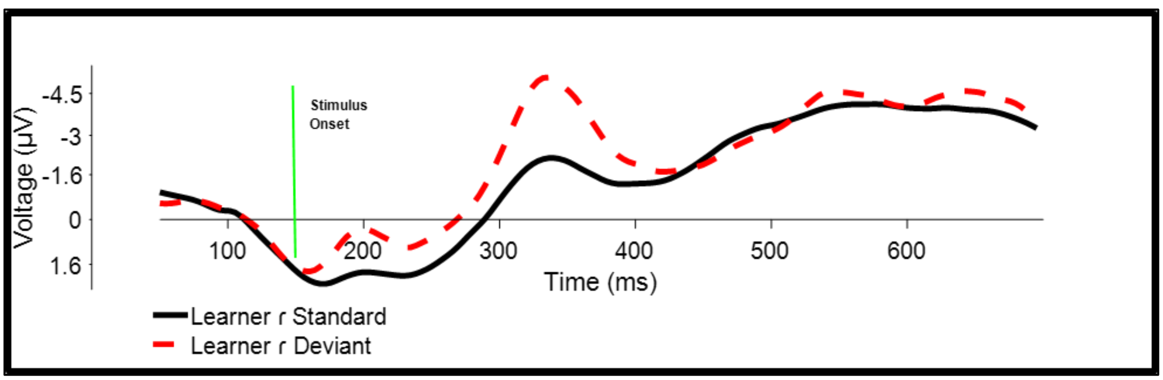

Figure 25: Averaged deviant [ɾ] and standard [ɾ] ERP for the advanced Spanish learners. Showing a large iMMN from 300-400ms

4.3.1.3.2 Results for [ð]

There was a significant classic MMN for [ð] (t(13)=6.04, p=4.15e-5) (Figure 26). There was also a significant iMMN for [ð] (t(13)=4.29, p=8.7e-4) (Figure 27).

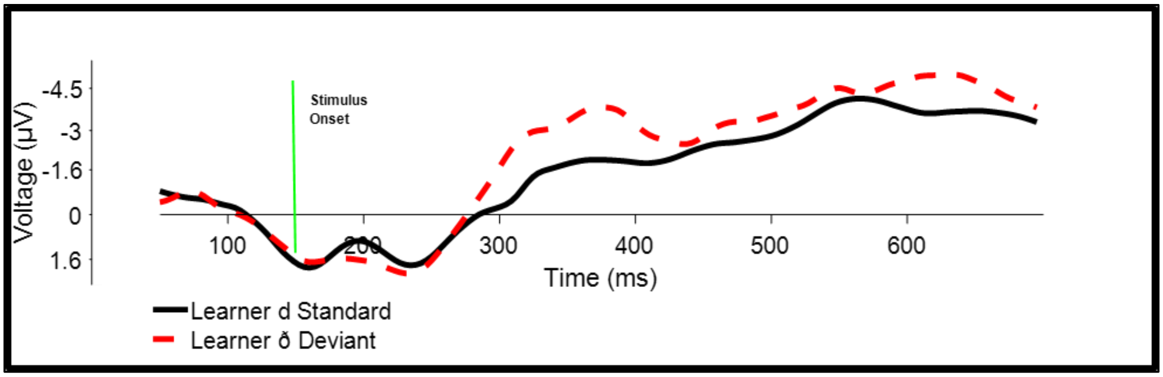

Figure 26: Averaged deviant [ð] and standard [d] ERP for the advanced Spanish learners. Showing a MMN from 300-400ms

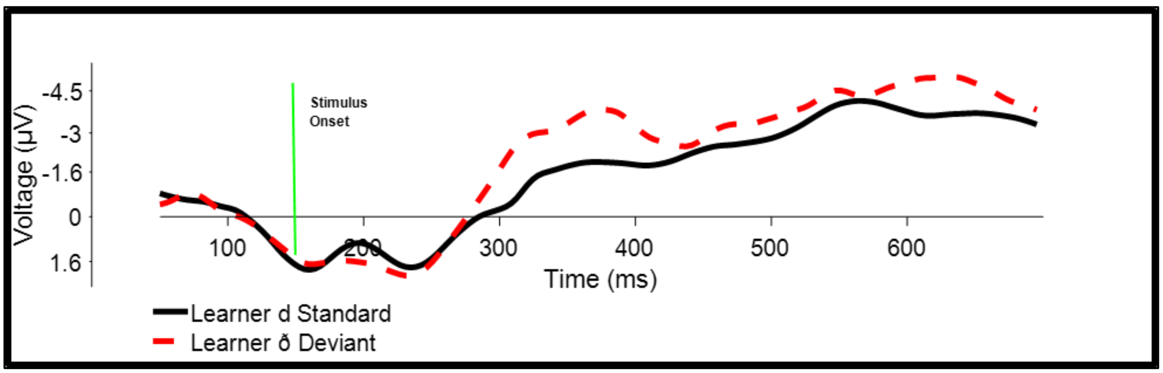

Figure 27: Averaged deviant [ð] and standard [ð] ERP for the advanced Spanish learners. Showing an iMMN from 300-400ms

4.3.1.4 Summary of within language EEG results

Contra expectations, all language groups showed significant mismatch responses for both comparisons. This was true of both the classic MMN and identity iMNN conditions. The only exception was the iMMN for the [ð] in the native Spanish speakers. This difference was marginally significant. Table 2 shows all the mean amplitude mismatch responses for each language group by sound and condition. Additional results are visualized in Figure 28.

Speaker Group	MMN [ɾ]	iMMN [ɾ]	MMN [ð]	iMMN [ð]
Monolingual English	-1.81 µV *	-2.07 µV *	-1.26 µV *	-1.51 µV *
Native English learner of Spanish	-2.07 µV *	-2.01 µV *	-1.79 µV *	-2.02 µV *
Native Spanish	-1.15 µV *	-0.99 µV *	-1.13 µV *	-1.36 µV

Table 2: Within language group mismatch mean amplitudes, * denotes significant difference from standard. The results show a significant mismatch in all comparisons for all language groups, except for the [ð] in the iMMN condition for the native Spanish speakers.

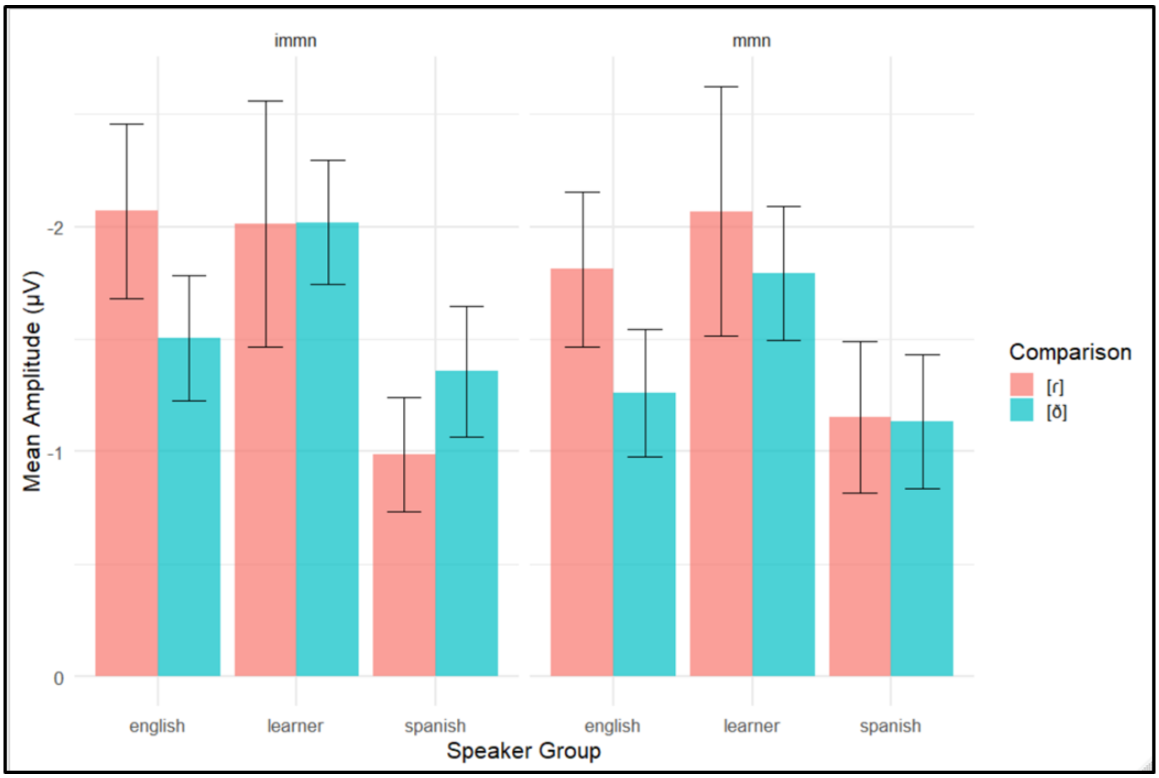

Figure 28: Mean amplitudes for all speaker groups by comparison and configuration. A discussion of comparisons across language groups is available in section 4.3.3

4.3.2 Peak latency & amplitude by segment

The standard practice in mismatch negativity studies is to compare the mean voltages within your chosen time window between conditions. This is what was done in the previous sections in both the classic MMN and iMMN comparisons. However, there are other ways in which one can measure the mismatch response, and a few have been used in prior cross-linguistic work. Peak latency, that is the time post-stimulus that the negative deflection of the mismatch response is at its voltage maximum, has been argued to be sensitive to allophony (Miglietta et al., 2013). In the following sections I will present the peak latency and peak amplitude results for each language group, by segment and phonological status.

4.3.2.1 Monolingual English Speakers

The ANOVA on peak amplitude revealed main effects for phonological status, allophonic vs. phonemic contrast ($F(1,96)=5.41$, $p=.002$) and by condition. ($F(3,96)=13.41$, $p=<.001$). No interaction between status and condition was observed.

The ANOVA on peak latency revealed a main effect for phonological status, allophonic vs. phonemic contrast ($F(1,96)=5.65$, $p=.02$), but no other main effects or interaction effects.

In terms of peak amplitude, in the MMN configuration, the allophonic contrast flap (-3.68μV), was significantly more negative than the phonemic contrast fricative (-2.45μV) ($t(12)=-3.40$, $p=.005$). For the iMMN difference wave, the results were similar with the flap being more negative (-3.46μV) than the fricative (-2.27μV) ($t(12)=-1.80$, $p=.06$).

Turning to peak latency, for the classic MMN, while the allophonic contrast flap (330.5ms) was on average faster than the phonemic contrast fricative (339.2ms), this difference did not reach significance (t(12)=-1.16, p=.27). In the iMMN configuration, peak latencies were even less different with flap (336.8ms) and fricative (336.3ms) showing no difference (t(12)=0.09, p=.93).

4.3.2.2 Native Spanish speakers

The ANOVA on peak amplitude revealed a main effect for condition. (F(1,80)=7.12 , p=.0002) but not for status. No interaction between status and condition was observed.

The ANOVA on peak latency revealed a main effect for phonological status, allophonic vs. phonemic contrast (F(1,80)=9.53, p=.002), and a main effect for condition (F(3,80)=3.75, p=.014). There was no interaction effect between status and condition.

For peak amplitude, in the classic MMN configuration, the phonetic contrast flap was more negative (-3.15 μV) than the allophonic contrast fricative (-2.55 μV), however this difference was not significant. (t(10)=-0.96, p=.23). The iMMN was similar, with flap (-2.14 μV) not significantly different from fricative (-2.53 μV) (t(10)=-0.03, p=.98).

For peak latency, in the classic MMN configuration, the phonetic contrast flap was faster (331.32ms) than the allophonic contrast fricative (339.48ms), however this difference was not significant. (t(10)=-0.81, p=.43). In the iMMN, flap had a peak latency of 340.55ms and fricative a peak latency of 340.91ms. These were not different (t(10)=-0.03, p=.98).

4.3.2.3 Advanced Spanish learners

The ANOVA on peak amplitude revealed a main effect for condition, (F(1,104)=9.84 , p<.0001) but not for status. No interaction between status and condition was observed.

The ANOVA on peak latency revealed a main effect for phonological status, allophonic vs. phonemic contrast (F(1,104)=9.15, p=.003), but no main effect for condition. There was no interaction effect between status and condition.

For peak amplitude, in the classic MMN, the allophonic contrast flap was more negative (-4.18 μV) than the phonemic contrast fricative (-3.27 μV), however this difference was not significant (t(13)=-1.43, p=.17). For the iMMN, flap was again more negative (-3.68 μV) than fricative (-3.07 μV), but this was not significant (t(13)=-0.89, p=.39).

For peak latency, the allophonic contrast flap was faster (337.89ms) than the phonemic contrast fricative (352.12ms) in the classic MMN, however this difference was not significant (t(13)=-1.55, p=.14). In the iMMN configuration, flap again had a slightly earlier peak latency (347.01ms) than fricative (350.72ms), but this was not significant (t(13)=-0.33, p=.74).

4.3.2.4 Summary of peak latency & amplitude results

The peak latency and peak amplitude results show a consistent pattern across all language groups with regard to segment. The differences in amplitude and latency do not seem to be driven by allophony, but rather by segment. While the only difference that reached statistical significance was the difference in peak amplitude between the flap and fricative in the monolingual English speakers, in most cases, the flap contrast consistently patterned more negative in amplitude peak and consistently earlier peak latency than the fricative contrast. Table 3 summarizes these results.

	Peak Amplitude MMN (μV)		Peak Amplitude iMMN (μV)		Peak Latency MMN (ms)		Peak Latency iMMN (ms)	
	[d]/[ð]	[d]/[ɾ]	[d]/[ð]	[d]/[ɾ]	[d]/[ð]	[d]/[ɾ]	[d]/[ð]	[d]/[ɾ]
Monolingual English Speakers	-2.45*	-3.68*	-2.27*	-3.46*	339.20	330.50	336.30	336.80
Native Spanish Speakers	-2.55	-3.15	-2.53	-2.14	339.48	331.32	340.91	340.55
Advanced Spanish Learners	-3.27	-4.18	-3.07	-3.68	352.12	337.89	350.72	347.01

Table 3: Within Language Group Peak Amplitude and Latency Results, * denotes significant differences between sound comparisons

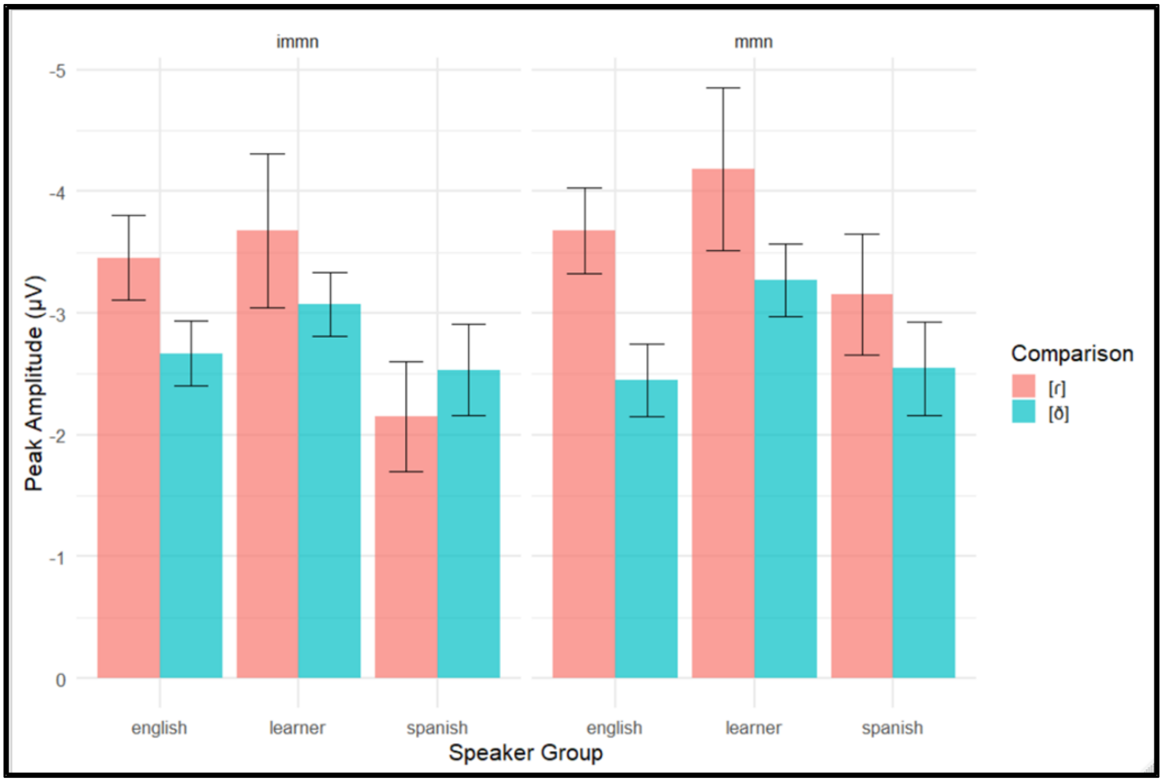

Figure 29: Peak amplitudes for each speaker group and sound comparison in both the classic MMN and iMMN conditions. In both conditions, the only significant differences in peak amplitude between the flap and fricative was for the monolingual English speakers.

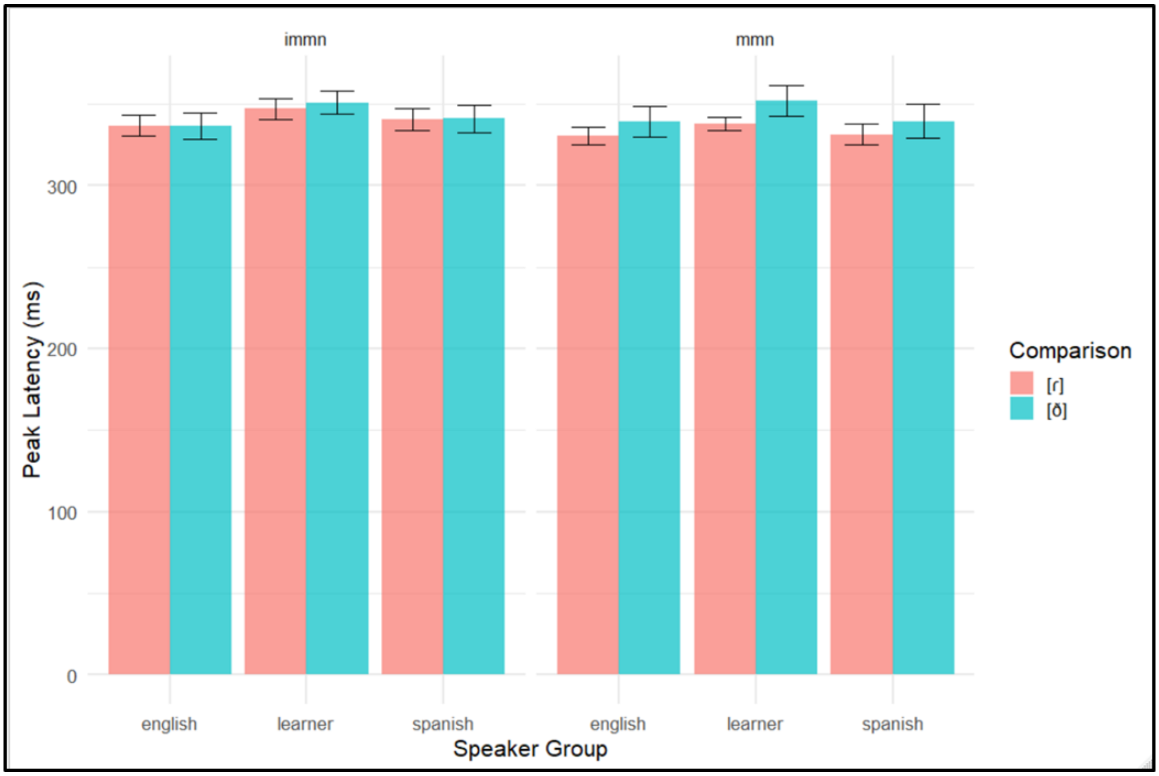

Figure 30: Peak latencies for each speaker group and sound comparison in both the classic MMN and iMMN conditions. While the results show consistently faster peak latencies for the flap in all groups and in both conditions, none of the differences reach statistical significance.

4.3.3 Across-language group results

When comparing the results across the language groups, it is instructive to look at the grand-average difference waves for each group on each contrast. This allows us to see quite clearly how amplitude and peak latency changes for each group. Again, I present the waveforms in both classic MMN (where the deviant response is compared directly to the standard [d] response, e.g. deviant [ɾ] compared to standard [d]) and the iMMN (where the deviant response is compared to the response where said segment is presented as a standard, e.g. deviant [ɾ] compared to standard [ɾ]). These difference waves are show in Figures 31 - 34 below.

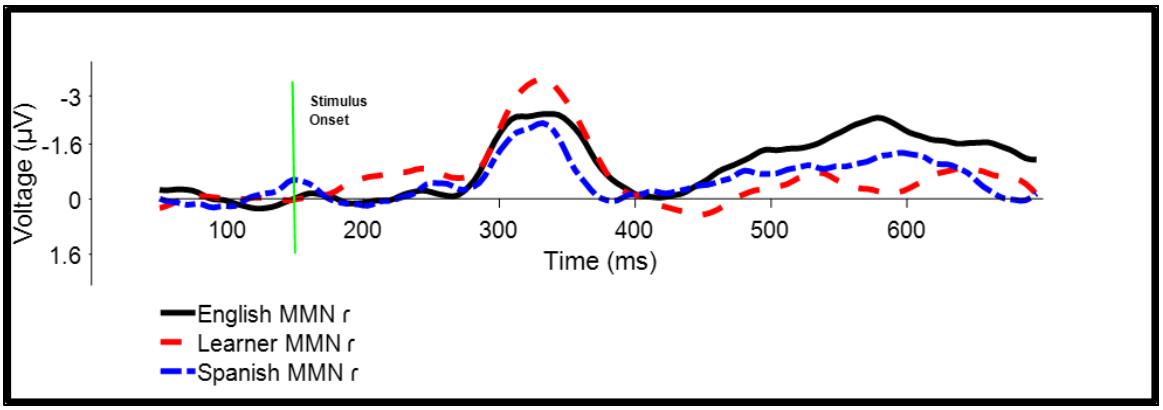

Figure 31: Classic MMN difference waves for all speaker groups. Deviant [r] subtracted from standard [d].

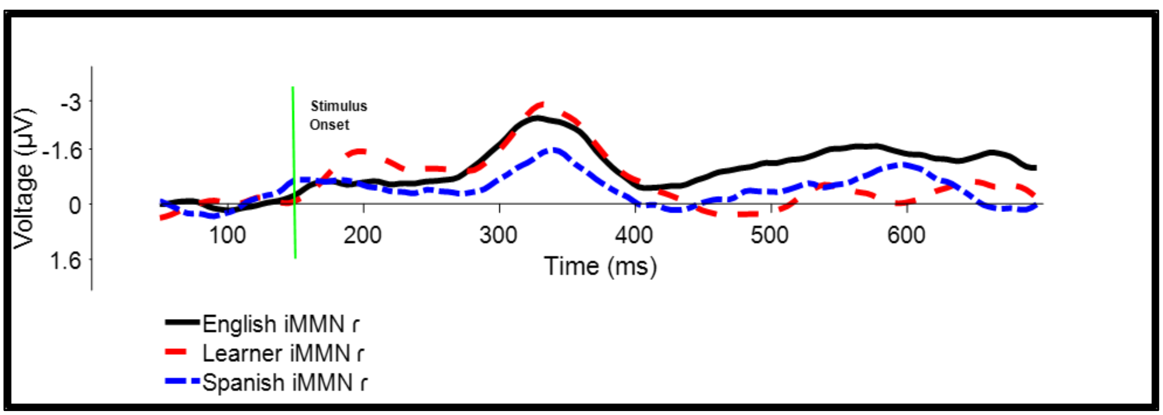

Figure 32: Identity MMN difference waves for all speaker groups. Deviant [r] subtracted from standard [r].

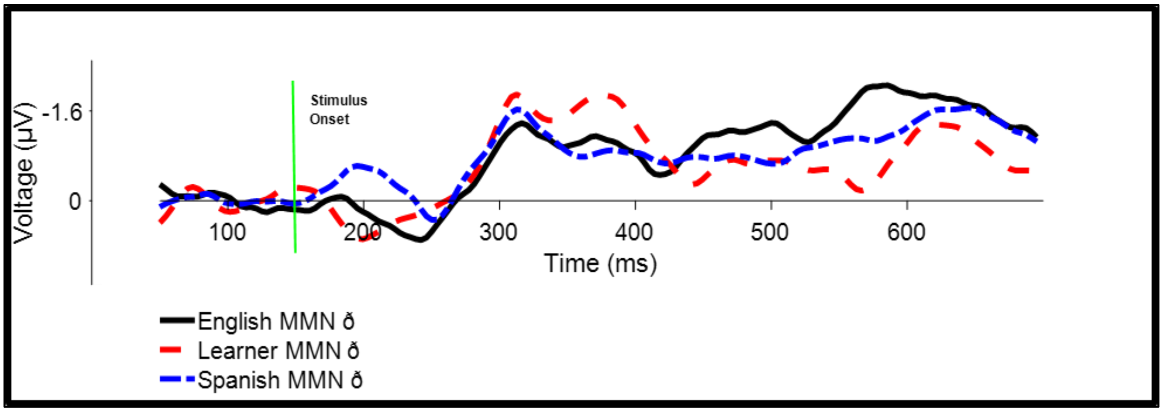

Figure 33: Classic MMN difference waves for all speaker groups. Deviant [ð] subtracted from standard [d].

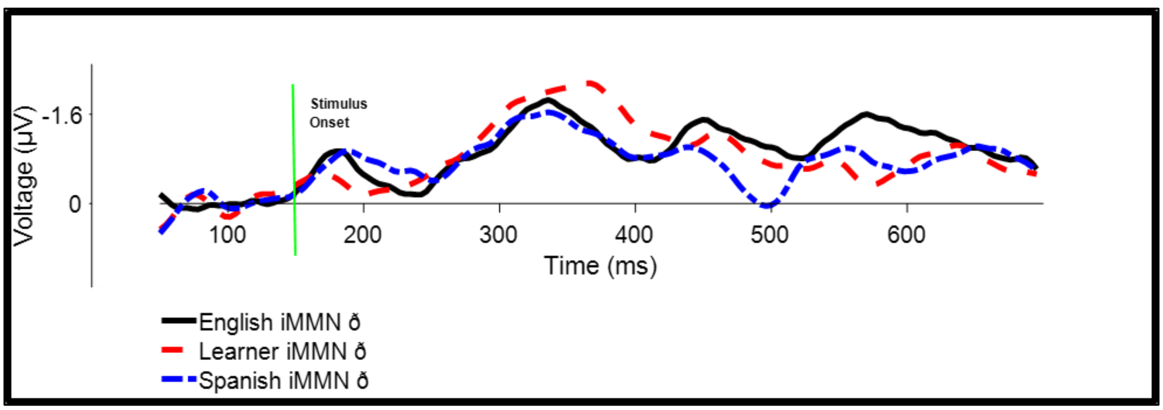

Figure 34: Identity MMN difference waves for all speaker groups. Deviant [ð] subtracted from standard [ð].

4.3.3.1 English compared to Spanish

In terms of mean amplitude, the iMMN response to the [ɾ]-[d] comparison was significantly stronger in the monolingual English speakers (-2.07 μV) than that of the native Spanish speakers (-0.98 μV) ($t(20)=2.34$, $p=.03$). However, in the classic MMN condition, while the pattern is similar to the iMMN condition, the statistical difference is only marginal. Monolingual English speakers had a mean amplitude of -1.81μV whereas the native Spanish speakers had a mean amplitude of -1.15μV. ($t(21.9)=1.36$, $p=.09$).

For the [ð]-[d] contrast, the monolingual English speakers had slightly more negative mean amplitudes (-1.51μV) in the iMMN condition than the native Spanish speakers (-1.36μV), but this difference was not significant ($t(21.6)=0.37$, $p=0.36$). The results for the classic MMN condition were similar. With the mean amplitude of the monolingual English speakers (-1.26μV) being slightly more negative than that the native Spanish speakers (-1.13μV), and no statistical difference in means ($t(21.6)=0.31$, $p=0.38$).

Looking only at peak amplitude, we see slightly different results. For the [ɾ]-[d] contrast we see a peak amplitude for monolingual English speakers in the iMMN condition of -3.46μV compared to -2.15μV for the native Spanish speakers. This result was statistically significant ($t(19.5)=2.29$, $p=0.02$). The MMN condition did not reach statistical significance with the monolingual English speakers having a peak amplitude of -3.7μV and the native Spanish speakers having a peak amplitude of -3.15μV ($t(18.6)=0.86$, $p=0.2$).

The [ð]-[d] comparison shows the monolingual English speakers with a peak amplitude of -2.67μV and the native Spanish speakers with a peak amplitude of -2.53μV in the iMMN condition. These were not significantly different ($t(18.7)=0.29$, $p=0.39$). The

MMN condition was similar with the monolingual English speakers having a slightly lower peak amplitude of -2.45μV than the native Spanish speakers at -2.55μV. This was also not significantly different ((t(19.7)=-0.19, p=0.58).

4.3.3.2 English compared to Learners

For mean amplitude, the iMMN response to the [ɾ]-[d] comparison was nearly identical in the monolingual English speakers (-2.07μV) than that of the advanced Spanish learners (-2.01) (t(23)=-0.09, p=.93). In the classic MMN condition, the advanced Spanish learners trended toward a larger mismatch response than the monolingual English speakers, but this difference did not reach statistical significance. Monolingual English speakers had a mean amplitude of -1.81μV whereas the advanced Spanish learners had a mean amplitude of -2.07μV (t(21.5)=0.39, p=.35).

For the [ð]-[d] comparison, the monolingual English speakers had less negative mean amplitudes (-1.51μV) in the iMMN condition than the advanced Spanish learners (-2.02μV), but this difference was not significant (t(24.9)=1.31, p=0.2). The results for the classic MMN condition were similar. With the mean amplitude of the monolingual English speakers (-1.26μV) being slightly less negative than that the advanced Spanish learners (-1.79μV), and no statistical difference in means (t(24.9)=1.29, p=0.1).

Turning to peak amplitude, we see similar results. For the [ɾ]-[d] comparison we see a peak amplitude for monolingual English speakers in the iMMN condition of -

3.46μV compared to -3.67μV for the advanced Spanish learners. There was no statistically significant difference between these groups (t(19.9)=0.31, p=0.76). The MMN condition also did not reach statistical significance with the monolingual English speakers having a peak amplitude of -3.7μV and the advanced Spanish learners having a slightly higher peak amplitude of -4.18μV (t(19.7)=0.67, p=0.51).

The [ð]-[d] comparison shows the monolingual English speakers with a peak amplitude of -2.67μV and the advanced Spanish learners with a peak amplitude of -3.07μV in the iMMN condition. These were not significantly different (t(24.9)=1.07, p=0.29). The MMN condition, however, showed the monolingual English speakers having a more positive peak amplitude of -2.45μV than the advanced Spanish learners at -3.27μV. This difference was marginally significant (t(24.9)=1.95, p=0.06).

4.3.3.3 Spanish compared to Learners

Mean amplitude for the iMMN response to the [ɾ]-[d] comparison was significantly stronger in the advanced Spanish learners (-2.01μV) than that of the native Spanish speakers (-0.98μV), t(18.1)=1.7, p=.05. In the MMN condition, advanced Spanish learners had a mean amplitude of -2.07μV whereas the native Spanish speakers had a mean amplitude of -1.15μV. This difference was marginally significant (t(20.7)=1.41, p=.08).

For the [ð]-[d] comparison, the advanced Spanish learners had a significantly more negative mean amplitudes (-2.02μV) in the iMMN condition than the native Spanish speakers (-1.36μV) ($t(22.2)=1.66$, $p=0.05$). The results for the MMN condition were similar. With the mean amplitude of the advanced Spanish learners (-1.79μV) being slightly more negative than that the native Spanish speakers (-1.13μV), and this difference was marginally significant($t(22.5)=1.56$, $p=0.06$).

In the peak amplitude analysis, we see for the [ɾ]-[d] comparison a peak amplitude for advanced Spanish learners in the iMMN condition of -3.68μV compared to -2.15μV for the native Spanish speakers. This result was marginally significant ($t(22.1)=1.97$, $p=0.06$). The MMN condition did not reach statistical significance with the advanced Spanish learners having a peak amplitude of -4.18 and the native Spanish speakers having a peak amplitude of -3.15 ($t(22.5)=1.24$, $p=0.2$).

The [ð]-[d] comparison shows the advanced Spanish learners with a peak amplitude of -3.07μV and the native Spanish speakers with a peak amplitude of -2.53μV in the iMMN condition. These were not significantly different ($t(18.9)=1.17$, $p=0.13$). The MMN condition showed a marginally significant differences with advanced Spanish learners having a more negative peak amplitude of -3.27μV than the native Spanish speakers at -2.55μV ($t(20.1)=1.5$, $p=0.07$).

4.3.3.4 Summary of across-language group results

Beginning with mean amplitudes over the 300ms-400ms time window, we find that the monolingual English speakers have significantly larger mismatch responses to the [ɾ] in the iMMN condition than do the native Spanish speakers, though this difference is attenuated in the classic MMN condition. For the [ð], the monolingual English speakers tended to have larger mismatch responses than the native Spanish speakers, but these differences were not significant in either the classic MMN or iMMN condition. These results persisted in the advanced Spanish learners. On both comparisons they patterned more similarly with the monolingual English speakers than the native Spanish speakers. The only significant difference was, like the monolingual English speakers, a stronger classic MMN for the [ɾ] than the native Spanish speakers, though the iMMN condition was marginally significant.

Peak amplitudes paint a slightly different picture. The monolingual English speakers again show a stronger mismatch response for the [ɾ] comparison than the native Spanish speakers in the iMMN condition, but no significant difference on the [ð] comparison. The advanced Spanish learners diverge on some measures from both the monolingual English and native Spanish groups. On the [ɾ], the advanced Spanish learners pattern more like the monolingual English speakers, showing a marginally statistical difference in the iMMN condition than the native Spanish speakers. This did not persist in the classic MMN condition. For the [ð], the advanced Spanish learners showed marginally stronger classic MMN responses when compared to both the monolingual English speakers and the native Spanish speakers. However, the iMMN condition showed no such difference.

To summarize, both the monolingual English speakers and the advanced Spanish learners show stronger mismatch responses to the [ɾ] comparison than the native Spanish speakers, both in mean amplitude and peak amplitude. This is contra expectations given that the [ɾ] is a phoneme in Spanish, but not in English. If anything, the advanced Spanish learners trend toward stronger mismatch responses than the monolingual English speakers. For the [ð], the monolingual English speakers and the native Spanish speakers do not differ significantly in mean or peak amplitude, but the advanced Spanish learners had marginally stronger mismatches than both the monolingual English and native Spanish groups in the classic MMN condition.

Chapter 5: Discussion

In this chapter I'll discuss and interpret the results of the experiments in chapters 3 and 4, and discuss their contributions to the broader field. I'll begin with the behavioral study in chapter 3.

5.1 Interpretation of the Behavioral results

The behavioral study detailed in Chapter 3 was an AXB discrimination task run on three speaker groups. Groups of monolingual English speakers, native Spanish speakers, and native English speaking advanced learners of Spanish all participated in the same AXB task where they were asked to discriminate between multiple tokens of [ɾ], [ð], and [d]. The expectation in this study was that these groups would differ in their ability to discriminate the consonants from one another based on the phonological status of the sounds in their native language. The purpose of testing the advanced Spanish learner group was to determine whether that group's discrimination ability would become more 'native Spanish-like' when compared to their monolingual English counterparts. We predicted that the monolingual English speakers would have difficulties discriminating [ɾ] from [d] because [ɾ] is an allophone of [d] in English, while at the same time discriminating very well [ð] from [d]. For full results, see section 3.3. In summary, the predictions were borne out. In every measure; reaction time, accuracy, and A′, the language groups more easily discriminated the pair that was phonemically contrastive in their respective language. That is, the flap [ɾ] was more easily discriminated from the [d] by the native Spanish group and the [ð] was more easily discriminated from the [d] but the monolingual English group. These results replicate those found in Boomershine et al., (2008). The current study adds an additional novelty of a native English-speaking group of advanced Spanish learners. These learners appear to perform like the native Spanish speakers when discriminating flap [ɾ] from [d], significantly outperforming the monolingual English group. They also appear to show some bi-directional influence of their L2 as their discriminability of the fricative [ð] from [d] was, while better than the native Spanish speakers, significantly impaired compared to the monolingual English group. This decrease in discriminability on a L1 phonemic contrast, is apparently due to having to juggle more than one representation of [ð] in their minds. While this result was not expected, it is not without precedent. Influence of the L2 on the L1 has been shown in previous work, including in EEG work (Caramazza et al., 2005; Guion, 2003; Peltola et al., 2003; Sancier & Fowler, 1997)).

Returning to the initial questions laid out in Chapter 1 of this dissertation; Do we see evidence of L2 category formation in second language learners? It would appear from the behavioral data that this is a possibility, and that the advanced Spanish learners have created a category for the flap [ɾ]. However, reaction times in this type of discrimination task are on the order of seconds; an eternity when it comes to the brain.

There could be all kinds of scaffolding being built to get to the correct answer in the three seconds or so it takes a participant to provide their answer. This motivated the need for the EEG study.

Do we see evidence that perception is modulated by level or representation? Yes, like Boomershine et al. (2008), we clearly see that when sounds are in allophonic (surface) contrast in a language they are more difficult to perceive than when they are in phonemic contrast. Using multiple standard tokens in the EEG experiment task and comparing results from the MMN and iMMN condition is an attempt at sussing out these differences experimentally.

Are we able to use the MMN to probe phonemic representations? The EEG experiment in Chapter 4 was designed to probe this question specifically. I will now turn to the Electrophysiological study.

5.2 Interpretation of the Electrophysiological results

MMN studies, specifically studies involving L2, have provided disparate results that can be difficult to reconcile. Generally, these studies have been designed to probe phonological information via the MMN, usually assumed to mean phonemic representations. Perhaps most clearly in the Kazanina et al., (2006) study, the differences in the MMN between Russian and Korean speakers are said to be due to differences in the status of the [t] and [d] sounds in those languages. Russian contrasts

[t] and [d] phonemically whereas Korean does not; consequently, the Russian speakers had a large MMN to this contrast whereas Korean speakers did not. Results such as these tend to mirror behavioral results on perception, where listeners more readily perceive phonemic contrasts than allophonic ones. The prediction then has been that the MMN should be larger in amplitude when a contrast is very salient to the listener, and attenuated (or nonexistent) on contrasts that are difficult to perceive. However, cross-linguistic MMN results have not been so clear. For example, Winkler, et al., (1999) tested Hungarian and Finnish speakers on both in-category and across-category vowel contrasts and found that both contrasts resulted in an MMN in both languages with the across-category contrast producing a larger MMN. In Finnish, the across-category contrast elicited a double peak MMN (presented as distinct early and late responses). In Hungarian, the across-category had only a single peak consistent with the early response of the Finns. Näätänen's work on Finnish and Estonian produced similar results for contrastive pairs (large MMNs) but with deviants not present in the listener's L1, an attenuated MMN was produced (Naatanen & Alho, 1997).

In my study, I contrasted [d] with [ɾ] and [ð] in both English and Spanish speakers, along with advanced learners of Spanish. These contrasts provide both a phonemic and allophonic pair to the speakers. If the EEG results were to mirror the behavioral results, we would expect small mismatch effects for the flap-stop [ɾ]-[d] comparison in the monolingual English speakers along with large mismatch responses to the fricative-stop [ð]-[d] comparison. We would expect this pattern to be reversed in the native Spanish speakers. For the advanced Spanish learners, we would expect to see an increased MMN for the flap-stop [ɾ]-[d] comparison and perhaps a slightly attenuated MMN for the fricative-stop [ð]-[d] comparison when compared to the monolingual English group. Essentially, they should become more 'Spanish-like'. This was both the predicted and borne out result of the behavioral study. However, I found that both the phonemic and allophonic contrasts produce MMNs in all subject groups, though there are some differences in amplitude and latency. This is surprising given the behavioral results where listeners are poor at discriminating the allophonic contrasts in their L1.

Compared to the results of the EEG study, the results of the discrimination study detailed in Chapter 3 are very straightforward. Each of the predictions made is borne out and it is a near replication of the Boomershine et al., (2008) study. This reinforces the validity of the study and allows us to be confident in the stimuli. It follows then that the unexpected results of the EEG study are not due to issues with the stimuli or methodology. So how then do we reconcile the unexpected EEG results with the comparatively beautiful behavioral results? First, it is important to acknowledge that given the past EEG work of this nature, these results are not that surprising. Herd, (2011) showed similar results for her English subjects. Her English subjects produced a large MMN for the allophonic contrast of [d] and [ɾ]. She offers that this may be due to [ɾ] being an allophone of both [t] and [d] in English, and that in the context of one of those stops, an English listener may be perceiving the other, essentially producing an MMN you would expect for a phonemic contrast. Barrios et al., (2016) ran an MEG study with monolingual English speakers, native Spanish speakers, and advanced L2 English learners on the same contrasts as the ones in this dissertation and found no statistical significance in the differences in groups. However, they found similar mismatch responses to both the [d]-[ɾ] comparison and the [d]-[ð] comparison in the English speakers, while Spanish speakers had similar mismatch responses to English speakers on the [d]-[ɾ] as well as an unexpected mismatch response for [d]-[ð]. Barrios does ultimately claim there is evidence for the advanced L2-English learners having acquired the proper phonemic contrast despite the allophonic status in L1, but this is made solely on the grounds that there is a statistically significant mismatch negativity for both the phonemic and allophonic contrasts among that language group. However, the L1-English and L1-Spanish group do not pattern as expected, nor do they match the behavioral data. The L1-English speakers show no significant mismatch for either the phonemic or allophonic contrast, and most surprisingly, the L1-Spanish speakers had only a significant mismatch for the allophonic contrast. These results are strikingly similar to mine, though that makes them no less difficult to interpret. Miglietta et al., (2013) also found significant MMNs for allophonic contrasts in Italian, both an allophonic and phonemic contrast elicited MMNs but the phonemic pair had a shorter peak latency. In my study, I found the opposite, that the allophonic contrasts displayed faster peak latencies than the phonemic ones.

So, the MMN results are not as simple as appealing to phonemic status where contrastive equates to an MMN and non-contrastive no MMN. It also seems that variations in phonetic detail used by the listener can alter the MMN by way of latency or amplitude. It is important to understand that mismatch negativity is simply a marker of change detection. It is not in any way 'linguistic' and seems to be stimulus agnostic. Although it can be used as evidence of categorical perception, it does not index allophony or phonemic categories. Those are simply aspects of the stimuli that could result in changes to the ERP waveform. It is predominately found on fronto-central electrodes, but localization of brain activity cannot be reliably inferred from scalp distributions. We must resist the temptation to ascribe meaning to the ERP components where there is none. So, it falls to the researcher to design stimuli and experiments in such a way to constrain the type of change that could be detected in the signal.

There are some amplitude differences in mismatch responses to the stimuli that vary by speaker group. Both the monolingual English speakers and the advanced Spanish learners showed stronger mismatch responses to the [ɾ]-[d] comparison than the native Spanish speakers. The advanced Spanish learners exhibited a stronger mismatch response than the native Spanish speakers on the [ð]-[d] comparison.

Regardless, we must confront the fact that the segments themselves pattern alike no matter the group. That is, the flap sound in all groups appears to elicit a higher amplitude mismatch response than the fricative sound. Also, peak latencies are consistent across language groups. The flap response peaks earlier than the fricative response in all cases. These facts point to the ERP waveform being a product of the phonetic detail of the individual segments themselves rather than phonological status such as allophone or phoneme. One way to test this is to calculate acoustic distances between the stimuli to see if mismatch response correlates with distance. To this end I calculated the acoustic similarity of the stimuli and compared the mean distances of the [ɾ] and [ð] to the standard [d]. We find that the mean acoustic distance between [ɾ] and [d] is larger than the mean acoustic distance between the [ð] and [d]. This difference in acoustic distance correlates with a larger mismatch response for the [ɾ] than the [ð]. Acoustic distance was calculated using Phonological CorpusTools (Hall et al., 2017; Mackie, et al., 2014)acoustic similarity function to compute mel-fruquency cepstrum coefficients (MFCC) (Delvaux & Soquet, 2007; Mielke, 2012). Distances were computed using dynamic time warping (Sakoe & Chiba, 1971). Figure 35 shows the mean acoustic distance for each pair of sounds. Essentially, these tools map the spectrum of the sounds onto a two-dimensional space so that a linear distance between them can be calculated. You can see in Figure 35 that the Euclidean distance between sounds is higher when the segments are different as in the case when comparing my [ɾ], [d], and [ð] stimuli. When different tokens of identical segments are compared, the acoustic distance is relatively low. When tokens of different segments are compares, the acoustic distance is much higher, with the [ɾ]-[d] comparison being the most distant from one another.

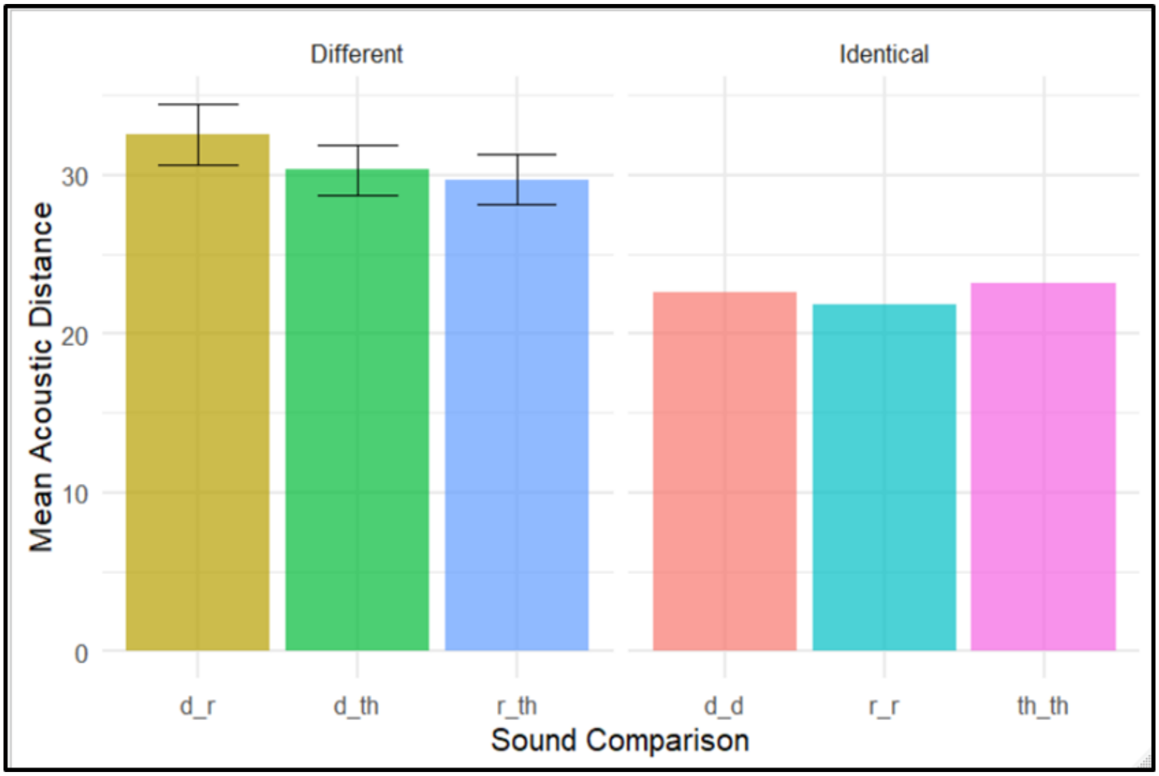

Figure 35: Acoustic distance calculations for each sound comparison showing that identical sounds have relatively low acoustic distance while [ɾ] is further from [d] than [ð] is from [d].

Tying the differences to the MMN response to the phonetic categories of the segments themselves, rather than appealing to phonemic status seems reasonable given my results, but what about previous work that has provided evidence that the MMN is sensitive to phonemic status? It is possible that for the given sound comparisons in my study subjects were simply unable to construct a standard based on phonemic information available in the stimuli. However, it may also be possible to reinterpret previous results as also showing sensitivity to phonetic, and not phonemic differences. The Näätänen et al., (1997) study on Finns and Estonians for example, it could be that the Finns simply assimilated the [õ] vowel as a poor exemplar of their existing phonetic category for [ö], causing the attenuated MMN to that stimuli when compared to the [ö] deviants. In their study on Russian and Korean stops, Kazanina et al, (2006) showed that in the Russian speakers, a large MMN was seen for the phonemic /t/-/d/ contrast using a variable standards paradigm. The lack of a similar MMN in the Korean speakers was claimed to be due to the voiced stop [d] being an allophone of the voiceless /t/ in that language. However, there is evidence for passive voicing in Korean (Jun 1996, Wright, 2007). If voicing is not phonological in Korean, then there would not be a clear phonetic category distinction between [t] and [d]. There would just be a single phonetic category for [t] that gets voiced only in production. Thus, if the MMN was sensitive to phonetic category, we would expect no MMN in the Korean speakers. In the Miglietta et al., (2013), where it was shown that a phonemic vowel contrast in Italian had a faster peak latency than an allophonic one, the authors claim that this is evidence for the MMN being sensitive to allophony. Because there was no cross-linguistic comparison, and the test was done on a single pair of contrasts, we cannot be sure that the phonetics of the sounds themselves are driving these results, especially given that both contrasts displayed a significant MMN.

One interesting finding is that the advanced Spanish learners have marginally higher peak amplitudes for both the [ɾ]-[d] (iMMN) and [ð]-[d] (MMN) comparisons than their native Spanish speaking counterparts (p=.06 and .07 respectively). The learners also have a marginally significant higher peak amplitude on the [ð]-[d] comparison than the monolingual English speakers in the classic MMN comparison. For the [ɾ]-[d] comparison, we would have expected that the advanced Spanish learners would have a stronger mismatch response than the monolingual English speakers if they were indeed creating a new category, and while the results do not reach significance, the trend is toward a higher peak amplitude for the learners. This increased amplitude could be indicative the advanced Spanish learners forming a new category for [ɾ]. Some caution should be taken with this interpretation due to the fact that the native Spanish speakers already show attenuated mismatch responses to the [ɾ]-[d] comparison when compared to both groups of native English speakers. It could be that the additional ways to use [ɾ], as an allophone to both [d] and [t] in English, as well as a new phonemic contrast for the advanced Spanish learners, results in even greater perceptual contrast in these learners. For the reasons stated above, and because the differences between groups are marginal if they exist at all, one should be wary of putting much stock in the peak amplitude results. This illustrates a more general issue of researcher degrees of freedom when using ERPs in language research that I discuss in the section below.

In light of the ERP results, I return to the original research questions highlighted in Chapter 1. Do we see evidence of L2 category formation in second language learners? Even though the advanced Spanish learners do show a strong mismatch response to the [ɾ]-[d] comparison, it would be difficult to claim that this is due to them creating a category for [ɾ], because the monolingual English speakers already show a similar mismatch response to this comparison. The learners also show a stronger mismatch response to the [ð]-[d] comparison than both the monolingual English speakers and the native Spanish speakers. So, while the advanced Spanish learners do appear to differ from their monolingual counterparts, they are not necessarily becoming more "Spanish-like" in their ERP results. The advanced Spanish learners showed less discriminability for the [ð]-[d] contrast than the monolingual English speakers in the behavioral task, something that could be indicative of the Spanish spirantization rule increasing the amount of processing needed, they also show slightly stronger mismatch responses to this contrast in the ERP results. The prediction was that the Spanish learners might show an attenuated mismatch response to this contrast with increased L2 influence, as was found in Peltola et al., (2003). Instead, the increased processing demand of the allophonic rule could instead be having an additive effect on the mismatch response.

Do we see evidence that perception is modulated by level or representation? Using multiple real-speech tokens to build the standard is an attempt to control for acoustic differences between the segments. If we are able to build a standard in a listener's mind from tokens that are not identical, we can be surer that the phonological properties of those sounds are being used to build that standard. Additionally, presenting both the classic MMN and iMMN results in theory would allow us to distinguish further mismatch responses based on acoustic differences between the standard and the deviant sound, and mismatches based on the phonological properties of those sounds. Using the variable standard paradigm coupled with measuring the iMMN was an attempt to get closer to the underlying level of representation. Ultimately, while there are some slight differences when looking at the classic MMN and the iMMN results, they generally pattern similarly. So, while the behavioral study in Chapter 3 provides evidence for this, the EEG results are less clear.

Are we able to use the MMN to probe phonemic representations? Despite measuring the MMN in multiple ways and this being previously claimed in the literature, I find no evidence for the MMN being sensitive to phonemic category, but clear evidence that it is sensitive to phonetic category.

5.3 Conclusion

The first human electroencephalograph was taken by Hans Berger in 1924 and since EEG has been utilized in a number of research and clinical settings to study the brain. In the last several decades, linguistics and other language sciences have taken to the technique. For language research, it is still early days. Various subfields have utilized other ERP components to test and refine their theories, such as the semantic N400 (Kutas & Hillyard, 1980) or syntactic P600 (Osterhout & Holcomb, 1992). The MMN is a relative newcomer, its use in studying phonology increasing in just the last twenty years. ERP research is time-consuming; initial set-up and subject preparation can be uncomfortable, the tasks are often monotonous, and usually recordings are quite long. Because of this, data is precious. From study design to analysis, researchers have many degrees of freedom when doing this work. One can choose from several measurements, electrode placements, recording settings, and a slew of post-processing procedures. Once the data is recorded and measured, there are additional choices to be made about statistical tests. It is extremely common in the literature for these choices to be made without mentioning the justification of one choice over another (Gelman & Loken, 2013; Gelman & Stern, 2006; Luck & Gaspelin, 2017; Simmons, Nelson, & Simonsohn, 2011). Part of the novelty of my study is simply acknowledging the myriad ways that researchers have analyzed ERP waveforms, specifically when utilizing the MMN, and attempting to be open about how that choice can affect the results in sometimes dramatic ways. This is the reason I chose to look at three different ERP measurements; mean amplitude over an analysis window, peak amplitudes in that window, and peak latencies in that window. Each measurement has been used in prior work and they have all been said to be modulated by phonological status, or at the very least, strength of perceptual contrast. In addition, I chose to be explicit about how I was calculating the mismatch response by presenting all measurements in both the classic MMN (standard response – deviant response) and the iMMN (response of the target sound presented in a standard context – response of the target sound presented in a deviant context). The latter being an attempt to tap into a more 'phonological' response by controlling for any effects the standard might have as it relates to the deviant, such as differences in acoustics between [d] and the target sounds. This type of "ERPology" is a necessary step to undertake before any EEG experiment can tell us something meaningful about the brain.

This dissertation contributes to the growing body of neurolinguistic work on second language acquisition and adds to the work on the mismatch negativity ERP component. It also extends previous speech perception work on second language learners by both replicating previous results and adding new results from electroencephalography that show incongruency with behavioral results. It provides evidence that advanced second language learners may be able to create categories that do not exist in the L1, and also that proficiency in an L2 may affect L1 perception. It also suggests that while the Mismatch Negativity is sensitive to phonetic category, it may not be sensitive to phonemic category. Previous studies that have claimed as much, can be re-interpreted as testing phonetic category.

APPENDICES

APPENDIX A

Electrode Cluster Locations

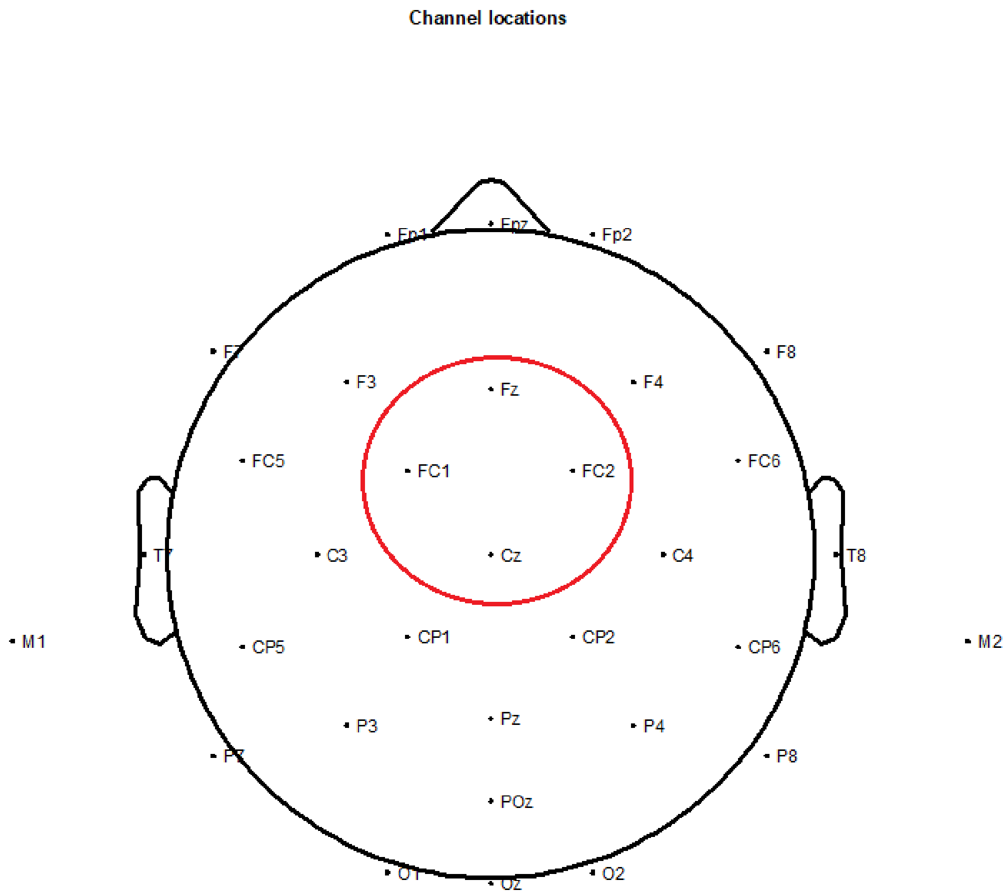

Figure 36: EEG Channel Locations for MMN Cluster Analysis

APPENDIX B

Participant Survey

Native Spanish Speakers

1. What is your native language?

2. Age/Sex/Handedness?

3. In what country were you born and raised?

4. At what age was your first exposure to English?

5. At what age did you take formal English classes?

6. How long have you been in the United States?

7. What is the primary language spoken in your home?

8. Level of education?

Advanced Spanish Learners

1. What is your native language?

2. Age/Sex/Handedness?

3. When was your first exposure to Spanish?

4. How much consecutive time have you had in a majority Spanish-speaking country?
5. At what age did you start formal Spanish classes?

6. How many years of Spanish instruction have you had?

7. How do you plan to use Spanish in the future?

8. Level of Education?

REFERENCES

REFERENCES

Aulanko, R., Hari, R., Lounasmaa, O. V., Näätänen, R., & Sams, M. (1993). Phonetic invariance in the human auditory cortex. *NeuroReport*.

Barrios, S., Jiang, N., & Idsardi, W. J. (2016). Similarity in L2 Phonology: Evidence from L1 Spanish late-learners' perception and lexical representation of English vowel contrasts. *Second Language Research*.

Barrios, S. L. . (2013). *Similarity In L2 Phonology*. Doctoral dissertation, University of Maryland.

Baudena, P., Halgren, E., Heit, G., & Clarke, J. M. (1995). Intracerebral potentials to rare target and distractor auditory and visual stimuli. III. Frontal cortex. *Electroencephalography and Clinical Neurophysiology*.

Beddor, P. S., & Strange, W. (2005). Cross-language study of perception of the oral–nasal distinction. *The Journal of the Acoustical Society of America*.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*.

Best, C. T. (1995). A direct realist view of cross-language speech perception. *Speech Perception and Linguistic Experience. Issues in Cross-Language Research*.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America*.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of Perceptual Reorganization for Nonnative Speech Contrasts: Zulu Click Discrimination by English-Speaking Adults and Infants. *Journal of Experimental Psychology: Human Perception and Performance*.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In *Language experience in second language speech learning: In honor of James Emil Flege*.

Birdsong, D., & Molis, M. (2001). On the Evidence for Maturational Constraints in Second-Language Acquisition. *Journal of Memory and Language*.

Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer [Computer program]. *Glot International*, Vol. 5, pp. 341–347.

Boomershine, A., Hall, K. C., Hume, E., & Johnson, K. (2008). The impact of allophony versus contrast on speech perception. *Contrast in Phonology - Theory, Perception, Acquisition*, 145–172.

Browman, C., & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook*, *3*, 219–252.

Browman, C., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, *6*(2), 201–251.

Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (2005). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *The Journal of the Acoustical Society of America*.

Celsis, P., Boulanouar, K., Doyon, B., Ranjeva, J. P., Berry, I., Nespoulous, J. L., & Chollet, F. (1999). Differential fMRI responses in the left posterior superior temporal gyrus and left supramarginal gyrus to habituation and change detection in syllables and tones. *NeuroImage*.

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience*.

Cheour, M, Korpilahti, P., Martynova, O., & Lang, A. H. (2001). Mismatch Negativity and Late Discriminative Negativity in Investigating Speech Perception and Learning in Children and Infants. *Audiology and Neurotology*, *6*(1), 2–11.

Cheour, Marie, Shestakova, A., Alku, P., Ceponiene, R., & Näätänen, R. (2002). Mismatch negativity shows that 3-6-year-old children can learn to discriminate non-native speech sounds within two months. *Neuroscience Letters*, *325*(3), 187–190.

Cook, V. J. (1992). Evidence for Multicompetence. *Language Learning*.

Csépe, V., Pantev, C., Hoke, M., Hampson, S., & Ross, B. (1992). Evoked magnetic responses of the human auditory cortex to minor pitch changes: localization of the mismatch field. *Electroencephalography and Clinical Neurophysiology/ Evoked Potentials*.

Deacon, D., Gomes, H., Nousak, J. M., Ritter, W., & Javitt, D. (2000). Effect of frequency separation and stimulus rate on the mismatch negativity: An examination of the issue of refractoriness in humans. *Neuroscience Letters*.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*.

Delvaux, V., & Soquet, A. (2007). The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica*.

Dien, J. (1998). Issues in the application of the average reference: Review, critiques, and recommendations. *Behavior Research Methods, Instruments, and Computers*.

Donaldson, W. (1993). Accuracy of d′ and A′ as estimates of sensitivity. *Bulletin of the Psychonomic Society*.

Dupoux, E., Pallier, C., Sebastian, N., & Mehler, J. (1997). A destressing "deafness" in French? *Journal of Memory and Language*.

Eulitz, C., & Lahiri, A. (2004). Neurobiological Evidence for Abstract Phonological Representations in the Mental Lexicon during Speech Recognition. *Journal of Cognitive Neuroscience*, *16*(4), 577–583.

Eulitz, C., & Lahiri, A. (2004b). Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition. *Journal of Cognitive Neuroscience*, *16*(4), 577–583.

Face, T. L. (2006). Intervocalic rhotic pronunciation by adult learners of Spanish as a second language. *Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*.

Face, T. L., & Menke, M. R. (2009). Acquisition of the Spanish voiced spirants by second language learners. In *Selected proceedings of the 11th Hispanic Linguistics Symposium*.

Flege, J. E. (1991). *Perception and production: The relevance of phonetic input to L2 phonological learning* (T. Heubner & C. W. Ferguson, Eds.). Amsterdam: John Benjamins.

Flege, J. E. (1995). Second Language Speech Learning: Theory, Findings, and Problems. *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*.

Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions. In J. Cole & J. I. Hualde (Eds.), *Laboratory Phonology 9* (pp. 353–381). Berlin, New York: Mouton de Gruyter.

Flege, J. E. (2011). Assessing constraints on second-language segmental production and perception. In *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities*.

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America*.

Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics*, *14*, 3–28.

Fowler, C. A. (1989). Real Objects of Speech Perception: A Commentary on Diehl and Kluender. *Ecological Psychology*.

Fowler, C. A. (2005). Auditory perception is not special: We see the world, we feel the world, we hear the world. *The Journal of the Acoustical Society of America*.

Gelman, A., & Loken, E. (2013). *The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time∗*.

Gelman, A., & Stern, H. (2006). *The difference between "significant" and "not significant" is not itself statistically significant*.

Gerrits, E., & Schouten, M. E. H. (2004). Categorical perception depends on the discrimination task. *Perception & Psychophysics*, *66*(3), 363–376.

Guion, S. G. (2003). The vowel systems of Quichua-Spanish bilinguals: Age of acquisition effects on the mutual influence of the first and second languages. *Phonetica*.

Halgren, E., Baudena, P., Clarke, J. M., Heit, G., Liégeois, C., Chauvel, P., & Musolino, A. (1995). Intracerebral potentials to rare target and distractor auditory and visual stimuli. I. Superior temporal plane and parietal lobe. *Electroencephalography and Clinical Neurophysiology*.

Halgren, E., Marinkovic, K., & Chauvel, P. (1998). Generators of the late cognitive potentials in auditory and visual oddball tasks. *Electroencephalography and Clinical Neurophysiology*.

Hall, K. C., Allen, B., Fry, M., Johnson, K., Lo, R., Mackie, S., & McAuliffe, M. (2017). *Phonological CorpusTools*.

Hari, R., Hämäläinen, M., Ilmoniemi, R., Kaukoranta, E., Reinikainen, K., Salminen, J., Sams, M. (1984). Responses of the primary auditory cortex to pitch changes in a sequence of tone pips: Neuromagnetic recordings in man. *Neuroscience Letters*.

Harris, J. w. (1969). *Spanish Phonolgy*. Cambridge, MA: MIT Press.

Hartshorne, J. K., Tenenbaum, J. B., & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition*.

Herd, W. (2011). *The preceptual and production training of /d flap r/ in L2 Spanish: Behavioral, psychologistics, and neurolinguistic evidence*. Doctoral dissertation, University of Kansas.

Herd, W., Jongman, A., & Sereno, J. (2010). An acoustic and perceptual analysis of /t/ and /d/ flaps in American English. *Journal of Phonetics*, *38*(4), 504–516.

Hestvik, A., & Durvasula, K. (2016). Neurobiological evidence for voicing underspecification in English. *Brain and Language*.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*.

Jun Zhang, & Shane T. Mueller. (2005). A note on ROC analysis and non-parametric estimate of sensitivity. *Psychometrika*.

Kahn, D. (1976). *Syllable-based generalizations in English phonology*. Doctoral dissertation, Massachusetts Institute of Technology.

Kazanina, N., Phillips, C., & Idsardi, W. (2006). The influence of meaning on the perception of speech sounds. *Proceedings of the National Academy of Sciences*, *103*(30), 11381–11386.

Kenstowicz, M. J. (1994). *Phonology in generative grammar*. Oxford: Blackwell.

Kilner, J. M., Stephan, K. E., Friston, K. J., & Garrido, M. I. (2009). The mismatch negativity: A review of underlying mechanisms. *Clinical Neurophysiology*, *120*(3), 453–463.

Kutas, M., & Hillyard, S. A. (1980). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology*, *11*, 99–116.

Ladefoged, P., & Johnson, K. (2011). A Course in Phonetics (6th ed). In *Wadsworth*.

Lenneberg, E. H. (1967). Biological foundations of language. New York: John Wiley and Sons, 1967. *Behavioral Science*, *13*(6), 493–495.

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*.

Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, Massachusetts: MIT Press.

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*.

Mackay, I. R. A., & Flege, J. E. (2004). Effects of the age of second language learning on the duration of first and second language sentences: The role of suppression. *Applied Psycholinguistics*.

Mackie, S., Currie Hall, K., Allen, B., Michael, M., & Fry, M. (2014). Phonological CorpusTools: A free, open-source tool for phonological analysis. *Presented at the 14th Conference for Laboratory Phonology*. Tokyo,Japan.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection Theory: A User's Guide*. Taylor & Francis.

Mäntysalo, S., & Näätänen, R. (1987). The duration of a neuronal trace of an auditory stimulus as indicated by event-related potentials. *Biological Psychology*, *24*(3), 183–195.

Martínez-Celdrán, E., Fernández-Planas, A. M., & Carrera-Sabaté, J. (2003). Castilian Spanish. *Journal of the International Phonetic Association*.

McCarthy, J. J. (2010). Derivations and levels of representation. In *The Cambridge Handbook of Phonology*.

Mielke, J. (2012). A phonetically based metric of sound similarity. *Lingua*.

Miglietta, S., Grimaldi, M., & Calabrese, A. (2013). Conditioned allophony in speech perception: An ERP study. *Brain and Language*, *126*(3), 285–290.

Molholm, S., Martinez, A., Ritter, W., Javitt, D. C., & Foxe, J. J. (2005). The neural circuitry of pre-attentive auditory change-detection: An fMRI study of pitch and duration mismatch negativity generators. *Cerebral Cortex*.

Näätänen, R., & Alho, K. (1995). Mismatch negativity--a unique measure of sensory processing in audition. *Int J Neurosci*, *80*(1–4), 317–337.

Näätänen, R., & Alho, K. (1997). Mismatch negativity--the measure for central sound representation accuracy. *Audiol Neurootol*, *2*(5), 341–353.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., … Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, *385*(6615), 432–434.

Näätänen, R., Paavilainen, P., Alho, K., Reinikainen, K., & Sams, M. (1989). Do event-related potentials reveal the mechanism of the auditory sensory memory in the human brain? *Neurosci Lett*, *98*(2), 217–221.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, Vol. 118, pp. 2544–2590.

Näätänen, R., Paavilainen, P., Titinen, H., Jiang, D., & Alho, K. (1993). Attention and mismatch negativity. *Psychophysiology*.

Näätänen, R, Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, *42*, 313–329.

Näätänen, R., & Michie, P. T. (1979). Early selective-attention effects on the evoked potential: A critical review and reinterpretation. *Biological Psychology*.

Näätänen, R., Pakarinen, S., Rinne, T., & Takegata, R. (2004). The mismatch negativity ({MMN}): towards the optimal paradigm. *Clinical Neurophysiology*, *115*(1), 140–144.

Osterhout, L., & Holcomb, P. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*, 785–806.

Paradis, M. (1985). On the representation of two languages in one brain. *Language*

Sciences.

Patterson, D., & Connine, C. M. (2001). Variant frequency in flap production: A corpus analysis of variant frequency in American english flap production. *Phonetica*, *58*(4), 254–275.

Peltola, M. S., Kujala, T., Tuomainen, J., Ek, M., Aaltonen, O., & Näätänen, R. (2003). Native and foreign vowel discrimination as indexed by the mismatch negativity (MMN) response. *Neuroscience Letters*, *352*(1), 25–28.

Peltola, M. S., Kuntola, M., Tamminen, H., Hämäläinen, H., & Aaltonen, O. (2005). Early exposure to non-native language alters preattentive vowel discrimination. *Neuroscience Letters*, *388*(3), 121–125.

Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., Roberts, T. (2000). Auditory cortex accesses phonological categories: An MEG mismatch study. *Journal of Cognitive Neuroscience*, *12*(6), 1038–1055.

Pulvermüller, F., & Shtyrov, Y. (2006). Language outside the focus of attention: The mismatch negativity as a tool for studying higher cognitive processes. *Progress in Neurobiology*.

Sabri, M., & Campbell, K. B. (2001). Effects of sequential and temporal probability of deviant occurrence on mismatch negativity. *Cognitive Brain Research*.

Sakoe, H., & Chiba, S. (1971). A dynamic programming approach to continuous speech recognition. *Proceedings of the Seventh International Congress on Acoustics*.

Sams, M., Paavilainen, P., Alho, K., & Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalography and Clinical Neurophysiology/ Evoked Potentials*.

Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*.

Schouten, B., Gerrits, E., & van Hessen, A. (2003). The end of categorical perception as we know it. *Speech Communication*, *41*(1), 71–80.

Schrager, E. (1998). APPLICATION OF ERP MEASURES IN PSYCHOLOGICAL RESEARCH Measurement and interpretation of the mismatch negativity. *Behavior Research Methods, Instruments, & Computers*.

Selinker, L. (1974). Interlanguage. *International Review of Applied Linguistics in Leanguage*

Teaching, *10*(1), 209–232.

Shestakova, A., Huotilainen, M., Ceponiene, R., & Cheour, M. (2003). Event-related potentials associated with second language learning in children. *Clinical Neurophysiology*, *114*(8), 1507–1512.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society, Inc*, *31*(1), 137–149.

Strange, W. (1995). Cross-langage studies of speech perception A historical review. *Speech Perception and Linguistic Experience : Issues in Cross-Language Research*.

Strange, W., & Jenkins, J. J. (1978). Role of Linguistic Experience in the Perception of Speech. In *Perception and Experience*.

Sussman, E., Winkler, I., & Wang, W. (2003). MMN and attention: Competition for deviance detection. *Psychophysiology*.

Tavabi, K., Elling, L., Dobel, C., Pantev, C., & Zwitserlood, P. (2009). Effects of place of articulation changes on auditory neural activity: A magnetoencephalography study. *PLoS ONE*.

Tervaniemi, M., Medvedev, S. V., Alho, K., Pakhomov, S. V., Roudas, M. S., Van Zuijen, T. L., & Näätänen, R. (2000). Lateralized automatic auditory processing of phonetic versus musical information: A PET study. *Human Brain Mapping*.

The Mathworks Inc. (2018). MATLAB 2018a.

Tiitinen, H., May, P., Reinikainen, K., & Näätänen, R. (1994). Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature*.

Waltmunson, J. C. (2005). *The relative degree of difficulty of L2 Spanish / d , t /, trill , and tap by L1 English speakers: Auditory and acoustic methods of defining pronunciation accuracy*. Doctoral dissertation, University of Washington.

Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., … Näätänen, R. (1999). Brain responses reveal the learning of foreign language phonemes. *Psychophysiology*, *36*(5), 638–642.

Winkler, I., Lehtokoski, A., Alku, P., Vainio, M., Czigler, I., Aaltonen, O., Naatanen, R. (1999). Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations. *Cognitive Brain Research*, *7*, 357–369.

Yeni-Komshian, G. H., Flege, J. E., & Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean–English bilinguals. *Bilingualism: Language and Cognition*.

Zampini, M. L. (1994). The Role of Native Language Transfer and Task Formality in the Acquisition of Spanish Spirantization. *Hispania*, *77*(3), 470.

Zanow, F., & Knösche, T. R. (2004). ASA - Advanced Source Analysis of continuous and event-related EEG/MEG signals. *Brain Topography*.

Zue, V. W., & Laferriere, M. (2005). Acoustic study of medial /t,d/ in American English. *The Journal of the Acoustical Society of America*, *66*(4), 1039–1050.