Nonnative Phonetic Perception in Adult L2 Learners

by

Alba Tuninetti

Bachelor of Arts, University of Pennsylvania, 2009

Master of Science, University of Pittsburgh, 2012

Submitted to the Graduate Faculty of the

Kenneth P. Dietrich School of Arts & Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2015

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS & SCIENCES

This dissertation was presented

by

Alba Tuninetti

It was defended on

November 24, 2015

and approved by

Charles Perfetti, Distinguished Professor, Department of Psychology

Tessa Warren, Associate Professor, Department of Psychology

Marta Ortega-Llebaria, Assistant Professor, Department of Linguistics

Dissertation Advisor: Natasha Tokowicz, Associate Professor, Department of Psychology

Nonnative Phonetic Perception in Adult L2 Learners

Alba Tuninetti, PhD

University of Pittsburgh, 2015

Even with years of practice, adult learners have trouble perceiving and producing sounds in a second language (L2). Adults tend to need more focused and targeted input to achieve native-like perception and production of L2 sounds than children. The present study aims to clarify the mechanisms through which L2 perception is influenced by first language (L1) sounds, the neural basis of this perception, how learner differences influence learning, and how different training paradigms modulate both the neural and behavioral basis of L2 sound perception. Native English and native Spanish speakers participated in a five-day training paradigm during which they learned to discriminate Hindi sounds that do not belong to their L1 sound categories. Participants underwent electroencephalogram (EEG) recordings from the scalp, baseline discrimination tasks, training, and several memory and attention individual measures. We expected that the L1 would modulate the EEG waveform known as the mismatch negativity (MMN) at approximately 150-200ms after sound onset. This measure indexes early phonetic learning and previous research has shown that the waveform's amplitude can change or shift with new phonetic learning, indicating a reorganization of early acoustic and phonetic processing with new input. Furthermore, we examined how the L1 and different training and feedback paradigms influence this MMN change. Results demonstrate that both learner groups showed a modulation in the MMN waveform after training, but the change was eclipsed by the native contrast that was tested as a control, depending on how well they performed during training. Furthermore, participants in the feedback condition performed better on the training than those in the no-feedback condition but

this was not related to the ERP results, suggesting that feedback may be useful for overt behavioral responses, but not necessary for pre-attentive neural responses. These results are examined in light of the Perceptual Assimilation Model (PAM; Best, 1991, 1995), the Speech Learning Model (SLM; Flege, 1995), the Native Language Magnet model (NLM; Kuhl & Riviera-Gaxiola, 2008), and the Unified Competition Model (UCM; MacWhinney, 2005), examining similarity between L1s, neural hardwiring in the brain, and competition between phonetic contrasts.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

There are innumerable people who have contributed to make this dissertation, and all the work that came before it, possible. To these people, I offer my deepest gratitude because I would not be here without you:

First and foremost, to my parents, who left their home to pursue their own educational goals and, in doing so, gave my siblings and I the opportunity to do so as well. Without their constant support, guidance, and love, I would not be here.

To my advisor, Natasha Tokowicz, who has supported me and guided me throughout my entire graduate school career, allowing me to explore new methodologies and areas of research, and giving me the confidence to wholeheartedly jump into research with both feet.

To my wonderful husband, Martin Doppelt, without whom this (graduate school and a dissertation) never would have been completed; thank you for putting up with me, cooking for me, make sure I stayed sane, and letting me know that it was all going to be ok, no matter what. You made this worth it.

To my partner in crime, Chelsea Eddington; we've been attached at the hip since we started together and finishing together is the only appropriate way to end this. Thank you for the laughing, crying, commiserating, coffee runs, work dates, stats dates, computer help, hugs,

family time, and everything else in between; it made it all better to know that you were right there with me.

To my committee members, Drs. Charles Perfetti, Tessa Warren, and Marta Ortega-Llebaria, without whom this work would not be what it is today; countless hours of committees, suggestions, debates all helped to make sure this work achieved its full potential. Thank you! Special thanks to Marta Ortega-Llebaria who very patiently taught me how to manipulate speech tokens in Praat and made sure the stimuli were created correctly.

To the PLUM and RLG labs, in particular Katherine Martin, Evelyn Milburn, Caitlin Rice, Tamar Degani, Leida Tolentino, Julie Fiez, Adeetee Bhide, Gina Calloway, Xiaoping Fang, Liz Hirshorn, and many more; and to all the research assistants who helped collect data with me over the past year to make this possible. Your support, feedback, and guidance throughout this have helped me immeasurably.

To Sudhir Pathak, who very patiently and kindly recorded stimuli in Hindi for me.

To my funding sources, the Language Learning Dissertation Grant and the National Science Foundation Doctoral Dissertation Research Improvement Grant, for making this work possible.

And finally, to my participants, who participated in a long and sometimes tedious experiment for five days straight. Without you, I would have no dissertation; thank you!

1.0 INTRODUCTION

It is well documented that successful second language (L2) acquisition differs widely during childhood vs. adulthood; children seem to acquire an L2 seemingly without problems as long as they begin and continue the L2 acquisition fairly early in life (DeKeyser, 2000; Flege, Yeni-Komshian, & Liu, 1999; Johnson & Newport, 1989; Lenneberg, 1967; Newport, 1990). The most prominent or apparent advantages to beginning L2 acquisition from a young age appears in speech perception and production; children who have learned an L2 from a young age have excellent perception and production of the L2 phonetics, accent, prosody, etc., which seem to diminish as the age of L2 acquisition increases (Oyama, 1975; Pallier, Bosch, & Sebastian-Gallés, 1997; Piske, MacKay, & Flege, 2001; Seliger, Krashen, & Ladefoged, 1975, Tahta, Wood, & Loewenthal, 1981).

Extensive research with infants and children has shown that language-specific phonetic perception is learned in the first six months of life, and children retain the ability to perceive phonetic differences relevant to all languages for variable amounts of time, depending on cognitive individual differences and environmental linguistic influences (see Werker & Tees, 1999, for a review of the infant literature). However, adult L2 speech acquisition is different than that of children's in crucial ways because, unlike children, adult learners need more time and intensive, focused input to achieve native-like speech production and perception in an L2 (Best & Tyler, 2007; Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1999).

The goal of the present study was to examine what factors influence speech perception in adult L2 learners using event-related potentials (ERPs) derived from EEG recordings to examine how nonnative contrasts are perceived in naïve listeners over the course of a training paradigm; importantly, this study addressed how cross-linguistic similarity across two different L1 speaker groups influenced perception and acquisition of the same nonnative phonetic contrast. Furthermore, the effects of feedback and training type were manipulated to examine how behavioral input would influence neural responses to the same stimuli. With this study, we hope to clarify how the L1 can help in perceiving nonnative contrasts through transfer or competition between the L1 and the nonnative language, as well as how feedback and training can influence both the behavioral and neural responses as participants learn to perceive nonnative sounds.

It is well established that adult speakers who have little experience with an L2 have difficulty discriminating and categorizing nonnative phonetic contrasts (e.g., Best, McRoberts, & Goodell, 2001; Flanagan, 1972; Kuhl et al., 1992; Werker & Logan, 1985; Werker & Tees, 1984; see Best & Tyler, 2007, for a comprehensive review). However, the relative level of performance also varies with the specific contrast (e.g., Best & Strange, 1992; Best et al. 2003) and with different L1s (e.g., Best & Strange, 1992; Best, Traill, Carter, Harrison & Faber, 2003; Flege, 1989). Specifically, if a speech sound does not occur in the native language contrastively with another speech sound, listeners have trouble perceiving it as a separate phonetic category. For example, native Japanese speakers have trouble distinguishing the English liquids /r/ and /l/ because these sounds do not occur contrastively as phonemes in their native language; that is, they do not occur in Japanese as two different linguistically relevant sounds, as they do in English. These language-specific speech contrasts hinder perception and production during learning because the perceptual categorization of nonnative listeners is not organized to distinguish between these two phonemes as meaningful linguistic contrasts.

One of the main factors influencing how people perceive and acquire these contrasts is similarity to the first language, as evidenced by extensive behavioral data and models (Best & Tyler, 2007; Flege, 1995). However, the level at which this occurs, acoustic or perceptual, has yet to be fully detailed because behavioral results cannot provide as fine-grained measures of sensitivity from overt responses. To address this lack of sensitivity, work examining the neural bases of speech perception and production has attempted to clarify how the brain uses incoming information to categorize and identify novel phonetic contrasts.

Current research has examined data from neuroimaging methodologies ranging from functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), and electroencephalography (EEG) in an effort to capture the neural mechanisms of speech perception and acquisition. The present study built on previous work in this field to examine how the factors of cross-language similarity, training, and feedback affect both behavioral and neural responses to speech contrasts not present in the L1.

Participants in our study underwent two different training conditions with three feedback conditions to investigate the benefit of feedback during nonnative phonetic perception and acquisition. In general, previous work has demonstrated that performance feedback is beneficial during phonetic training paradigms, for both fixed and adaptive training (see McCandliss et al., 2002; Tricomi et al., 2006); here, the influence of feedback on both behavioral and neural responses was examined to further elucidate how feedback aids in speech perception and acquisition. Therefore, by combining two different language groups learning to perceive the same nonnative contrast with different training paradigms and feedback, this study addresses how perception of nonnative speech contrasts can be influenced by overt behavioral tasks, both at the behavioral level and the neural level with the added factor of cross-linguistic similarity. The results were examined in light of four current models of L2 speech perception, detailed below.

1.1 MODELS OF L2 SPEECH PERCEPTION

Examining the underlying structures and mechanisms by which L2 speech perception occurs can provide more information regarding the development and the state of the speech perception system generally, as well as answer questions regarding its plasticity and adaptation to novel stimuli. Here, we will review four relevant models of L2 speech perception and acquisition: the Perceptual Assimilation Model (PAM; Best, 1991, 1995; Best & Tyler, 2007), the Speech Learning Model (SLM; Flege, 1995), the Native Language Magnet (NLM) model (Kuhl & Riveria-, 2008), and the Unified Competition Model (UCM; MacWhinney, 2005). The mechanisms underlying these models will be addressed with respect to the present study.

Earlier models, such as the Perceptual Assimilation Model (PAM; Best, 1991, 1995; Best & Tyler, 2007) and the Speech Learning Model (SLM; Flege, 1995), addressed nonnative speech perception using the available research at the time, which, due to methodological constraints, concentrated primarily on behavioral methodologies and linguistic theories. The PAM proposes that speakers use their L1 system to perceive L2 input, assimilating the L2 phones into the L1 phones that are the most similar to the L2. The SLM posits that how well a speaker will perceive and produce nonnative speech is a function of the amount of L2 input. The amount and quality of input is measured as the age of arrival of the speaker to a country and how often the speaker relies on or is exposed to the native or nonnative language system. For both models, similarity between the L1 and L2 is one of their main tenets, although the SLM also highlights experience and exposure.

The PAM proposes that L1 speakers use their L1 speech system to perceive L2 input, assimilating the L2 phones into the L1 phones that are the most similar to the L2. Nonnative speech is assimilated either as an exemplar of a native phonemic segment (categorized), as an exemplar of a nonnative language phoneme (uncategorized; that is, not in the native phoneme category), or, more rarely, as a nonlinguistic sound (non-assimilated; that is, not incorporated into a phonemic category). Within the categorized exemplars, a good or poor exemplar of a native phonological segment corresponds to better or worse "goodness of fit". To measure goodness of fit, researchers have listeners categorize L2 input as pertaining to a native category, and then rate them on how "well" they fit the prototypical L1 category (e.g., Schmidt, 1996). Thus far, evidence has demonstrated that when listeners rate two speech contrasts as more similar to an L1 category, those contrasts receive lower discrimination scores because they were rated as more similar (Best, 1990; Polka, 1995). It is important to note that this reasoning is circular – those contrasts that are rated as similar to an L1 are then harder to discriminate because of the fact that they are similar.

The SLM has been compared frequently to the PAM, but the SLM critically different in that it addresses not only perception, but production and learning as well. The SLM posits that speakers extract information from acoustic-phonetic information directly, and use the statistical regularities in this information to build up information for a speech sound repertoire, more in line with the general approach to speech perception. Secondly, how well a speaker will perceive and produce nonnative speech is a function of the age of arrival of the speaker to a country and how often the speaker relies on or is exposed to the native or nonnative language system. A study examining native Italian speakers who spoke English as an L2 showed that there is a linear relationship between the age at which the speaker learned English and the perceived degree of foreign accent, such that the earlier the speaker learned their L2, the more "native-like" they were rated (Flege, Munro, & MacKay, 1995). Critically, the SLM states that the mechanisms used for learning a new speech system stay intact over the lifespan, such that new nonnative representations can form, but they depend strongly on the similarity to the already entrenched speech representations. Behavioral training studies have demonstrated that with enough exposure to a new speech contrast, subjects can gradually learn to correctly categorize nonnative contrasts that they were not able to categorize previously (Logan, Lively, & Pisoni, 1991; Strange & Dittmann, 1984). The PAM and the SLM focus primarily on adult nonnative speech perception and production, but with naïve listeners (the PAM) or with experienced listeners (the SLM).

The NLM model (Kuhl & Rivera-Gaxiola, 2008) and the Unified Competition Model (UCM; MacWhinney, 2005, 2012), began expanding on the ideas of competition, hardwiring, and neural encoding. The NLM (Kuhl, 1993; Kuhl & Rivera-Gaxiola, 2008) focuses primarily on nonnative speech perception during childhood - that is, how assimilation of a nonnative speech system occurs during early exposure to more than one language - and addresses how this assimilation develops over the lifespan as well. The perceptual space that is occupied by the language system can be reshaped with early language experience such that the representational structure changes as a function of the statistical distribution of the input. Native language-specific perception in infants begins around 6 months of age; by 11 months, there is a decline in foreign language phonetic perception (Kuhl & Rivera-Gaxiola, 2008). Infants’ neural networks are still malleable enough to change with early exposure to a new speech system, but these stabilize over time such that adults' networks are relatively unaffected by change and exposure to new stimuli. Given this, adult learners will have more difficulty achieving high proficiency when presented with L2 speech contrasts that require discrimination because their increased experience with their L1 speech system will block new phonetic learning and their perceptual space has been neurally committed to the L1 system.

Finally, the UCM (MacWhinney, 2005, 2012) is a data-driven connectionist model that uses cues as a basis for language comprehension. For L2 learners, the cues that adults use to process their L2 are obtained from their L1; this means that successful L2 processing depends on whether the L1 and the L2 are similar enough that the same cues can be applied from the L1 to the L2. These cues, which are present in the language and which listeners use inherently, form the crux of the UCM—they vary in strength based on their availability (how often they are present) and reliability (how often they lead to a successful interpretation). The UCM has mainly concentrated on learning grammar and syntax with this framework of cues and similarity, but it also touches on learning the phonetics and phonology of an L2. Cue strength can be objectively quantified from corpus data, and the UCM has used corpus data to show support for cue strength and similarity hypotheses (e.g., MacWhinney & Pleh, 1988). Huge corpora of language can be mined for different cues in phonetics and phonology and applied to learning an L2.

The UCM measures similarity in a relatively more objective way rather than depending on participant ratings. The UCM posits that similarity between the L1 and L2 will result in positive transfer between items (such as morphosyntactic endings, syntactic constructions) because they can transfer directly; differences between the L1 and L2 result in negative transfer and cause competition between the L1 and L2; finally, the learning of unique constructions in the L2 depends on cue strength. With respect to speech perception and processing, reliable and available cues may come from the natural environment or from laboratory training experiments, as used here.

In brief, models of L2 speech perception have attempted to explain the basis for speech perception in terms of L1-L2 similarity, plasticity, and neural commitment. Later models, such as the UCM and NLM, combine theories and mechanisms from earlier models, such as the PAM and SLM, to address how similarity to the L1 can influence the neural responses or cue weightings when responding to an L2. The current study seeks to clarify underlying assumptions in models of L2 speech perception, such as the level at which perception occurs, how native speaker and learner groups show differential responses to the same stimuli, and how manipulation of these stimuli affect both behavioral and neural responses. Theoretically, the UCM, with the competition mechanism, and the NLM, with neural hardwiring as a major component, are the most relevant with respect to the current experiment. Prior phonetic categorizations can provide competition when attempting to learn new phonetic categories, but previous research has shown that it is possible to have native-like responses to nonnative phonetic categories and this may challenge the NLM's construct of neurally-committed perceptual space.

1.2 CURRENT STUDY

Behavioral measures of speech perception have examined how individuals categorize phonetic stimuli with overt responses in an effort to quantify how factors such as cross-language similarity influence categorical discrimination and identification. However, to clarify the perceptual level at which categorization occurs, researchers turned to neural measures of speech perception and phonetic processing in an attempt to understand how linguistic factors can influence acoustic and phonetic perception.

Specifically, the mismatch negativity (MMN) response has been examined as the neural signature indicating sensitivity to phonetic contrasts. The MMN is a negative-going response seen particularly at the frontocentral and central electrodes that indexes a change in frequent stimuli. For nonnative speech perception, the MMN captures pre-attentional perception of infrequent stimuli and is used to test whether subjects can perceive the difference between two stimuli that differ either acoustically or phonetically. It is obtained by subtracting the ERP response to a frequent, or standard, stimulus from the ERP response to an infrequent, or deviant, stimulus and occurs between 150-250 ms after the change in stimuli. The MMN response is helpful to study because it indexes an involuntary pre-attentive response to changes in incoming auditory stimuli (Näätänen, 2001). Therefore, it is essential to note that the MMN is a difference waveform yielded by comparing waveforms elicited by different incoming sensory stimuli. Importantly, the MMN can be elicited only after the auditory system has been habituated (or has been able to form a representation of a standard stimulus) before hearing the deviant stimulus. That is, the MMN indexes the change in stimuli because the standard forms a short-term memory trace due to its repetition; upon hearing the deviant stimulus, the change in input causes the perceptual system to recalibrate the response. The change from the standard to the deviant stimulus is responsible for the MMN response (Näätänen, 2001; Näätänen & Winkler, 1999; Sams, Alho, & Näätänen, 1984). Finally, the MMN is elicited independent of attentional processes, so behavioral tasks are not needed to detect this waveform. In fact, many studies that investigate MMNs during auditory processing use behavioral tasks to direct subjects’ attention away from the MMN-eliciting stimulus to dissociate attention-dependent ERP responses that may overlap the MMN, usually in the form of watching a silent movie while listening to the stimuli (for a review, see Näätänen, Paavilainen, Rinne, & Alho, 2007).

Because the MMN is sensitive to changes in stimuli (from standard to deviant), it lends itself to study how sensitive individuals are to either acoustic or phonetic changes in stimuli. Investigating the timecourse of neural responses to native and nonnative speech contrasts in both children and adults was facilitated by the use of the MMN response to quantify differences in perception (Mueller, 2005; see Näätänen, 2001, for a review of the MMN in speech perception). For example, if examining the perceptual categorization of voice onset times (VOTs) of stop consonants in English, a 50ms VOT stimulus and a 90ms VOT stimulus would both be categorized as /p/ because the perceptual category of /p/ in English ranges in VOT from 20ms VOT to 120ms VOT (Lisker & Abramson, 1964). If the 50ms VOT were the standard stimulus, the 90ms VOT would be considered the *within-category* deviant. A *between-category* deviant would have a VOT outside of the range of English VOTs. Crucially, the within-category and between-category deviants are controlled such that the physical acoustic difference between them and the standard is equal, so the response is solely based on perceptual phonetic differences and not acoustic differences. Conversely, the *difference* between 50ms VOT and 90ms VOT is 40ms. That same difference of 40ms would have to be applied to the *between-category* deviant, resulting in a VOT of 10ms (50ms – 40ms). Hence, the standard 50ms VOT would have a between-category deviant of 10ms VOT, which also fulfills the requirement of being outside the range of English VOTs for /p/. This is done to ensure that the perceptual categorization of the /p/ continuum is not due solely to noticing a physical difference in the stimuli, but to linguistically and perceptually relevant differences.

Dehaene-Lambertz (1997) showed that there is a larger-amplitude MMN in response to an acoustic change that corresponds to a phonemic boundary in subjects' native language. Subjects heard two contrasts: one native to their language (French /ba/-/da/) and one not native to their language (Hindi dental /da/- retroflex /Da/) and were instructed to press a button in response to a perceived change. Importantly, the acoustic distance between the two sets of contrasts was controlled to be the same. This was done to ensure that brain responses to the stimuli were elicited by the phonetic contrast and not the acoustic difference. There was a larger MMN response for the native contrasts, suggesting that subjects could perceive only the difference in stimuli for contrasts that were native to their L1, but not for nonnative contrasts. The authors concluded that native language speech characteristics are encoded in memory and are used to compare incoming stimuli to stored representations during the task (Dehaene-Lambertz, 1997). Importantly, the researchers did not control for attentional influence in this study by making participants make overt responses to changes they heard. The MMN can be influenced by attention to the stimuli, leading to a heightened response (Pisoni, 1973).

Näätänen and colleagues (1997) controlled for attentional influence by having particpants direct their attention away from the incoming auditory stimuli. They had native Finnish and native Estonian speakers listen to vowels that were prototypical in Estonian and Finnish with deviant vowels that existed in both languages (a between-category contrast) or a vowel that only existed in Estonian (no category for native Finnish speakers). The vowels only varied in the second formant (F2) and all other frequencies were kept constants. The participants showed MMN responses only when the deviant stimulus was a native language phoneme prototype. For example, native Finnish speakers only showed an MMN response when the deviant stimulus was a Finnish /ö/, but not when it was an Estonian /õ/ (the standard stimulus was a Finnish /e/), suggesting that the MMN response was language-specific (Näätänen et al., 1997; see also Peltola et al., 2003; Sharma & Dorman, 2000). The authors also concluded that language-specific memory traces of phonemic representations are activated during the tasks to judge incoming stimuli, much like Dehaene-Lambertz (1997; see also Cheour et al., 1998).

These canonical MMN studies showed the MMN response within a language group; across language groups, Sharma and Dorman (2000) had native English speakers discriminate between two stimuli that differed in VOT, /baar/ and /paar/, a speech contrast native to Hindi speakers. Their stimuli were produced by a native Hindi speaker and then edited at zero crossings to create stimuli with pre-voicing from 0 ms VOT to 90 ms VOT. Their native Hindi control group showed an MMN response to the contrasts but the native English speakers did not, demonstrating that the MMN is, in fact, language-specific. In late L2 learners, Winkler and colleagues (1999) showed that native Hungarian speakers who had been living in Finland and were fluent in Finnish (learned after age 13) had comparable MMN responses to native Finnish speakers to two Finnish speech contrasts. Native Hungarian speakers with no prior experience in Finnish did not show this neural response. These results suggest that learning a foreign language can generate long-term changes in the brain, such that the perceptual system of L2 learners shows the same sensitivity to L2 speech contrasts as native speakers.

However, there may still be differences between proficient L2 learners and native speakers in the quality of their neural responses. Nenonen et al. (2003, 2005) showed that L2 speakers did not have the same “fine-tuning” as native speakers because their L2 speaker group showed an MMN with smaller amplitude compared to the native speakers for a phonetic contrast. They had native Russian and native Finnish participants listen to vowel duration differences, keeping all other acoustic measures (formant frequency, pitch) constant across deviant and standard stimuli. Standard stimuli were presented with a 200 ms vowel duration and deviants were presented with a 150 ms vowel duration. Cross-linguistic influence was controlled in a later study that examined how similarity between languages could influence the responses. Highly-proficient Russian-Finnish bilinguals showed a greater-amplitude MMN for a dissimilar speech contrast than a similar speech contrast (confirmed by norming) suggesting that similarity to the L1 and mediation through L1 speech systems could influence neural brain responses to L2 contrasts. However, this study was done with children; examining this effect with late or naïve L2 learners could elucidate on the neural mechanisms behind perceiving and learning new nonnative phonetic contrasts.

More recently, Brandmeyer et al. (2012) showed that highly proficient adult L2 speakers (Dutch-English) still show L1 influence in perceiving differences between L1 and L2 speech contrasts when presented with a /pa/ and /ba/ VOT contrasts. The authors manipulated one token of an aspirated /pa/ to create additional stimuli by removing 11 ms sections of aspiration to create a continuum that sounded like /b/-/p/ for native English speakers. Behaviorally, native English speakers show an earlier shift in the category boundary than L2 speakers, suggesting that the L2 influenced the perception of the boundary, and neurally, larger amplitude MMNs corresponded to the respective perceived category boundary of each language group, such that native speakers had larger MMNs at an earlier category boundary than L2 speakers. This suggests that native language experience affects how acoustic features of speech are weighted within a language to provide salient information within the perceptual speech representations of the native language.

Training studies have used the MMN to examine how brain responses change from pre- to post-exposure to nonnative phonetic contrasts. Kraus et al. (1995) trained adults to discriminate between synthesized variants of /da/ varying in the formant frequencies for F2 and

F3, but keeping all other acoustic features equal (duration, amplitude, fundamental frequency, steady-state vowel, and consonant portion). They found that the MMN showed longer duration and larger magnitude after training than before training, coinciding with an improvement in behavioral performance for most subjects. Furthermore, Tremblay and colleagues (1997) also showed that behavioral improvement in speech contrast discrimination transferred to novel nonnative speech discrimination tasks. Participants were tested on synthesized labial (/ba/-/pa/) and alveolar (/da/-/ta/) speech sounds in a continuum from -50 ms VOT to 50 ms VOT in 10 ms steps. These contrasts constituted a difference only in place of articulation; participants were trained on the labial contrast and were tested on discrimination for the alveolar contrasts. Participants showed showed enhanced MMN responses for novel stimuli (Kraus, McGee, Carrell, King, Tremblay, & Nicol, 1995; Tremblay, Kraus, Carrell, & McGee, 1997; see also Tremblay, Kraus, & McGee, 1998). Studies with non-speech stimuli have shown that the same type of MMN responses can be elicited after extensive training, suggesting that the formation of meaningful auditory categories is not a phenomenon special to human language, but can be done through exposure and statistical regularities (Liu & Holt, 2011).

To date, studies examining behavioral training and MMN ERP responses have failed to clarify how the L1 influences this perceptual adaptation and learning and how training and feedback could interact with L1 similarity to play a role in the behavioral and neural changes during learning. Specifically, the present study investigated how L1 speech categories influence acquisition of L2 speech categories with behavioral discrimination accuracy judgments and pre-attentive neural responses (i.e., the MMN). Native English and native Spanish speakers were tested on their discrimination of Hindi pre-voiced contrasts that do not exist as contrastive linguistic distinctions in English or Spanish. Furthermore, both native speaker groups learned the same pairs of contrasts, but those contrasts are categorized differently in English and Spanish. This manipulation will help elucidate the influence of L1 similarity on perceiving and learning new phonetic categories in an unknown language. Previous work has shown that similarity between the L1 and L2 can influence MMN responses in children and adults, such that more dissimilar sounds show a larger amplitude MMN response, and more similar sounds show a smaller MMN response (Dehaene-Lambertz, 1997; Díaz, Baus, Escera, Costa, & Sebastián-Gallés, 2008; Näätänen et al., 1997; Nenonen et al., 2005; Sharma & Dorman, 2000). However, most studies have only looked at two languages in contrast with each other (i.e., one language group learning to perceive an L2 or nonnative contrast), and have not examined behavioral and neural responses co-occurring in an experimental design with more than one language group learning to perceive similar nonnative contrasts.

Furthermore, our study also examines the effect of feedback on speech perception learning. Previous research has shown that consciously attending to auditory stimuli improves perceptual discrimination (Francis & Nusbaum, 2002), but explicitly drawing attention to the beginning of each sound (instead of allowing participants to do it on their own) may "speed up" the process of successfully discriminating between nonnative contrasts. Therefore, the present study examined if and how feedback influences the individuals' capacity to perceive and learn new nonnative contrasts with three feedback groups: no feedback, standard feedback, and attentional feedback. Standard feedback was defined as receiving a "correct" or "incorrect" response to their answer, whereas attentional feedback provided participants with information regarding the point at which to listen to the differences between the sounds (e.g., "Please pay attention to the *beginning* of the sound"). We hypothesized that participants in the attentional feedback condition would improve their discrimination earlier in training than those who received standard feedback because they would be better able to notice differences between the speech sounds (Schmidt, 1990, 1993).

Native English and native Spanish speakers were trained on the same Hindi phonetic contrasts: /b/-/p/ and /p/-/p^h/. Importantly, the first contrast, /b/-/p/, is native to both English and Spanish, and will act as a control to ensure that the MMN response is elicited by across-category phonetic stimuli. The contrast of interest is /p/-/p^h/ because these phonemes are categorized differently by native English and native Spanish speakers (see Figure 1 for categorization continuums for language groups). Native English speakers have a much wider category for the /p/ phoneme which encompasses both unaspirated and aspirated /p/ tokens. Although the aspiration is used systematically in English in certain phonemic contexts, it does not constitute a meaningful distinction from that of an unaspirated /p/, as it is in Hindi (i.e., they do not form a minimal pair distinction). For example, aspiration in /p/ regularly occurs at initial locations in English words and unaspirated /p/ occurs after /s/; an aspirated /p/ after /s/ cannot occur in English (Lisker, 1985). This predictable variation occurs in English, but because this variation can occur within the same semantic and lexical environment and not signal a minimal distinction (e.g., aspirated and unaspirated instances of /p/ within the same semantic and lexical context: 'ra/p^h/id' vs. 'ra/p/idity'), it is considered part of the same phonetic category. However, native Spanish speakers only have a very restricted range of aspiration when pronouncing /p/, from 0 ms VOT to approximately 15 ms VOT (Lisker & Abramson, 1964).

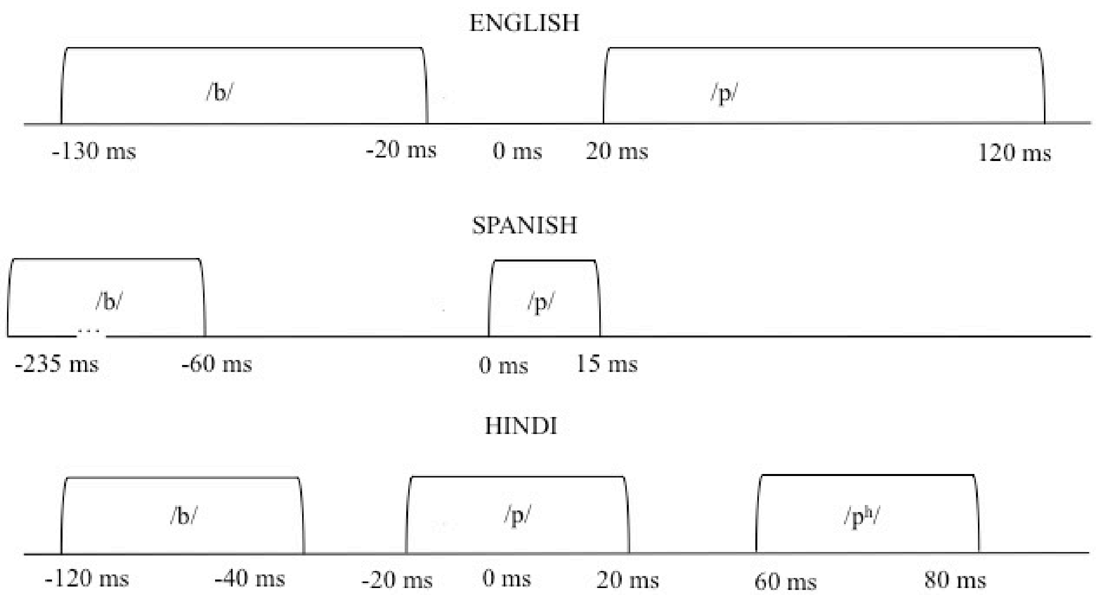

Figure 1. /b/ - /p/ continuum for three language groups.

Therefore, the current study investigated whether a prior existing representation, albeit organized differently, would make it easier (or harder) to learn a new categorization compared to having no categorization for the novel stimulus. Prior representations may make it easier to categorize the novel stimulus because no new representation needs to be formed within the phonemic repertoire; conversely, native English speakers may experience competition from the already-established /p^h/ that does not allow for re-categorization. Native Spanish speakers, who do not have an aspirated /p/ in their phonemic repertoire, may find it easy to form a new phonetic category for the novel stimulus and therefore show an earlier MMN response to the Hindi phonemes; conversely, it is unknown how perceptual representations of new phonetic stimuli are effectively formed so native Spanish speakers may not show the same benefit as native English speakers. The UCM predicts that similar constructions will engender positive transfer between languages, dissimilar constructions will cause competition, and learning of unique constructions depend on the cue validity; here, native English speakers should show positive transfer from the already-existing phonetic representation /p^h/ and be able to easily separate it from the larger /p/ category to make a new one. Alternatively, even though it already exists in English, it may provide competition because it is instantiated in the L1 differently than how it is expected in the L2. For this study, we hypothesize that competition between prior phonemic representations and the nonnative phonetic contrast would be a more powerful force than similarity between the speech sounds. Specifically, it would be more difficult for native English speakers to break apart an existing phonemic category into two phonetic categories. For native Spanish speakers, the aspirated /p^h/ would be a unique construction, such that the successful perception and acquisition of the phone depends on the type of cues available throughout training. In this case, because they were exposed to the phonetic contrasts everyday, the cue was available 100% of the time. However, due to the fact that they were in an English-speaking environment (i.e., the native Spanish speakers were living in the United States), the reliability of the cue was mixed: during training, it was reliably presented as separate from the unaspirated /p/, but in the environment outside the laboratory, they received input in which it was categorized within the larger English /p/ category.

The L1 of participants could influence the acquisition of nonnative contrasts in different ways depending on how the L1 phonetic categorization modulates the L2 speech contrast acquisition. Specifically, the present study could show a higher-amplitude MMN response to those phonetic speech pairs that are easier to learn because they would be instantiated more natively in the perceptual system and a correlation between behavioral responses and neural responses that could expand on the similarities and differences between the overt behavioral responses and underlying neural data; although it may be expected that a fairly close relationship would result, previous studies have shown that dissociation between behavioral and neural responses occurs in phonetic discrimination tasks (Kraus et al., 1995) and higher-level language tasks (e.g., Tokowicz & MacWhinney, 2005). This is the first training study, as far as we know, that has examined the statistical relationship between these two types of responses as a way to more fully understand the developmental timecourse of adult nonnative perception. This study also sought a clearer understanding of how nonnative phonetic perception develops over the course of adult L2 acquisition; three ERP sessions were done to test naïve perception, perception after half the training had been completed, and perception after the full training. Finally, how explicit metalinguistic feedback influences acquisition of nonnative speech contrasts and how effective this type of feedback can be in real-world applications was investigated.

The training paradigm consisted of an AXB phonetic discrimination task wherein participants are asked to choose which of two sounds are more similar to each other (decide if sound A is more similar to sound X or if sound B is more similar to sound X). Specifically, participants heard three sounds in a row and were asked to decide if the first sound they heard or the third sound they heard was more similar to the second sound they heard. For example, participants heard two very similar /p/ tokens (one with more aspiration, e.g., 10 ms VOT) and a fully aspirated /p/ (e.g., 80 ms VOT) and were asked to choose which of the /p/ sounds was most similar to the aspirated /p^h/ sound. Their response was correct if they answered that the /p/ token with *more* aspiration was more similar to the fully aspirated /p/, in order to shift their categorization to consider aspiration as a relevant feature for discrimination. The AXB task is used more often to ensure that listeners are using phonemic memory to correctly categorize the stimuli as opposed to being biased to answer same or different in an AX task (Gerrits & Schouten, 2005). The task forces listeners to compare each of two sounds to a third one, and although the three tokens are acoustically different, two belong to the same category. In this way, the AXB task examines the similarity between *categories* and not between acoustic stimuli. Importantly, this also addresses the problem of having variation within a phonemic category; that is, there is inherent variation in categories because it encompasses a range of acoustic dimensions that when perceived, all lead to perception of the same phonetic category.

Therefore, the acoustic differences *within* a category do not provide a *meaningful* difference between each one. When those acoustic differences cross a category boundary, and they become *between* category differences, they index a meaningful distinction. Therefore, when asked to discriminate between sounds, listeners are assumed to make a meaningful distinction; however, when examined with a simple AB task, it is unclear if listeners are making a meaningful distinction or just listening for acoustic differences. In this case, a task like the AXB task provides better evidence for the categorical discrimination. Out of the three sounds, two are in the same category though they may vary acoustically within the same category—therefore, listeners are making a responding based on categorical evidence, not acoustic. Here, we attempted to actively shift participants' categorization of the sounds to construct a new phonetic category of aspirated /p/ as separate from /b/ and unaspirated /p/, like native Hindi speakers.

Participants had exposure to two different oddball paradigms to test the reliability of the MMN response to more variation within a phonetic category. Two blocks of the oddball paradigm had the traditional MMN design such that there was one standard stimulus repeated with a within-category deviant and between-category deviant pseudo-randomly distributed throughout the block. In addition to these blocks, there were two blocks in which the standard stimulus was varied within the category; for example, instead of having the same /pha/ phone with 40 ms VOT repeated throughout the standard stimulus, the standard stimulus varied from 0 to 30 ms VOT to examine if the normal VOT variation captured in natural speech would elicit the same MMN response to within and between-category deviants. Varying standard presentation has been done with non-speech stimuli (sinusoidal tones), changing the frequency and intensity of the deviant stimuli; the same response was found to intensity and frequency changes although the amplitude of the MMN decreased with increasing variability (Winkler, Paavilainen, Alho, Reinikainen, Sams, & Naatanen, 1990). However, due to the natural variation inherent in speech (e.g., production of an aspirated /p/ is not the exactly the same from one instance to another), it would behoove us to know if the MMN can capture the phonetic status of the new phone within more variable presentation. This would add to evidence regarding how the MMN captures phonetic, and not acoustic, information about the stimuli if it can also be elicited within a more natural variable speech environment. To our knowledge, the current study was the first one to examine how natural speech variation within the standard presentation could elicit the same type of MMN response.

Specifically, the present experiment builds on existing literature by examining multiple language groups acquiring the same nonnative phonetic contrasts; these contrasts are instantiated differently in the L1 or do not exist in the L1, allowing us to examine how L1 similarity affects phonetic perception behaviorally and neurally. For the traditional oddball paradigm designed to elicit the MMN response, native English and native Spanish speakers were expected to show an MMN response to the between category contrast, /ba/-/pa/, but not to the second deviant pair, /pa/-/pha/, because it was a within-category change in English and because it does not exist as a native contrast in Spanish. By the end of training, they should show more comparable MMN responses to both contrasts if they have learned to perceive the /pa/-/pha/ difference as a meaningful distinction. For the manipulated oddball paradigm with varying standard presentation, we expected the same type of MMN response, but with a lower MMN amplitude for the /pa/-/pha/ change because there was more variability within the presentation, as seen in previous research varying the standard presentation (Winkler et al., 1990).

All participants who underwent training also completed various individual difference measures to ensure that variables exogenous to the study, such as cognitive differences between groups, would not account for differences in training performance or perceptual acquisition. Previous research has found that individual differences such as working memory and measures of non-spatial intelligence correlate with the ability to produce L2 speech more fluently, with fewer hesitations and pauses (Mota, 2003; Weissheimer & Mota, 2009) due to the enhanced ability of those with higher working memory to manipulate items and plan speech. Although we did not test speech production here, the effects of higher working memory might be seen in training performance, for example, which required participants to choose between three stimuli, keeping them in mind and actively comparing them. Jakoby, Goldstein, and Faust (2011) divided participants into two groups depending on well they performed on an L2 standardized test, to distinguish between "better" and "worse" learners; better learners had shorter MMN and P3a latencies, which the authors suggest may reflect a more efficient language learning mechanism. Ortiz-Mantilla, Choudhury, Alvarez, and Benasich (2012) showed that early exposure to another language may modulate acoustic abilities, as evidenced by performance on a nonspeech task, contributing to better language learning later in life. However, controlling for other types of individual differences has not yet been done, and it is possible that these may interact with perceptual learning.

Having two language groups learning the same nonnative phonetic contrast with behavioral and neural responses will better clarify the underlying bases of L1 cross-language influence for models of L2 speech perception, providing an additional measure of cross-language influence (MMN) that has yet to be incorporated into models of nonnative speech processing. Furthermore, we expected that differences between feedback types would inform theories and practices of foreign language teaching and learning in real-world classrooms and instruction, allowing for more efficient and effective language learning experiences among adults. Explicitly drawing attention to dissimilar constructions between the L1 and L2 has been shown to be helpful in learning grammatical constructions (Norris & Ortega, 2000; Tolentino & Tokowicz, 2014); this same type of explicit attention could aid in perceiving other types of L2 input, such as phonetics.

2.0 METHODS

2.1.1 Participants

Participants were 25 native English, 17 native Spanish, and 16 native Hindi speakers from the University of Pittsburgh, Carnegie Mellon University, and the surrounding community. All participants had normal or corrected vision, no self-reported hearing problems, and had no implanted brain devices. Native English and native Spanish speakers had no knowledge of Hindi or any other South Asian language.

Native English speakers and native Spanish speakers came in for five days of testing. Native English speakers were recruited through the University of Pittsburgh Introduction to Psychology subject pool, wherein undergraduate students are required to participate in experimental research for credit. All participants were paid $7/hour for tasks completed on the computer and $10/hour for time spent in the EEG setup, except when they were part of the University of Pittsburgh undergraduate psychology subject pool; those participants received 1 credit per hour for the first four hours of their participation, and were paid for additional time with the same rates as the other participants. Native Hindi speakers came in for one day of testing and were paid $7 per hour for behavioral tasks and $10 per hour for time spent in the EEG setup.

The data from two native Hindi speakers were not analyzed due to malfunctioning EEG caps during the testing and from another two due to being left-handed. Therefore, a total of 12 native Hindi speakers were analyzed as the control group. Three native English speakers did not complete the entire protocol; two were due to cap malfunctions and lack of EEG signal, one was due to drop-out after the first two days. They received credit for their time. Additionally, three native English speakers were left-handed; therefore, the data included a final set of 19 native English speakers. Two native Spanish speakers dropped out during testing and one was left-handed, for a total of 14 native Spanish participants.

Table 1. Participant LHQ Data.

	Native English speakers		Native Spanish speakers		Native Hindi speakers	
Gender	M	F	M	F	M	F
N	9	10	9	5	10	2
Age (years)	19.45 (2.36)		27.83 (4.09)		25.90 (5.52)	
Time in USA (months)	233.40 (28.32)		48 (64.20)		22.32 (66.36)	
	Self-rated proficiency					
L1 Reading	9.86 (0.35)		9.80 (0.54)		8.29 (2.40)	
L1 Writing	9.59 (0.66)		9.40 (1.02)		7.86 (2.60)	
L1 Speaking	9.73 (0.63)		9.87 (0.34)		9.54 (0.78)	
L1 Listening	9.82 (0.50)		10.00 (0)		9.71 (0.61)	
L2 Reading	4.65 (2.41)		9.33 (0.87)		8.57 (2.06)	
L2 Writing	4.15 (2.50)		8.80 (0.91)		8.64 (1.86)	
L2 Speaking	4.60 (2.16)		9.07 (0.99)		8.79 (0.89)	
L2 Listening	5.30 (2.45)		9.20 (0.75)		9.29 (0.61)	

2.1.2 Design

The study used a 2 training type (within-category training, between-category training) x 3 feedback type (no feedback, standard feedback, attentional feedback) x 3 testing day (pre-training, middle of training, post-training) mixed design.

2.1.3 Stimuli.

Stimuli consisted of the first syllable of a nonword conforming to Hindi phonotactics (confirmed by two native Hindi speakers): *paka*, *p^{h}aka*, *baka*, and *b^{h}aka*. A native Hindi speaker was recorded pronouncing the nonwords into a Marantz PMD670 solid-state recorder in an electrically-shielded and sound-attenuated booth (International Acoustics, Inc.). He recorded the nonwords with a lead-in sentence, such as “Please say *paka* now”, and as isolated words. All the words and sentences were spoken in Hindi. Analysis in Praat software (Boersma & Weenink, 2013) confirmed that there were no significant differences between the frequency, pitch, or amplitude of the words pronounced in isolation or in the sentences. Therefore, the isolated words were used for ease of manipulation. Each word was isolated in Praat and the first syllable of each word (*pa*, *p^{h}a*, *ba*, *b^{h}a*) was cut from the nonword. The syllable *p^{h}a* was selected to construct the range of syllables that participants would hear in both the training and the MMN protocol to ensure parity across the stimuli in terms of frequency, pitch, and volume. Furthermore, the syllable *p^{h}a* can be manipulated to cover the range of phonetic stimuli to which participants were exposed; that is, with full aspiration (20 to 80 ms VOT), the syllable is perceived as /p^{h}a/; with minimal aspiration (-10 to 10 ms VOT), it is perceived as /pa/ and with no aspiration and prevoicing (-20 to -80 ms VOT), it is perceived as /ba/ for native Hindi speakers. The /p^{h}a/ token had 10 ms of aspiration removed at zero crossings to form a new stimulus for a total of 10 exemplars of /p^{h}a/, ranging from 100 ms of aspiration to 0 ms of aspiration. Each subsequent removal of 10 ms of aspiration was balanced with an added 10 ms of silence to the beginning of the syllable to ensure that the length of each syllable would be comparable. To form the /ba/ syllables, voicing was added to the beginning of each subsequent stimulus in 10 ms sections, such that it began with the voiceless /pa/ sound and gradually increased to a reliable /ba/ sound with 100 ms of voicing onset. This was confirmed with norming done with native English and native Spanish speakers (different from those who participated in current experiment) in a simple identification task; they were asked to label a sound presented to them as beginning with a "p" or "b" sound (see Appendix A for details on the norming study). The native English and native Spanish speakers who participated in this training study also had the appropriate distributions for categorizing /p/ and /b/ according to their native language. This was confirmed with a second norming task done with a simple identification task; native English and native Spanish speakers in the current study were asked to label a sound presented to them as beginning with a "p" or "b" sound (see Figures 1-3 for continuum categorizations).

2.1.4 Procedure

Native English-speaking and native Spanish-speaking participants came in for five consecutive days to complete the experiment. Please refer to Table 2 for the order of tasks completed by each native language group.

Table 2. Summary and order of tasks for all language groups.

Language Group	Day 1	Day 2	Day 3	Day 4	Day 5

English	ERP ID Task LHQ	Training Day 1 Operation Span Task Stroop Task	ERP Training Day 2	Training Day 3 Flankers Raven's Matrices	ERP ID Task Payment
Spanish	ERP ID Task LHQ	Training Day 1 Operation Span Task Stroop Task	ERP Training Day 2	Training Day 3 Flankers Raven's Matrices	ERP ID Task Payment
	Day 1				
Hindi	ERP Identification (ID) Task Language History Questionnaire (LHQ) Payment				

On the first day, native English-speaking and native Spanish-speaking participants completed the first EEG session to capture their baseline response to the native Hindi sounds. They also completed an identification task, during which they simply listened to the speech sounds presented and responded if they believed the stimuli began with a /p/ or /b/ sound. These sounds were the same stimuli heard during the EEG tasks. They also filled out a language history questionnaire (Tokowicz, Michael, & Kroll, 2004) in an online survey format. On the second day, participants completed the first session of training, an operation span task (LaPointe & Engle, 1990) to measure working memory, and a Stroop task (Stroop, 1935) in their native language. On the third day, they completed the second EEG session and the second training session. On the fourth day, they did the last training session, the Raven's Matrices test (Raven, 1981), and the Flankers inhibitory control task (Eriksen & Eriksen, 1974). On their final day, they completed the final EEG session, the same identification task they completed on the first day, received payment, and were debriefed. Native Hindi speakers completed only a single EEG session, the identification task, and the language history questionnaire.

2.1.4.1 Training.

The training had two conditions; participants were trained either on discrimination between /p/ and /p^h/ in the AXB training task, or between /b/ and /p/ to see if training on different phonemes from the same continuum range would transfer to the /p/-/p^h/ contrast. Furthermore, the training conditions had three separate feedback conditions: no feedback, standard feedback, and attentional feedback. Native English and native Spanish speakers completed three days of training. Each day of training was successively more difficult than the previous day by manipulating the stimuli to be closer together phonetically. That is, on the first day, stimuli in the AXB task were selected from the extreme ends of each training continuum. Participants in Training condition A heard stimuli from the /p/-/p^h/ continuum (/p/ with 0 ms VOT to /p/ with 90 ms VOT) and participants in Training condition B heard stimuli from the /b/-/p/ continuum (/p/ with 0 ms VOT to /b/ with -90 ms VOT). The second day included stimuli selected from the middle of the two spectrums, and the third day included stimuli selected from the center of the spectrum, so the stimuli were the most phonetically similar. This allowed for gradual adaptation to the phonemes in the continuum as they became more difficult to perceptually discriminate. Previous research has shown that increasing difficulty of the training task during the training procedure leads to more gains in discrimination of speech sounds due to exaggerating and highlighting the relevant features of the sounds to be learned, in this case, aspiration (McCandliss et al., 2002; McClelland et al., 2002; but see Iverson et al., 2005 for contradictory evidence).

Participants heard three sounds in a row and were asked to judge whether the second sound was more similar to the first sound or to the third sound. They made their response using a response box. The three feedback conditions were implemented during the training: participants in the standard and attentional feedback conditions received a “correct” or “incorrect” assessment when they did the AXB task depending on their answer. The correct or incorrect statements were in line with training the participants to perceive the /p/ - /p^h/ distinction as meaningful. For example, when asked to determine which stimulus was more similar to a non-aspirated /pa/, /ba/ at 10 ms VOT or /pa/ at + 30 ms VOT, the correct answer is /ba/ at 10 ms VOT, even though it crosses the category boundary. This was manipulated to ensure that the aspirated /pa/ was considered a separate phoneme from the unaspirated /pa/ so that participants would learn to categorize aspirated /pa/ as a separate phoneme. Blocks had 60 triplets of phonemes and there were four blocks for a total of 240 trials per participant.

2.1.4.2 Operation span task.

The operation span task (LaPointe & Engle, 1990) is a standard measure of working memory; participants are required to hold words presented to them in memory while judging whether provided completions are correct or not, and the sets of operations/words range in size from two to six with three sets of each size. Participants were asked to decide whether simple mathematical equations (e.g., (9*1) – 2 = 7) were answered correctly or incorrectly and were then presented with a word to remember. Participants were asked to recall the words in the order in which they appeared as best as possible after each set with a prompt on the screen. They had two sets to practice the task during which the experimenter stayed with the participant and then they were left alone during the rest of the task.

2.1.4.3 Flankers test.

The Flankers test measures inhibitory control, or how well participants are able to inhibit conflicting stimuli to focus on the target (Eriksen & Eriksen, 1974). Participants were required to indicate the direction of a middle arrow (left or right) when flanked by congruent or incongruent arrows. Participants were shown five arrows; the middle arrow either pointed congruously with the flanking arrows (i.e., the middle arrow pointed in the same direction as the arrows on the side) or incongruously (i.e., the middle arrow pointed in the opposite direction of the flanking arrows). They had to respond if the middle arrow pointed to the left or the right as quickly and accurately as possible.

2.1.4.4 Raven's matrices.

Participants completed a portion of the Raven's test of nonverbal intelligence which examines how well people can deduce the missing part of a larger pattern, independently of language, reading, or writing skill (Raven, 1981). This task asks participants to complete a six by six matrix of patterned items by choosing the missing element of the pattern from one of eight choices. The task gradually gets more difficult as participants progress. Participants were shown 12 sets of the Raven's matrices test, after three practice sets. There was no time limit on how long they could take to complete each pattern.

2.1.4.5 MMN Protocol.

The EEG sessions followed the traditional oddball paradigm to elicit the MMN response. Participants listened to the blocks of stimuli over noise-canceling headphones in a steady stream of one stimulus every 900 ms. The volume was set at a comfortable level. Participants watched a movie of their choice on mute while hearing the sounds to keep their attention occupied, in keeping with previous studies of the MMN (Näätänen, 2001; Näätänen et al., 1997). Participants listened to five blocks of native Hindi syllables: /pa/, /ba/, or /p^{h}a/ in a semirandom presentation, such that there were always at least three standard stimuli between presentations of a deviant stimuli. The standard stimulus was always /pa/ and the deviant stimuli were always /ba/ and /p^{h}a/.

There were five blocks of stimuli in the MMN protocol. Blocks one and three were set up with one standard stimulus of /pa/ with two deviant stimuli of /pha/ and /ba/. Blocks two and five had varying standard stimuli to investigate the effect of categorization on speech perception. Therefore, the standard stimuli varied between each other in increments of 10 ms of aspiration each. The deviant stimuli were kept the same as blocks one and three. Each block had approximately 1000 trials; the standard and deviant stimuli were presented in a pseudo-random order such that at least three standard stimuli were presented between each deviant stimulus. Approximately 85% of each block was a standard stimulus; the remaining 15% was divided equally between the two deviants. There was a final deviant-stimulus alone block in which participants listened to 400 trials of one deviant and 400 trials of the second deviant to ensure that that the responses to the deviants by themselves are not contaminated by the physical differences between the standard and deviant stimuli.

EEG activity was recorded continuously during each EEG block at a sampling rate of 1000 Hz and amplified with Neuroscan SynAmps2 amplifiers with 24-bit analog-to-digital conversion (Compumedics Neuroscan, Inc.). Participants wore a 64-channel Ag/AgCl electrode cap; hanging electrodes were placed on the right and left mastoid bones for re-referencing during preprocessing, on the outer canthi of the left and right eyes to monitor horizontal eye movements, and above and below the left eye to monitor eye blinks. Impedances were kept below 5kOhms as much as possible. After the five blocks of the MMN protocol, the electrode caps were removed and participants completed the other tasks on the appropriate days. See Figure 2 for a schemata of the EEG MMN recording paradigm.

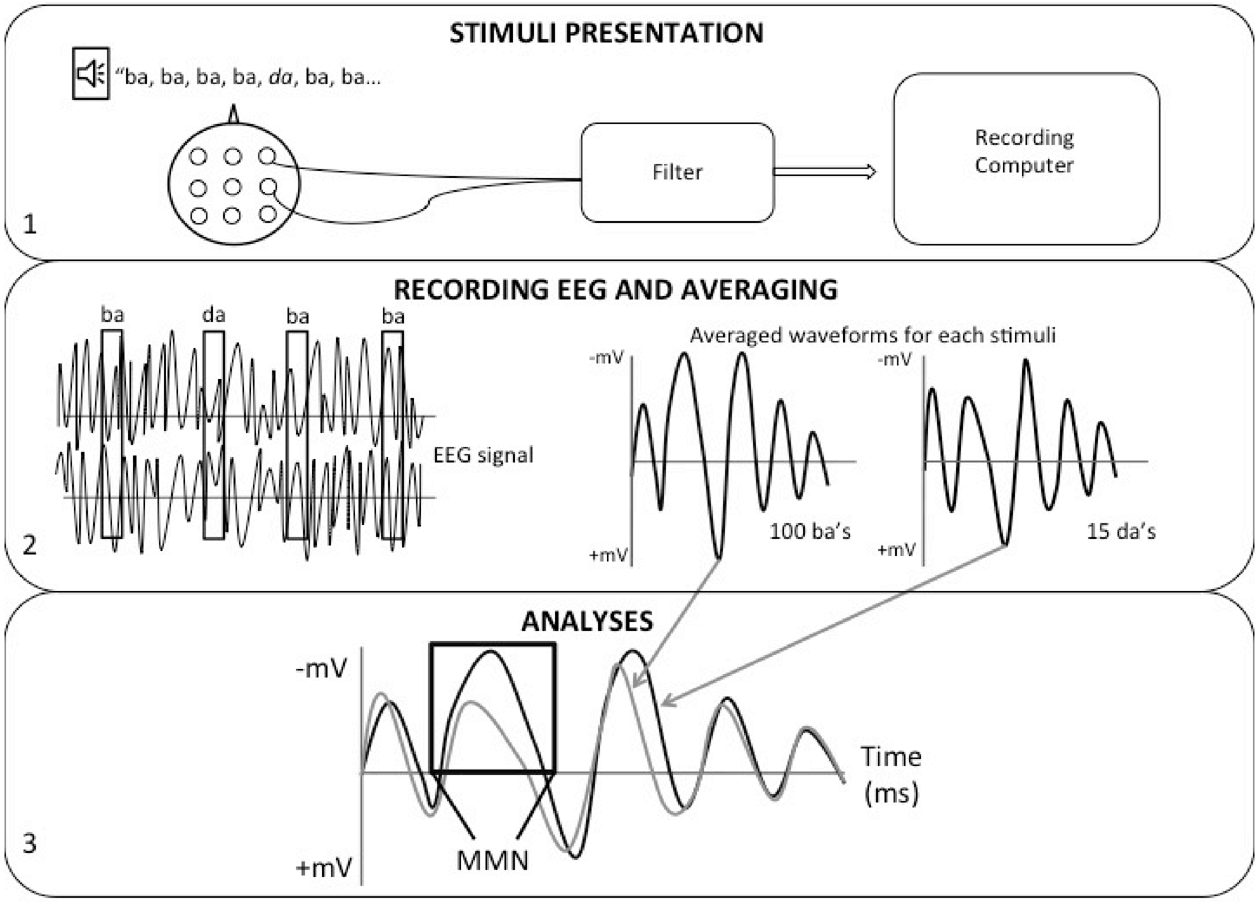

Figure 2. EEG and MMN recording paradigm.

2.1.4.6 EEG Preprocessing.

EEG data were processed off-line with Neuroscan Edit 4.3 software. All electrodes were re-referenced to the average of the right and left mastoids and low-pass filtered at 30 Hz with a slope of 24 dB per octave. Ocular artifact rejection was performed based on estimates of the average eye blink duration using standard algorithms (Neuroscan, Inc.). The ERP epoch was set from 100 ms pre-stimulus to 500 ms after the stimulus to encompass the MMN range of 150-250 ms after stimulus presentation and the possibility of finding a phonological mismatch negativity (PMN) from 250-350 ms. Channels that contained large artifacts were excluded from the averages; this corresponded to a maximum of two excluded channels in two participants and one electrode was excluded from all analyses due to large artifacts across all participants (FP1). Participants' data were included if they had a minimum of 50 trials in each deviant condition without artifacts; due to the repetition of stimulus blocks and a large number of trials, no participants' data were excluded on this basis.

3.0 RESULTS

3.1 BEHAVIORAL DATA

3.1.1 Training.

3.1.1.1 Native English speakers.

As stated above, native English speakers underwent three days of training with three different types of feedback: no feedback, standard feedback (correct or incorrect answer), or attentional feedback (correct or incorrect; if incorrect, they were given information regarding where to focus their attention to correctly distinguish between the phonemes: "Please pay attention to the beginning of the sound"). There were two training conditions, wherein one group received specific training on the /pa/-/pha/ continuum (A) and the other condition received training on the /ba/-/pa/ continuum (B). Training consisted of an AXB task in which participants heard three sounds (instances of the syllables) were asked to choose if the first sound or third sound they heard was more similar to the second. This was manipulated such that the correct answer shifted their categorization to align with the native Hindi speakers (e.g., less aspiration in the /pa/ to a /ba/ sound, indicating that /p^{h}a/ and /pa/ were actually not similar). Behavioral analyses were done to examine the effects of feedback on training accuracy. Only significant effects are reported and post-hoc tests are reported with Bonferroni corrections to correct for multiple comparisons. An ANOVA with training and feedback type as variables indicated that there was a significant effect of feedback type on overall accuracy, F (2,16) = 5.56, p = .02 η_p^2 = .41, such that participants with the standard feedback (M = 68%) and attentional feedback (M = 69%) conditions had more accurate performance than the participants in the no feedback condition (M = 55%) (standard vs. none: t (13) = 3.49, p = .004; attentional vs. none: t (14) = 2.79, p = .02). There was no difference between the standard and attentional feedback conditions, t < 1, and there were no differences between training conditions, t (20) = 1.00, p = .33. These results suggest that for native English speakers, receiving feedback of some sort was more helpful during the training to correctly identify similarity between phonetic contrasts than receiving no feedback.

Furthermore, there was a significant effect of feedback on day 1, F (2,16) = 7.02, p = .006, η_p^2 = .48, a marginal effect for day 2, F (2,16) = 3.14, p = .07, η_p^2 = .28, and no effect for day 3, F (2,16) = 1.07, p = .37, η_p^2 = .12, suggesting that the effect of feedback on total accuracy was driven by feedback type on the first and second days. Participants receiving attentional (M = 82%) and standard (M = 88%) were significantly more accurate on Day 1 compared to those receiving no feedback (M = 63%) (attentional vs. none: t (14) = 2.53, p = .02; standard vs. none: t (13) = 3.83, p = .002). The same pattern continued for Day 2: attentional (M = 63%) and standard (62%) conditions performed better than those receiving no feedback (M = 51%) (attentional vs. none: t (14) = 1.99, p = .07; standard vs. none: t (13) = 2.11, p = .06). This effect was no seen for Day 3: attentional (M = 59%), standard (M = 57%) did not have better accuracy than those in the no feedback condition (M = 52%) (all t's < 1). There were no significant differences between attentional and standard feedback conditions on any of the days (all t's < 1). These results suggest that both feedback conditions led to better accuracy during training but only for the first two days. As the difficulty of the training increased, the benefits seen by feedback decreased. See Figure 1 for mean accuracy for native English speakers across training and feedback type.

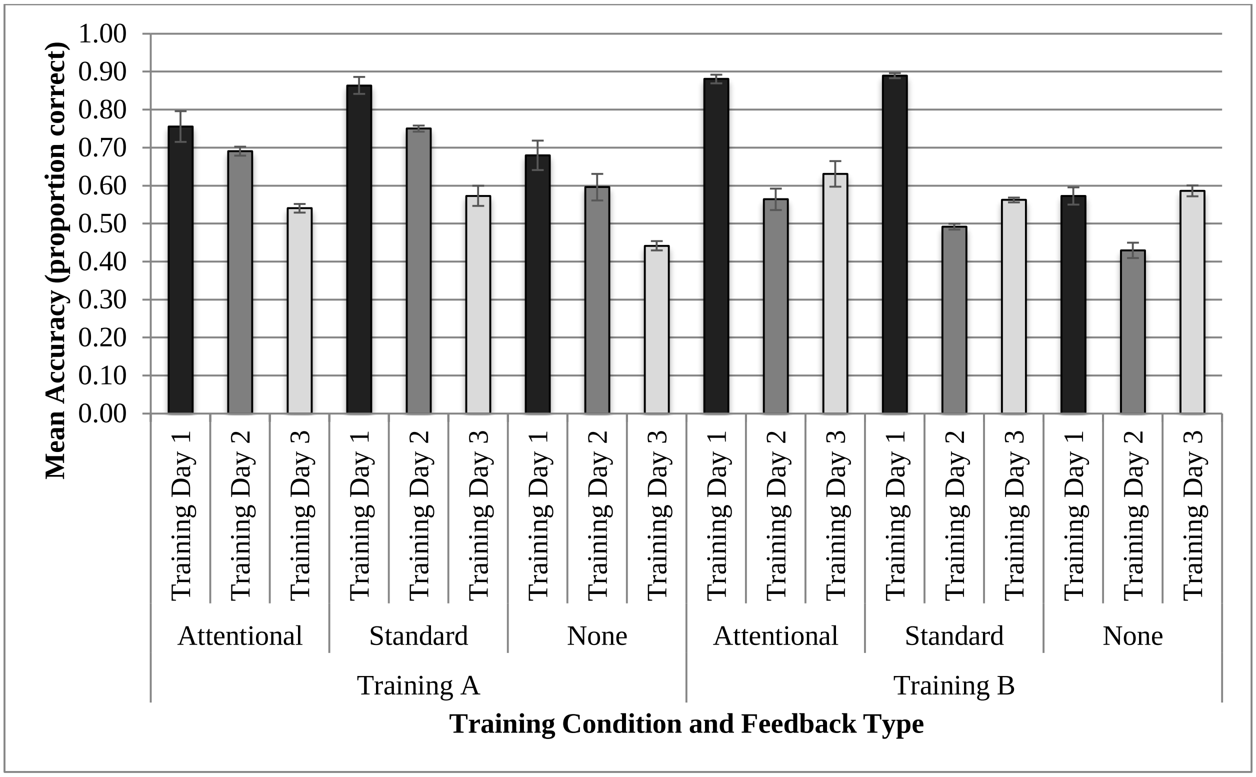

Figure 3. Native English speakers' training accuracy across condition and feedback type.

3.1.1.2 Native Spanish speakers.

Native Spanish speakers had the same three days of training as native English speakers. An ANOVA revealed a marginal effect of feedback type, F (2,9) = 2.93, p = .10, or training condition, $F < 1$, for native Spanish speakers. The marginal effect of feedback type follows the same pattern as native English speakers, such that the participants in the attentional (M = 69%) and standard (M = 72%) feedback conditions had better accuracy than participants in the no-feedback condition (M = 62%). Native Spanish speakers did not show an interaction for feedback on day, though there was a significant effect of training condition, such that the training on the

/b/-/p/ continuum led to better accuracy for the third day of training, F (1,9) = 8.16, p =.02, η_p^2 = .48. Participants in the /b/-/p/ training had a significantly higher accuracy (M = 65%) on Day 3 compared to participants in the /p/-/ph/ training (M = 50%), t (13) = 3.27, p = .006. See Figure 2 for mean accuracy across condition and feedback type.

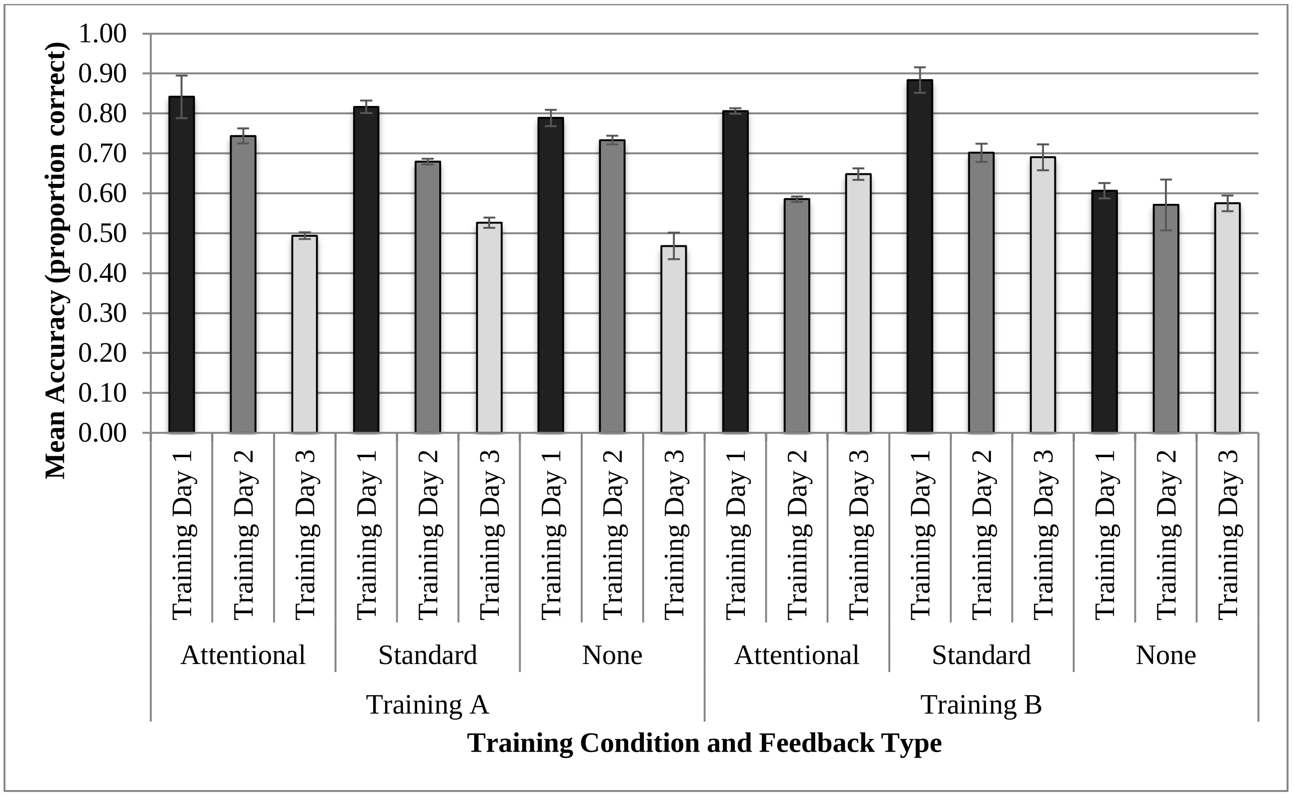

Figure 4. Native Spanish speakers' training accuracy across condition and feedback type.

3.1.2 Operation Span.

Operation span was analyzed with set size span which is the largest set size for which participants were able to get two out of three correct (range: 3 – 6 set size span). An ANOVA showed that there were no differences in set size span between native English (M = 5.09) and native Spanish (M = 4.93) speakers, F < 1.

3.1.3 Flanker's Test.

As a measure of inhibitory control, analyses for Flanker's were done using reaction times for correct trials only, such that the trial was only counted if they were correct in noting if the arrow pointed to the right or left. There were no significant differences between native English (M = 412 ms) and native Spanish speakers (M = 411 ms), $F < 1$.

3.1.4 Raven's Matrices.

Total accuracy on the Raven's matrices test (one point per problem if participants correctly answered out of a twelve possible total) showed significant differences between native English and native Spanish speakers, F (1, 35) = 7.71, p = .009, such that native Spanish speakers had significantly higher accuracy (M = 81%) than the native English speakers (M = 66%). Furthermore, native English speaking participants in the /ph/-/p/ training condition had higher Raven's accuracy (M = 72%) than those in the /b/-/p/ training condition (M = 56%), despite random assignment to conditions. No such difference was seen for the native Spanish speakers.

3.1.5 Identification Task.

All participants completed the identification task before training and after training; accuracy was measured by correct identification of the Hindi syllables. Figures 1-3 show the categorization continuums of each language group. Native Hindi speakers categorized from /ba/ + 20 ms of aspiration to the fully aspirated /pha/ as "p" sounds. Native English speakers categorized from /pa/ + 10 ms of aspiration onwards as "p" sounds both pre- and post-training. However, native

Spanish speakers showed a difference between pre- and post-training such that they showed a later categorization crossing pre-training, at /pa/ + 0 ms of aspiration, suggesting that they were able to move their categorization continuum with continued exposure to the /b/-/p/ training stimuli.

A post-hoc repeated measures MANOVA with testing day and VOT as within-subject factors revealed that there was a main effect of testing day for both the /b/, F (1,35) = 4.34, p =.04, η_p^2 = .11, and /p/ continuum, F (1,35) = 13.46, p = .001, η_p^2 = .28. There was no interaction with language and testing day, suggesting that both native English and native Spanish speakers showed a difference in how they categorized /p/ and /b/. There were main effects for VOT for both continuums as well; /b/, F (9,315) = 46.72, $p < .001$, η_p^2 = .57; /p/, F (9, 315) = 82.90, $p < .001$, η_p^2 = .70, suggesting a significant difference in where on the VOT continuum participants categorized b and p. This is not surprising, as both phones have different VOT measurements in Spanish and English. There was an interaction with language for /b/, F (9,315) = 2.73, p = .004, η_p^2 = .07, and /p/, F (9,315) = 7.84, $p < .001$, η_p^2 = .18, suggesting that native English and native Spanish speakers differed in where on the continuum they placed their /b/ and /p/ identification. Again, this is unsurprising because the native language groups use VOT differently for this range. There was also an interaction between testing day and VOT for both continuums, /b/, F (9,315) = 4.83, $p < .001$, η_p^2 = .12, and /p/, F (9,315) = 5.01, $p < .001$, η_p^2 = .13. This suggests that there was a difference in how both language groups categorized /b/ and /p/ from pre- to post-test, indicating that the training and exposure to the Hindi continuum shifted their representation.

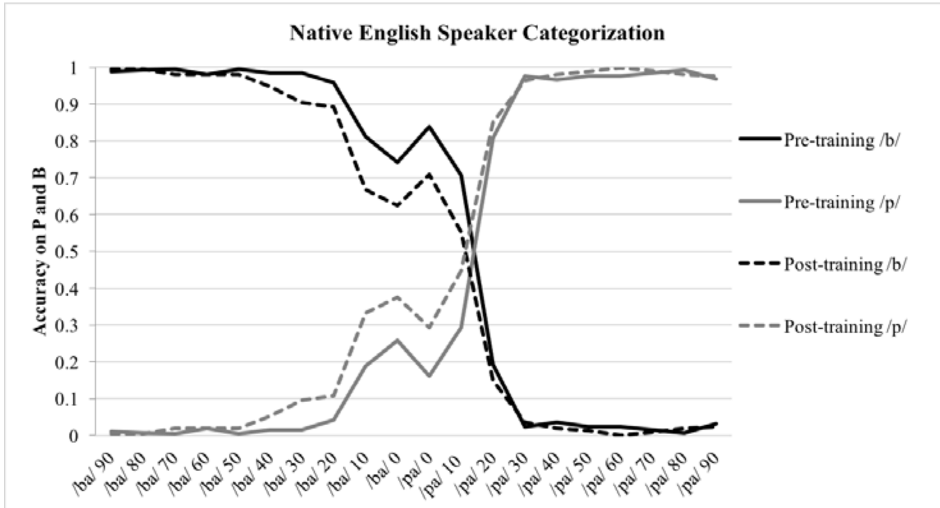

Figure 5. Native English speaker categorization.

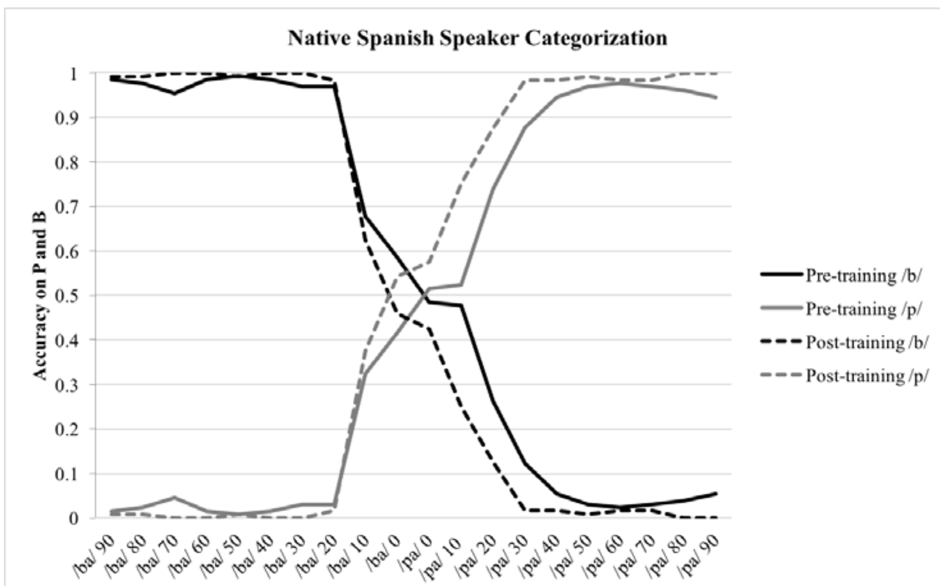

Figure 6. Native Spanish speaker categorization.

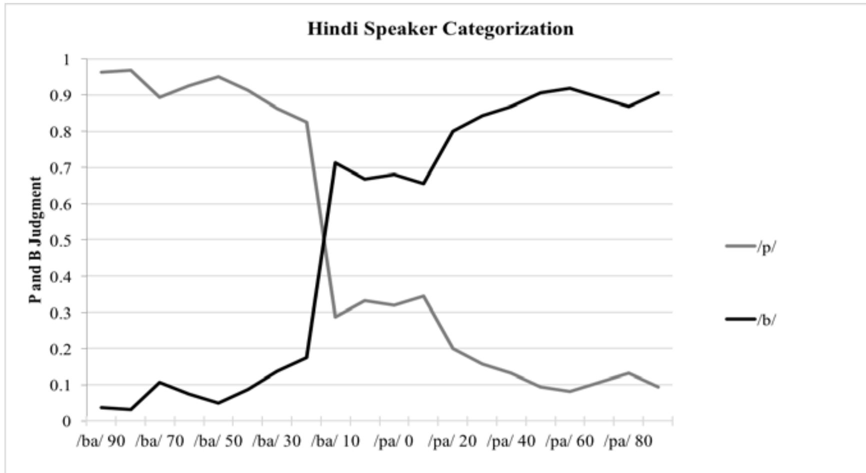

Figure 7. Native Hindi speaker categorization.

3.2 ERP DATA

Results from averaged ERPs in the 100 to 300 ms range are reported to test for a traditional MMN response. A repeated measures ANOVA included the within-subject factors of Difference (p^h-deviant (/p/ - /p^h/), b-deviant (/b/ - /p/)), Session (1, 2, or 3), Block (1 and 3; 2 and 4), laterality (left, midline, right), and electrode site (frontal, frontocentral, central) and the between subject factor of language (English, Spanish). Analyses were also conducted comparing native Hindi speakers to native English speakers and native Spanish speakers separately. Electrodes were selected to test for a fronto-central negativity, as per previous research (e.g., Brandmeyer, 2012; Nenonen et al., 2005); laterality was defined as left (electrodes F3, FC3, C3), midline (FZ, FCZ, CZ), and right (F4, FC4, C4). Lobes were defined within those electrodes tested: frontal (F), frontocentral (FC), and central (C). Blocks were combined for greater power and number of events per stimulus in the analyses; blocks 1 and 3 were combined and blocks 2 and 4 were combined, following the experimental setup of each one (i.e., blocks 1 and 3 were traditional oddball paradigm, blocks 2 and 4 were the manipulation of standard stimuli within the standard presentation). Results are reported within the 100-300 ms window separately for blocks 1 and 3 and blocks 2 and 4. Significant interactions were followed up using Duncan's multiple-range test to locate the source of the effect. Per convention, corrected *p*- and *F*-values are reported with Greenhouse-Geisser non-sphericity corrections; degrees of freedom are reported uncorrected (Keil et al., 2014). Effect size partial eta-squared measures are also reported.

3.2.1 Blocks 1 and 3 (traditional oddball paradigm).

An ANOVA revealed significant effects of difference type such that the b-deviant (/b-p/) difference was more negative than the p^h-deviant (/p^h-p/) difference for both native Spanish, F (1,12) = 10.40, p = .007, η_p^2 = .46, and native English speakers, F (1, 18) = 6.31, p = .02, η_p^2 = .26. This difference was not significant for native Hindi speakers, $F < 1$. This confirms our initial hypotheses that native Hindi speakers show comparable MMN responses for both standard-deviant pairs because they both exist as meaningful phonetic contrasts in their first language. However, native English and native Spanish speakers only have the Deviant 2 - Standard difference as a native contrast in their respective native languages and therefore show a larger MMN response to this contrast. Native Hindi speakers showed a significant main effect of block, F (1,11) = 4.89, p = .05, η_p^2 = .31, such that the first block showed more negative responses overall than the second block.

Native Spanish speakers also showed a main effect of block, F (1,12) = 11.28, p = .006, η_p^2 = .48, such that the second block of the traditional MMN paradigm showed more negative responses than the first block, an opposite effect of what was seen for native Hindi speakers. Furthermore, there was a main effect of lobe, F (2,24) = 4.20, p = .04, η_p^2 = .26, such that frontal electrode sites had more negative responses than central electrode sites. Native Spanish speakers also showed a marginal four-way interaction with difference type, session, block, and lobe, F (4,48) = 2.77, p = .07, η_p^2 = .19. Follow up tests indicated that the b-deviant difference was significantly more negative across all session except in the first block of Session 2. Furthermore, the p^h-deviant response was most positive in the first block of Session 1 compared to all other responses, excepting the second block in Session 2.

The results are in line with the hypothesis that the MMN response to the phonetic contrast already present in the native language would change very little over the course of the training; however, the response to the phonetic contrast *not* present in their native language changed, eliciting a more negative response after training, such that it was not significantly different in Sessions 2 and 3 from the native contrast negativity. This was confirmed with individual ANOVAs run between each session (Session 1 vs. Session 2, Session 1 vs. Session 3, and Session 2 vs. Session 3).

Native Spanish speakers only showed differences between Sessions 1 and 2, F (1,12) = 15.29, p = .002, η_p^2 = .56, and Sessions 1 and 3, F (1,12) = 8.62, p = .02, η_p^2 = .42, suggesting that native Spanish speakers showed the most difference between the MMN responses between pre- and post-training. Comparing between Sessions 1 and 3, there was an interaction between difference type and lobe, F (2,24) = 7.66, p = .006, η_p^2 = .39, which demonstrated that responses were most negative in Session 3 for the p^h-deviant difference across the frontal and frontocentral electrodes, confirming a classic MMN response to the nonnative contrast. This suggests that training did affect the morphology of the MMN response although not significantly compared to the b-deviant response (all p's < .05).

Native English speakers showed a significant interaction between difference type and lobe, F (2,36) = 4.14, p = .03, η_p^2 = .19, and between difference type and block, F (1,18) = 10.87, p = .004, η_p^2 = .38. Although follow-up tests did not show significant differences between means, the means for difference type and lobe suggest that the b-deviant difference was more negative over frontal, frontocentral, and central electrodes compared to the p^h-deviant difference; this reflects a standard MMN response. The interaction between difference type and block suggests that the b-deviant response was more negative in the second block within the traditional MMN paradigm compared to the first block. This could reflect the native contrast /b/-/p/ becoming more salient phonetically throughout training.

Analyses were done to compare differences between sessions within a language group to test the hypotheses of changes to the MMN responses as training progressed. Native English speakers showed differences between difference types in all sessions: Session 1 vs. Session 2, F (1, 18) = 3.79, p = .07, η_p^2 = .17, Session 1 vs. Session 3, F (1, 18) = 4.01, p =.06, η_p^2 = .18, Session 2 vs. Session 3, F (1, 18) = 7.66, p = .01, η_p^2 = .30. Native English speakers showed a change in MMN responses between standard-deviant pairs across all sessions although this was only marginally significant between Session 1 and Session 2 and between Session 1 and Session 3. The b-deviant response was more negative in all sessions compared to the p^h-deviant response.

For native English speakers, there was a significant interaction between difference type and block between Sessions 1 and 2, F (1,18) = 6.23, p = .02, η_p^2 = .26; follow-up tests did not reveal significant differences between means, but the effect appears to be driven by the second block being more negative for the b-deviant response compared to the second block of the p^h-deviant response. This same effect was found when comparing between Sessions 1 and 3, F

(1,18) = 7.66, p = .01, η_p^2 = .30 and Session 2 and 3, F (1,18) = 6.89, p = .02, η_p^2 = .28. This may be expected because this is the native contrast for native English speakers.

There was an interaction between difference type and lobe between Session 1 and Session 3, F (2,36) = 7.08, p = .006, η_p^2 = .28. Follow up tests did not reveal significant differences, but the b-deviant response appears to show a more negative response across midline and right electrode sites compared to the p^h-deviant response for native English speakers. Finally, there was an interaction between session, block, and laterality for Sessions 1 and 3 for native English speakers, F (2,36) = 4.61, p = .03, η_p^2 = .20. Post hocs did not show any significant differences between means; Session 2 appears to have more negative responses for the first block of the traditional MMN paradigm across midline and right electrodes compared to Session 1.

For native Spanish speakers, there was a main effect of lobe with Sessions 2 and 3, F (2,24) = 5.42, p = .01, η_p^2 = .31, such that frontal electrode sites were more negative than central electrode sites. This suggests that the MMN responses followed the typical frontal and frontocentral distribution. There was also an interaction between block, lobe, and laterality for the comparison between Sessions 2 and 3, F (4,48) = 3.21, p = .04, η_p^2 = .21; post-hoc tests did not significant differences between means. The interaction appears to be driven by more negative left and midline frontal electrodes for the second block, and more negative responses in general for the second block.

Comparing between Sessions 1 and 3 for native Spanish speakers, a main effect of block was found, F (1,12) = 6.93, p = .02, η_p^2 = .37, indicating the same pattern as the native English speakers, such that the second block in the traditional MMN paradigm showed a larger negativity compared to the first block. A four-way interaction between difference type, session, block, and lobe when comparing Sessions 1 and 3, F (2,24) = 7.66, p = .006, η_p^2 = .39, showed that the p^h-deviant response in the first block of Session 1 had less negative responses than the b-deviant response in both Sessions 1 and 3 (p's < .05). Figure 6 shows the mean amplitudes for native Hindi speakers for both the traditional and manipulated MMN paradigm. See Figures 7 and 8 for mean amplitudes across sessions and contrast types for both native language groups.

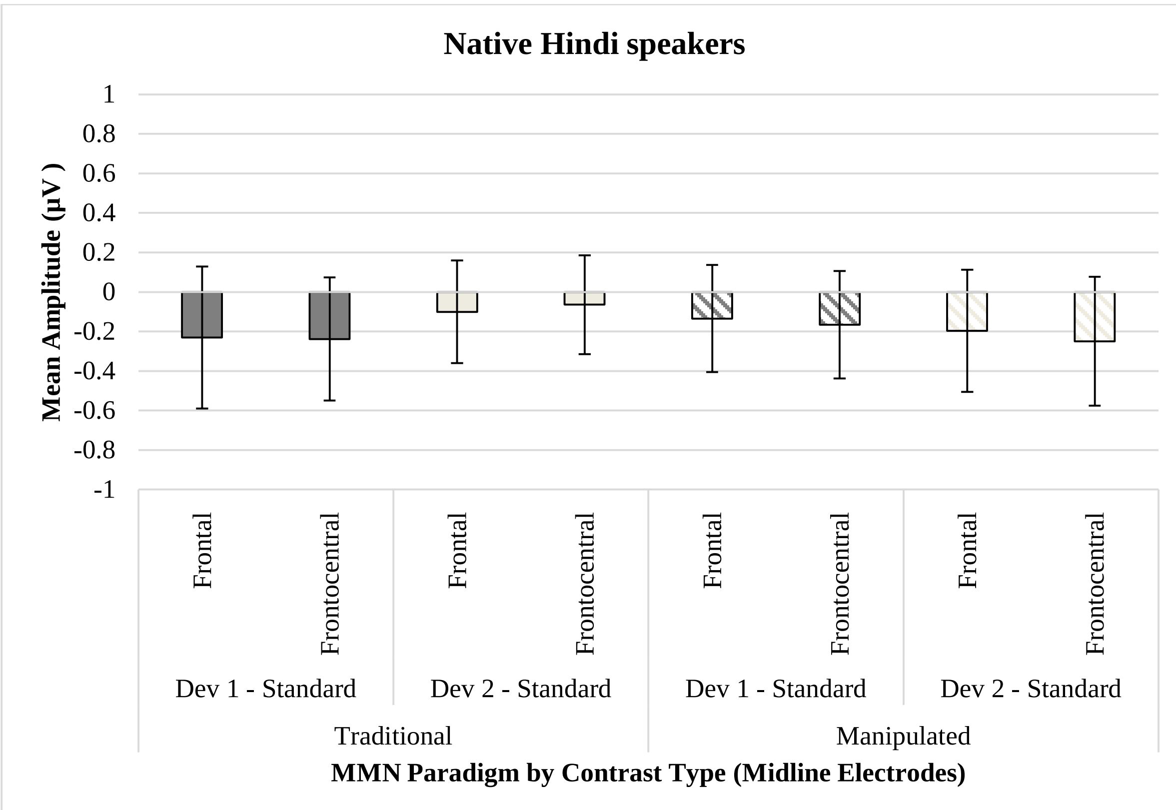

Figure 8. Native Hindi speakers, contrast type within traditional and manipulated MMN paradigm.

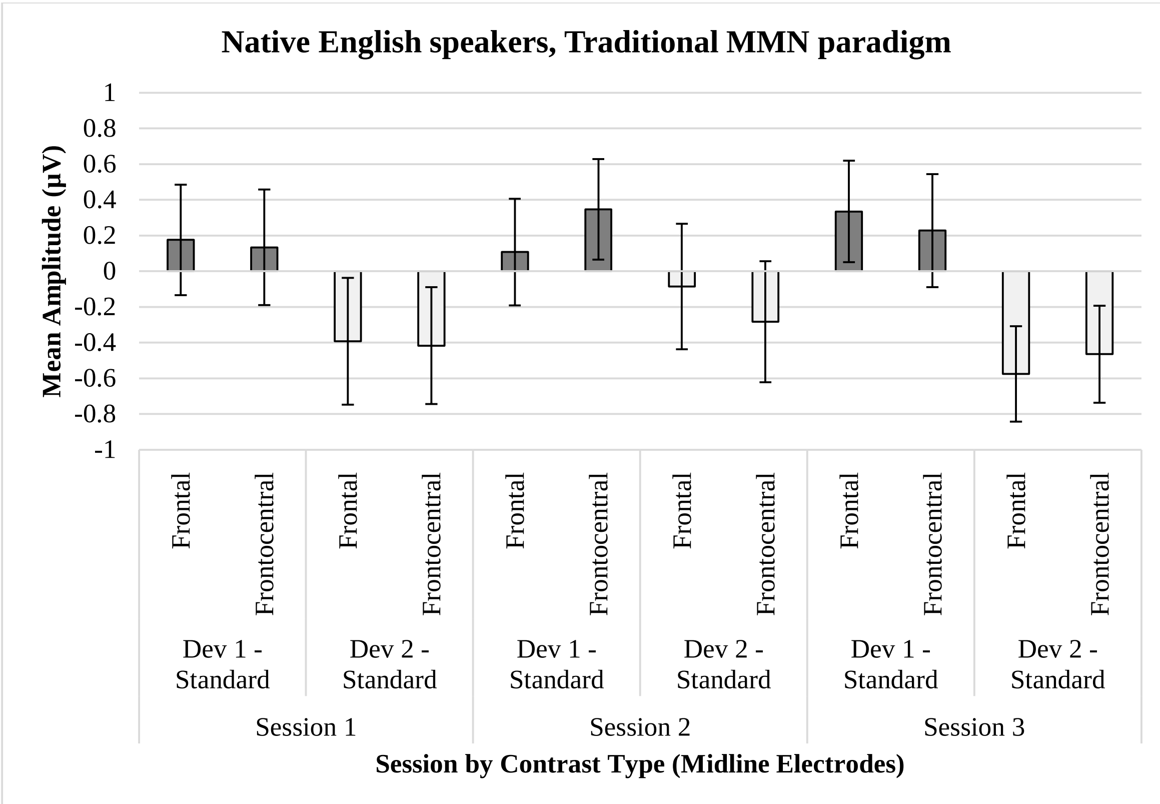

Figure 9. Native English speakers, session by contrast in traditional MMN paradigm.

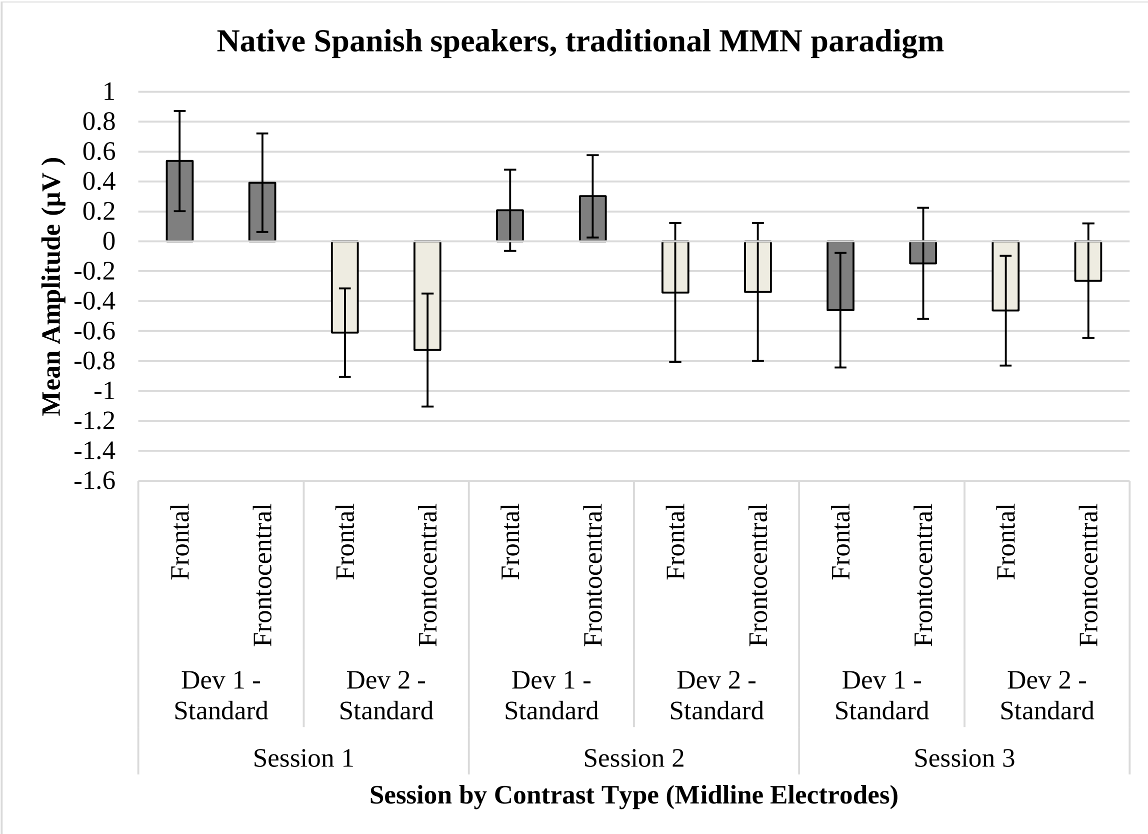

Figure 10. Native Spanish speakers, session by contrast in traditional MMN paradigm.

3.2.1.1 Hindi vs. English comparison.

Comparisons between Hindi and English speakers for both Sessions 1 and 3 were done to test native language differences pre- and post-training for native English speakers. When comparing Session 1, an ANOVA showed an interaction between difference type and laterality, F (2,58) = 3.99, p = .04, η_p^2 = .12; post-hoc tests showed no significant differences between the means but the interaction seems to show that the p^h-deviant response was more negative over the left electrodes and the b-deviant response was more negative over the right electrodes. There was also an interaction between difference type and laterality, F (2,58) = 4.00, p = .04, η_p^2 = .12. Follow up tests revealed no significant differences between means; visual inspection suggests that the p^h-deviant response showed more negativity over left electrode sites compared to more negative response over right electrode sites for the b-deviant response. Furthermore, an interaction between block and laterality, F (2,58) = 5.90, p = .01, η_p^2 = .17 suggests that there were more negative responses in the first block of the traditional MMN paradigm compared to the second block. This is in contrast to comparisons within native English speakers showing that the second block tends to be more negative; this effect may be driven by the native Hindi speakers, although there were no interactions with language. There was also an interaction between difference type, lobe, laterality, and language, F (4,116) = 3.29, p = .04, η_p^2 = .10, in Session 1. Post hoc analyses showed that the interaction was driven by the frontal right electrode (F4) in the p^h-deviant response being more positive than certain electrode sites within native English and native Hindi speakers (p < .05). These single electrode comparisons are not particularly important for our hypotheses; visual inspections of the means suggest that midline and right frontal and frontocentral electrodes are more negative for the p^h-deviant contrast in native Hindi speakers, but native English speakers had more negative responses for the b-deviant responses in general. For Session 3 comparing native Hindi and native English speakers, there was a main effect of block such that the first block of the oddball paradigm had more negative responses than the second block, F (1,29) = 6.48, p = .02, η_p^2 = .18.

3.2.1.2 Hindi vs. Spanish comparison.

Comparisons between Hindi and Spanish speakers in Session 1 showed a significant main effect of difference type, F (1,23) = 6.00, p = .02, η_p^2 = .21, such that the b-deviant response was more negative in general than the p^h-deviant response. There was also an interaction with language, F (1,23) = 9.04, p = .006, η_p^2 = .28; post-hoc tests showed no differences but the means suggested that there was a larger difference for the native Spanish speakers between the difference types than for the native Hindi speakers (as shown by the non-significant difference between difference types for the native Hindi speakers, $F < 1$). There was also a main effect of laterality, F (2,46) = 6.31, $p = .009$, $\eta_p^2 = .22$; midline and right electrode sites showed more negativity than left electrode sites. There was also a significant interaction between block and language, F (1,23) = 9.63, $p = .005$, $\eta_p^2 = .30$. Although follow up tests did not show significant differences between means, the means suggest that native Hindi speakers showed more negative responses in the first block, but native Spanish speakers showed more negative responses in the second block. In Session 3, there was an interaction with difference type, lobe, and laterality, F (4,92) = 3.52, $p = .04$, $\eta_p^2 = .13$; follow-up tests showed no differences between means, but the means suggest that native Hindi and native Spanish speakers had maximally negative responses over the midline and right electrodes for the p^h-deviant response.

3.2.2 Blocks 2 and 4 (manipulated oddball paradigm).

As with the traditional MMN paradigm, native Hindi speakers did not show a difference between responses to the p^h-deviant or b-deviant responses ($F < 1$). Native Spanish speakers showed a larger negative response for the b-deviant response than the p^h-deviant response, F (1,12) = 22.15, $p = .001$, $\eta_p^2 = .65$; however, this was only marginally significant for native English speakers, F (1,18) = 3.04, $p = .10$, $\eta_p^2 = .15$. For the native English speakers, there was a main effect of lobe, F (2,36) = 4.15, $p = .05$, $\eta_p^2 = .19$, such that the frontal electrode sites showed more negative responses than the central electrode sites. There was an interaction between block and laterality, F (2.36) = 5.59, $p = .01$, $\eta_p^2 = .24$; post-hoc tests did not reveal significant differences between means, but the interaction appears to be driven by more negative responses across left and midline electrodes for the second block within the manipulated MMN paradigm.

Comparing Sessions 1 and 3 for native English speakers showed an interaction between session and block, F (1,18) = 4.37, p = .05, η_p^2 = .20; means did not show significant differences in follow up tests, but the effect appears to be driven by the second block in Session 3 having a more negative response compared to the first block and to both blocks in Session 1. There was also an interaction between block and laterality when comparing between Sessions 1 and 3, F (2,36) = 3.49, p = .05, η_p^2 = .16; the second block appeared to show more negative responses overall and particularly over left electrode sites, although post-hoc tests did not show significant differences between the means.

There was also a significant interaction between session, lobe, and laterality for the comparison between Session 1 and Session 3, F (4,72) = 2.83, p = .05, η_p^2 = .14. Follow up tests indicated that frontal and frontocentral left and midline electrode sites showed more negative responses in Session 3 compared to central electrode sites in Session 1 (ps < .05). There was a main effect of block when comparing Sessions 2 and 3 for native English speakers, F (1,18) = 4.23, p = .05, η_p^2 = .19, such that the second block of the manipulated oddball paradigm showed more negative responses than the first block across both sessions. This suggests that the training led to more negative responses across sessions specifically for the second block.

For native Spanish speakers, there was a main effect of difference type when comparing Session 1 and Session 2, F (1,12) = 18.09, p = .001, η_p^2 = .60, such that participants had more negative responses for the b-deviant condition. This same pattern was seen when comparing Session 1 and Session 3, F (1,12) = 19.75, p = .001, η_p^2 = .62, such that the b-deviant response showed more negative responses in Session 3.

When comparing Session 1 and Session 2, there was an interaction between difference type and lobe, F (2,24) = 4.81, p = .03, η_p^2 = .29; follow up tests did not reveal any significant differences between means, but the results suggest that the b-deviant response was more negative on all lobes. There was an interaction between session, block, and lobe, F (2,24) = 4.72, p = .02, η_p^2 = .28; the means indicate that the frontal and frontocentral electrode sites showed more negative responses in Blocks 1 and 2 in Session 1, although these were not significant in post-hoc tests.

Native Spanish speakers also showed an interaction between session and laterality, F (2,24) = 4.46, p = .04, η_p^2 = .27. Follow up tests did not reveal a significant difference between means, but the means suggested that Session 1 showed more negativity across the right electrode sites compared to Session 2. An interaction between difference type, lobe, and laterality was found when comparing Sessions 1 and 2, F (4,48) = 4.63, p = .006, η_p^2 = .28; follow up tests showed that the b-deviant response was more negative over all electrode sites compared to the p^h-deviant response (all p's < .05). Furthermore, midline frontal and frontocentral electrode sites showed a larger difference for the p^h-deviant response compared to all other electrode sites regardless of session (ps < .05).

There was an interaction between difference type, lobe, and laterality when comparing Sessions 1 and 3 for native Spanish speakers, F (4,48) = 3.09, p = .04, η_p^2 = .21. Post hoc tests showed that the b-deviant response showed more negative responses across all electrode sites except for the central midline electrode (Cz) for the b-deviant response (all ps < .01).

Comparing Session 2 and 3 for native Spanish speakers showed a significant effect of laterality, F (2,24) = 4.88, p = .02, η_p^2 = .29, such that left electrode sites showed more negative responses than the midline and right electrode sites. There was an interaction with lobe and difference type, F (2,24) = 5.79, p = .02, η_p^2 = .33; post-hoc tests showed no significant differences between means, though they suggest that the b-deviant condition had more negative responses compared to the p^h-deviant response. There was an interaction with block and lobe as well, F (2,24) = 3.65, p = .05, η_p^2 = .23. Follow up tests revealed no significant differences between means; the means suggest that Block 2 was more negative over central electrode sites, and Block 1 had more negative responses was more negative over frontal and frontocentral electrode sites. See Figures 9 and 10 for mean amplitudes across session and contrast type for each language group.

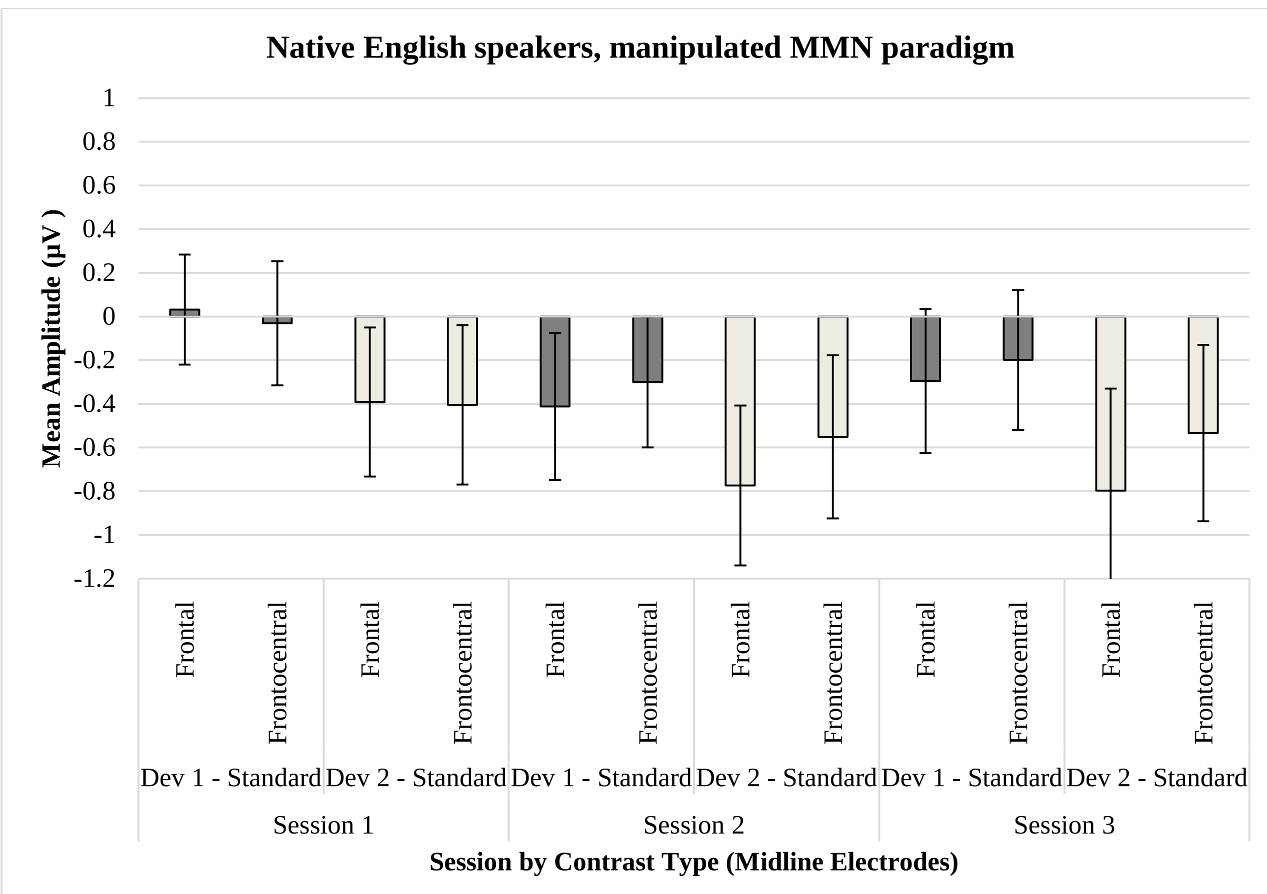

Figure 11. Native English speakers, session by contrast in manipulated MMN paradigm.

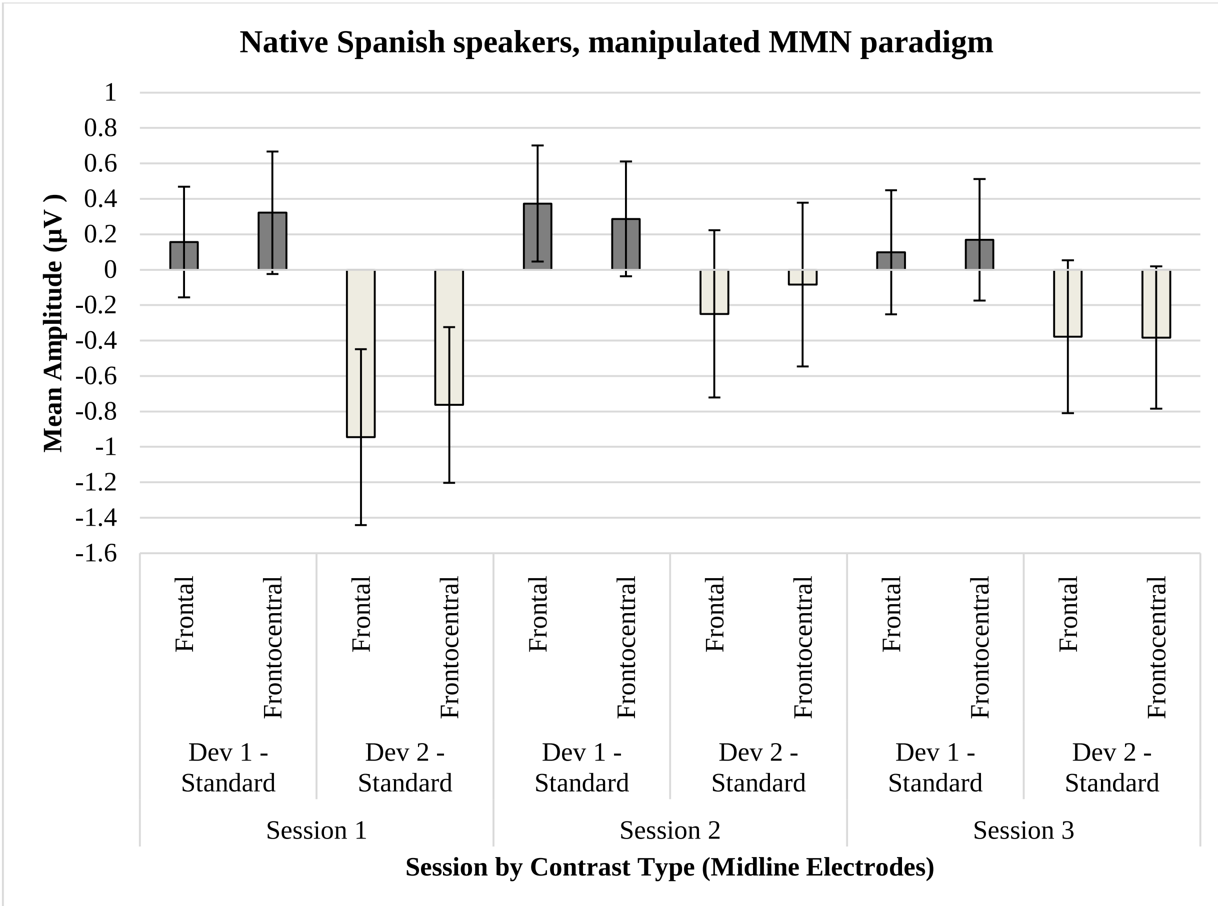

Figure 12. Native Spanish speakers, session by contrast in manipulated MMN paradigm.

3.2.2.1 Hindi vs. English comparison.

Comparisons between native English and native Hindi speakers for Session 1 showed a significant three-way interaction between language group, block, and laterality, F (2,58) = 7.53, p = .001, η_p^2 = .21; follow up tests revealed no significant differences between means but the amplitudes suggest that native English speakers showed more negative responses over midline and right electrode sites for the first block only compared to native Hindi speakers. Comparisons between language groups in Session 3 showed an interaction between language group and laterality, F (2,58) = 3.46, p = .05, η_p^2 = .11; there were no significant differences between means in the post hoc analyses, but native English speakers appeared showed more negative responses over left and midline electrode sites compared to native Hindi speakers.

3.2.2.2 Hindi vs. Spanish comparison.

Comparisons between native Spanish and native Hindi speakers in Session 1 showed a significant main effect of difference type, F (1,23) = 5.25, p = .03, η_p^2 = .19, such that the b-deviant response was more negative than the p^h-deviant response. This was qualified by an interaction with language group, F (1,23) = 4.42, p = .05, η_p^2 = .16; follow up tests did not show significant differences between the means but native Spanish speakers seemed to show a more negative response to the b-deviant response whereas native Hindi speakers showed no difference between the two difference responses, $F < 1$. There was also an interaction with lobe, laterality, and language group within Session 1, F (4,92) = 3.50, p = .02, η_p^2 = .13; post hoc tests did not reveal significant differences between means. Visual inspection of the means suggests that native Spanish speakers showed more negative responses overall, but possibly less negative in frontocentral and central electrodes compared to native Hindi speakers. Finally, there was an interaction with difference type, lobe, and laterality, F (4,92) = 3.01, p = .03, η_p^2 = .12; follow up tests showed that the frontal midline and central right sites had more negative responses to b-deviant response except to right electrode sites for the p^h-deviant response (all p's < .05).

3.2.3 Summary of ERP data.

See Figures 11 through 24 for ERP waveforms across all electrodes sites by session and native language group. Native Hindi speakers showed the predicted effect such that there were no differences between the p^h-deviant and b-deviant responses; this confirms that both contrasts were native to their first language (Hindi) and therefore they both showed negativity. Native English and native Spanish speakers showed a significant difference between the contrast pairs, such that the b-deviant response was more negative across all sessions compared to the p^h-deviant response. Although the initial difference was expected in the first session, the hypothesis also stated that this difference should lessen during training, as they learned to perceive the nonnative contrast. Analyses between sessions show that both native language groups showed the same pattern of responses: the b-deviant response was more negative across all sessions and blocks. Mean amplitudes suggest that the p^h-deviant response became more negative throughout the training for native English and native Spanish speakers, however it was not a large enough difference to overcome the negativity for the native contrast (Deviant 2 – Standard). Comparisons between native Hindi and learner groups confirmed that native Hindi speakers showed no differences between responses to the phonetic contrasts; native English speakers and native Spanish speakers consistently showed a more negative MMN to the b-deviant response. The lack of differences between native language groups for Session 3 may suggest that the native English and native Spanish speakers started shifting their perceptual representation of the nonnative contrast such that it was no longer significantly different from native Hindi speakers; however, this must be interpreted cautiously because the absence of a difference does not necessarily mean that responses were comparable between language groups. Finally, the significant effects of block throughout the analyses suggest that the second block of both MMN paradigms have more negative responses in general, indicating that the participants may have experience a training or practice effect from hearing the blocks more than once. Importantly, we are interested in examining how training and feedback affects the ERP responses because better performance on training may lead to a more negative MMN response for the predicted contrast; correlations between these variables are detailed below.

3.2.4 Correlational analyses.

Refer to Tables 3 and 4 in Appendix B for correlations between training, individual difference measures, and electrode sites. Correlations between individual difference measures and the difference types in the ERP measures showed that native English speakers had significant correlations between their performance on the training and the MMN response for the p^h-deviant response in Sessions 2 and 3 across several electrodes sites, including right frontal and frontocentral electrodes, suggesting that the training did influence responses on the p^h-deviant response. Additionally, native English speakers showed significant correlations between training performance and the b-deviant condition only in Session 1. The training correlation between p^h-deviant and training would only show up in either Session 2 or 3 because there was no training prior to Session 2. The significant correlations between training and the p^h-deviant difference in Session 3 may reflect the consolidation of the MMN response at the canonical electrode site for MMN responses. Additionally, there were significant correlations between performance on the Raven's Matrices test and p^h-deviant difference in all three sessions. This suggests that the better the participants performed on the Raven's matrices test, the larger difference was seen for the p^h-deviant condition. Finally, there were significant correlations between Flankers reaction time and the b-deviant responses only in Session 3 across the left and midline central electrode sites, suggesting that the faster the reaction time on Flanker's, the larger the difference seen between the b-deviant response.

For native Spanish speakers, significant correlations between training performance and the first difference pair, ph-deviant (/p^ha/-/pa/), were found only with left electrode sites. Training performance and the second difference pair, b-deviant (/ba/-/pa/), were only correlated in Session 3, such that higher accuracy in training was correlated with more negative responses across frontal and frontocentral electrode sites for the pair. This suggests that the training consolidated the between-category deviant pair within the native Spanish speakers by the end of training, which was contrary to what was seen with the native English speakers. The native Spanish speakers showed significant correlations with the Raven's matrices in Session 1 for only the Deviant 1- Standard response right and midline frontal, frontocentral, and central electrode sites, suggesting that higher accuracy on the Raven's test showed a more negative response to the p^h-deviant pair (the unique condition for native Spanish speakers). Finally, there were significant correlations between performance on the Raven's matrices and the b-deviant responses only in Sessions 2 and 3, suggesting that higher accuracy on the Raven's tests led to more negative ERP responses for the between-category condition. Correlations between individual difference measures that are not of theoretical interest in the present study were also found; see Tables 5 and 6 in Appendix B for a summary.

Figure 13. Native English speakers, Session 1, traditional MMN paradigm.

Figure 14. Native English speakers, Session 1, manipulated MMN paradigm.

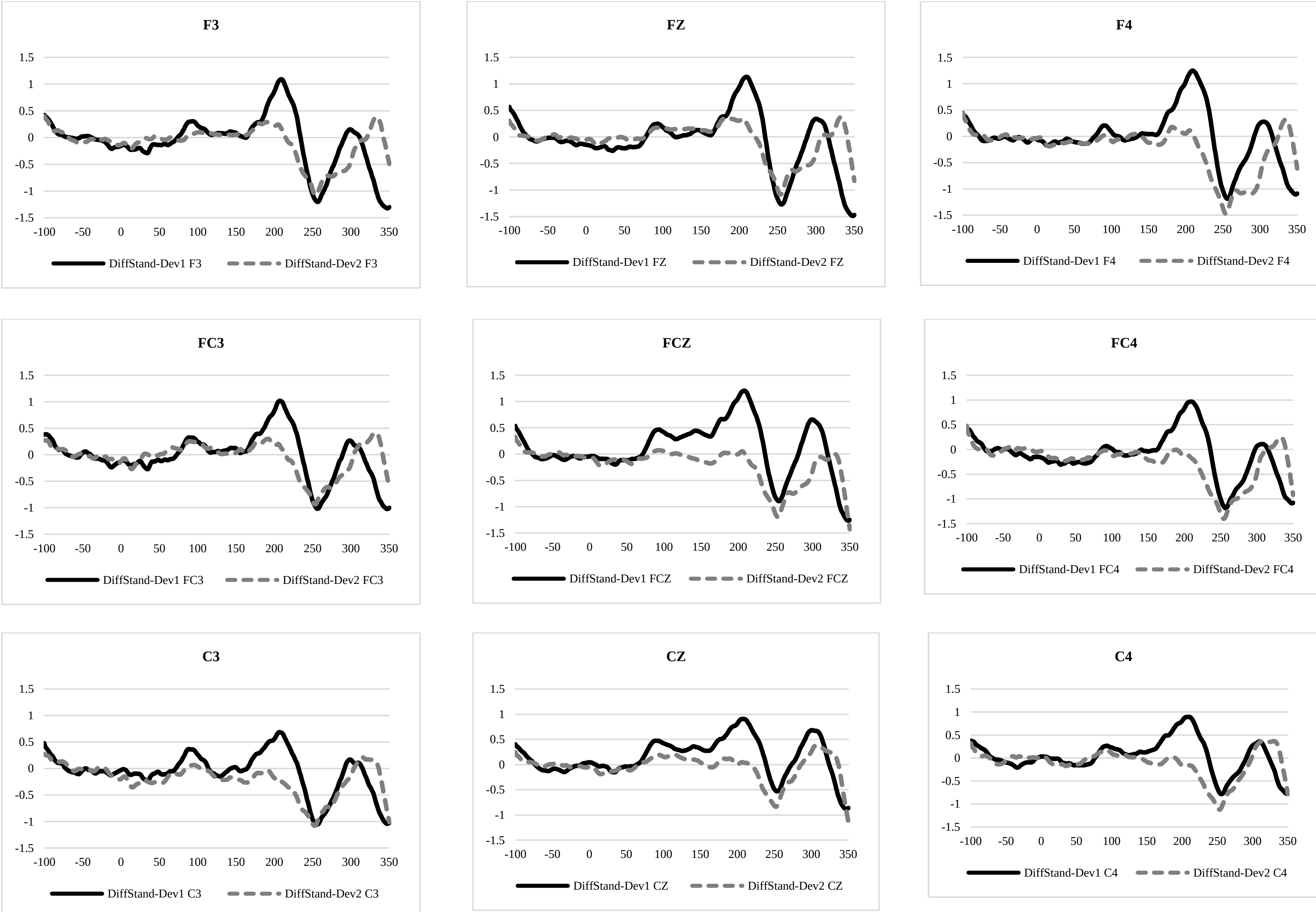

Figure 15. Native English speakers, Session 2, traditional MMN paradigm.

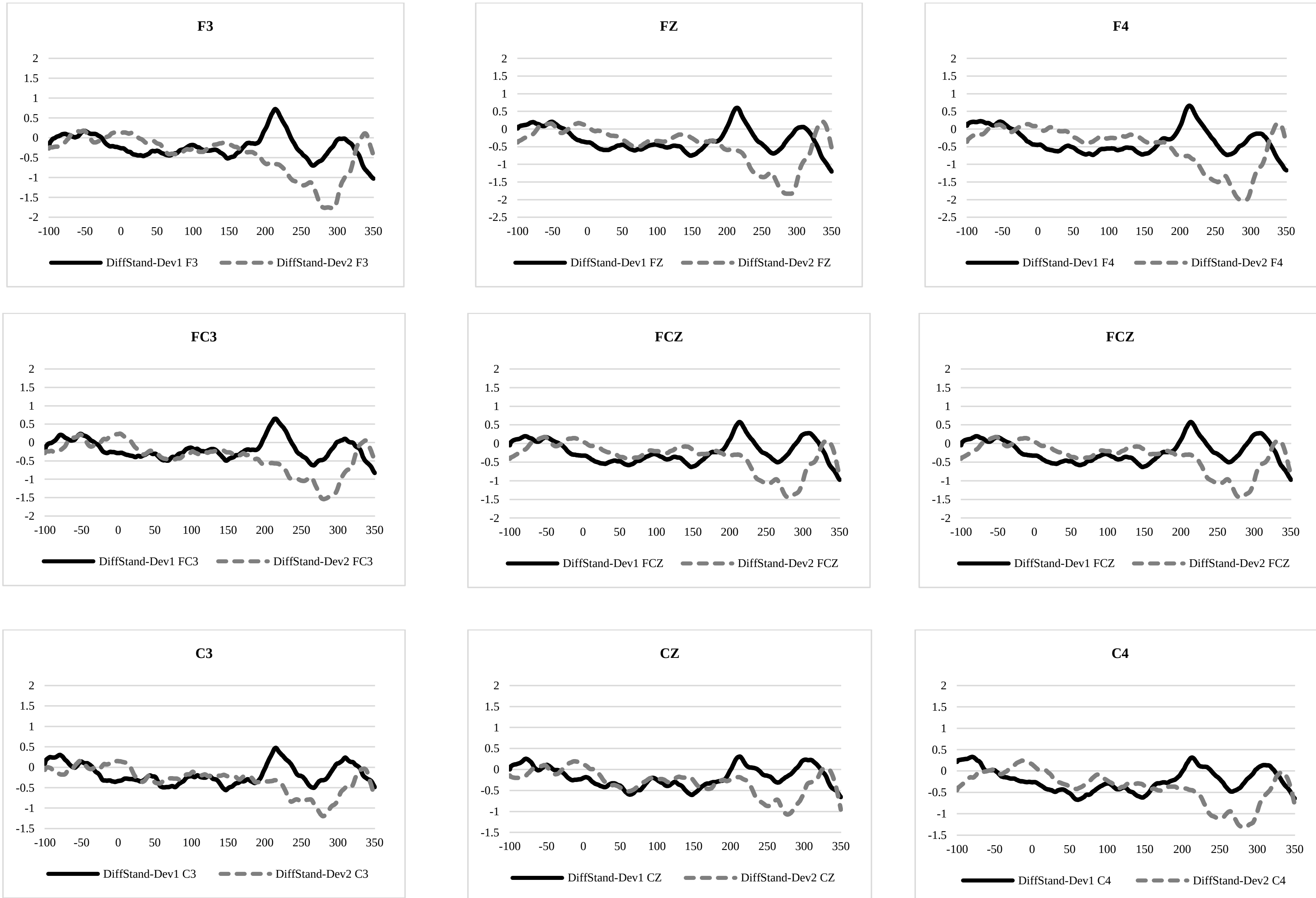

Figure 16. Native English speakers, Session 2, manipulated MMN paradigm.

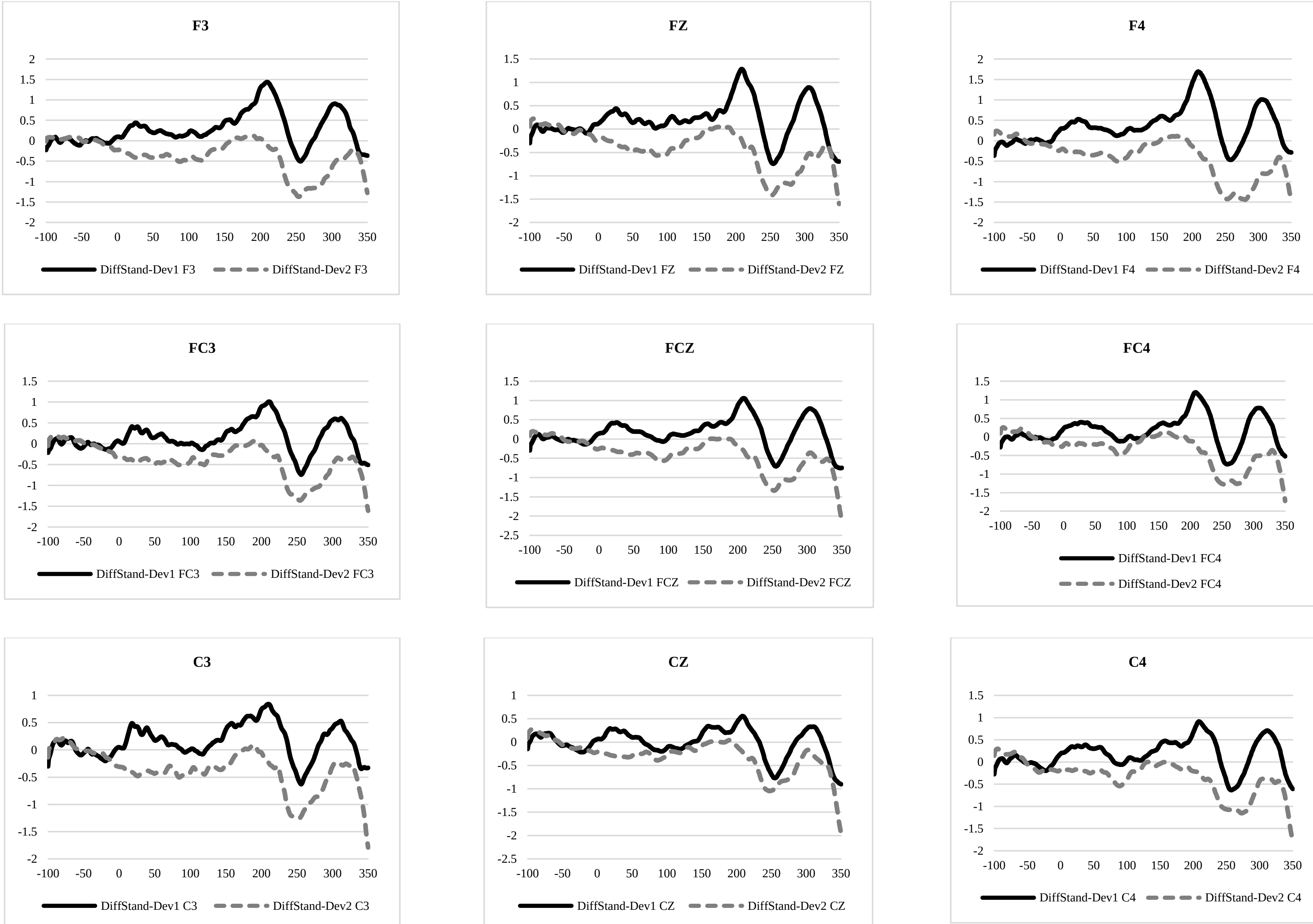

Figure 17. Native English speakers, Session 3, traditional MMN paradigm.

Figure 18. Native English speakers, Session 3, manipulated MMN paradigm.

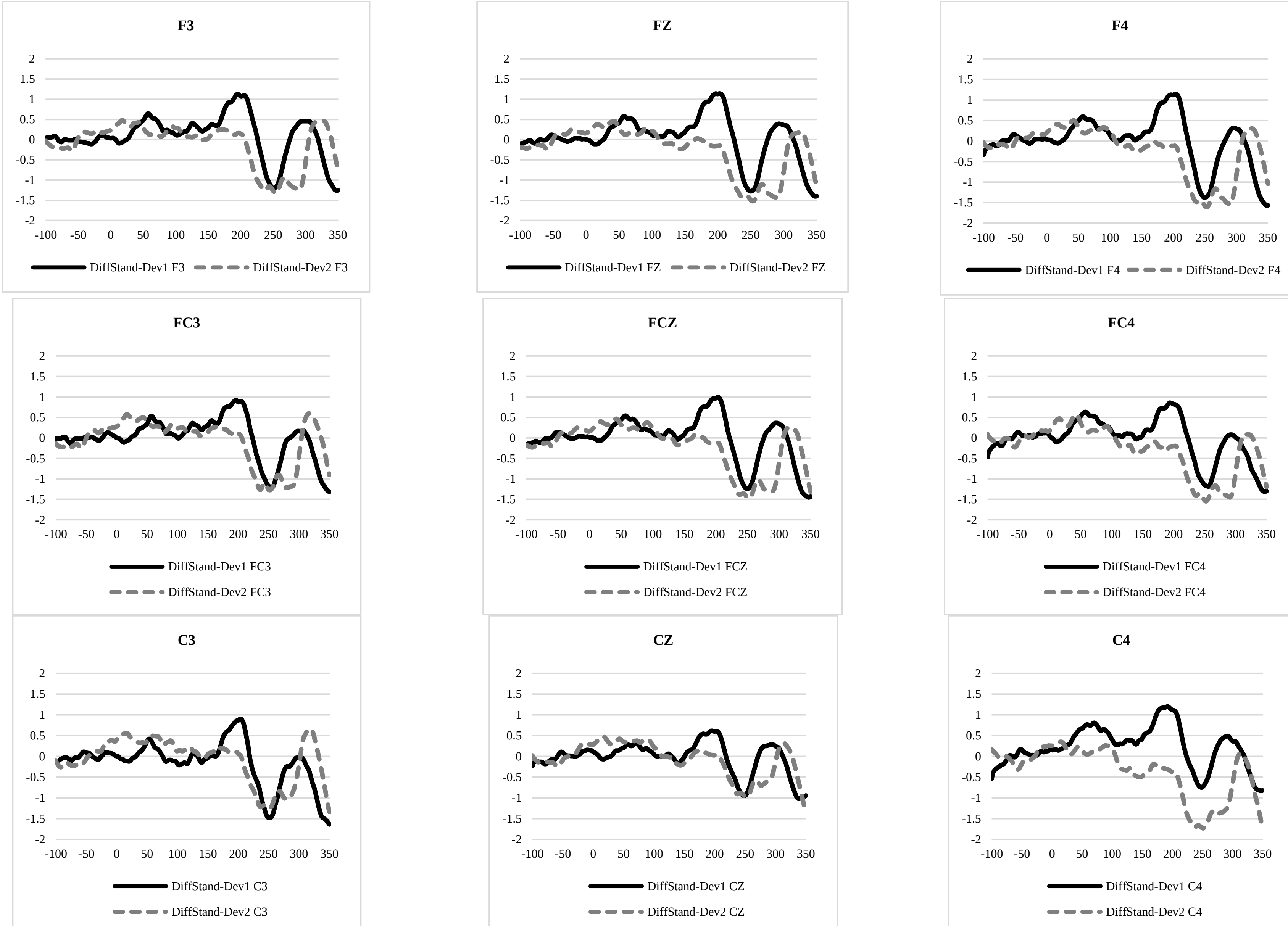

Figure 19. Native Hindi speakers, traditional MMN paradigm

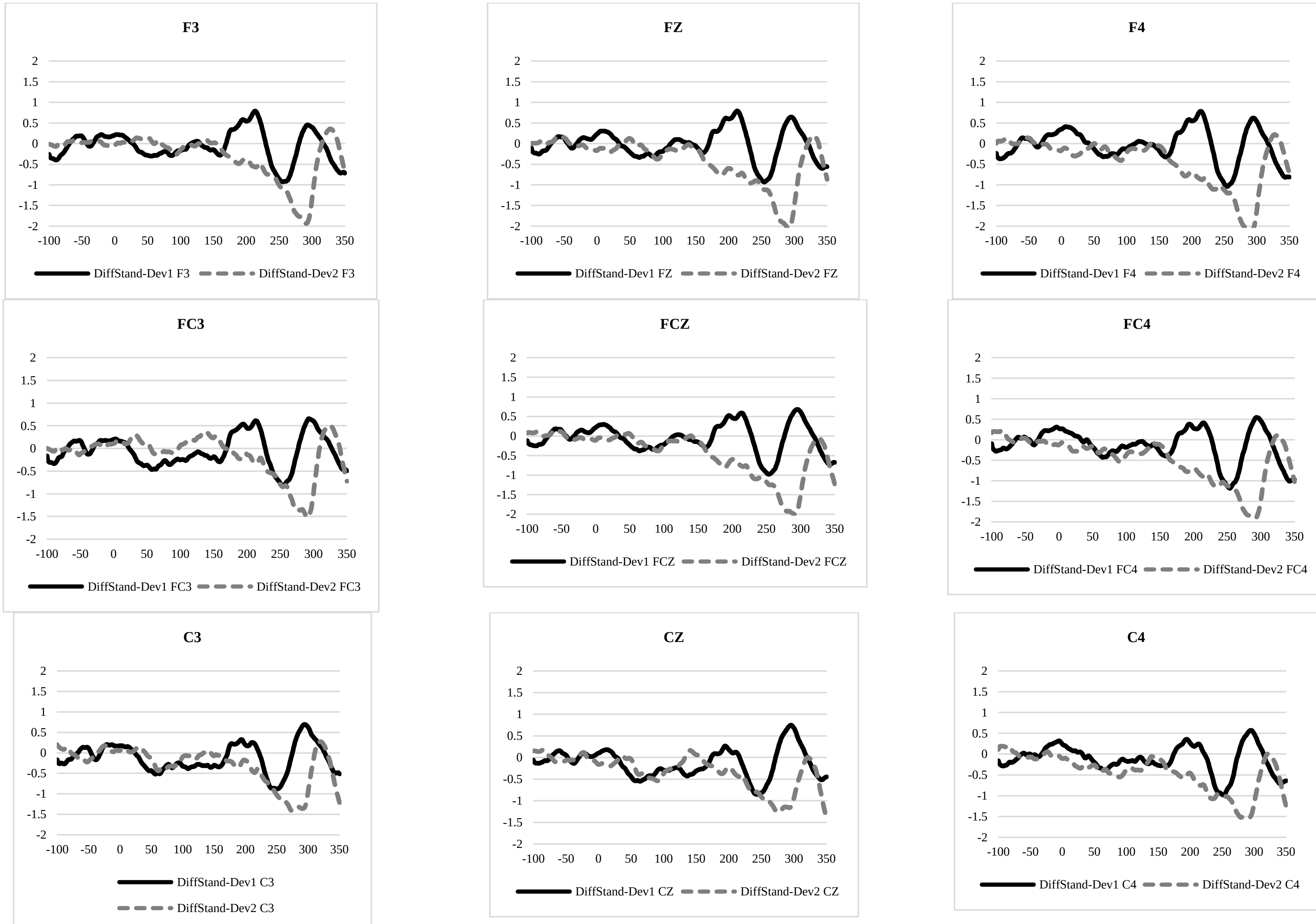

Figure 20. Native Hindi speakers, manipulated MMN paradigm.

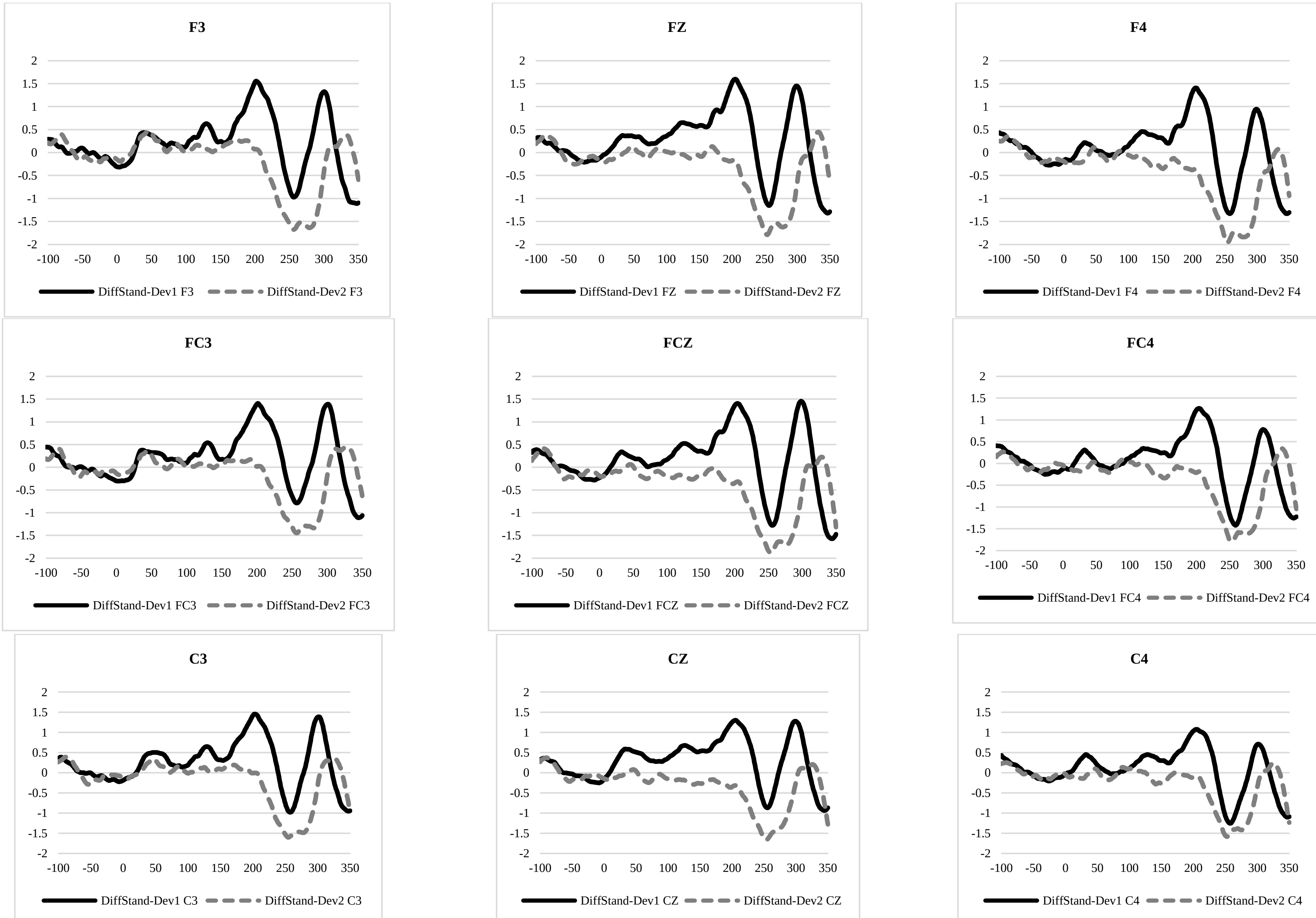

Figure 21. Native Spanish speakers, Session 1, traditional MMN paradigm.

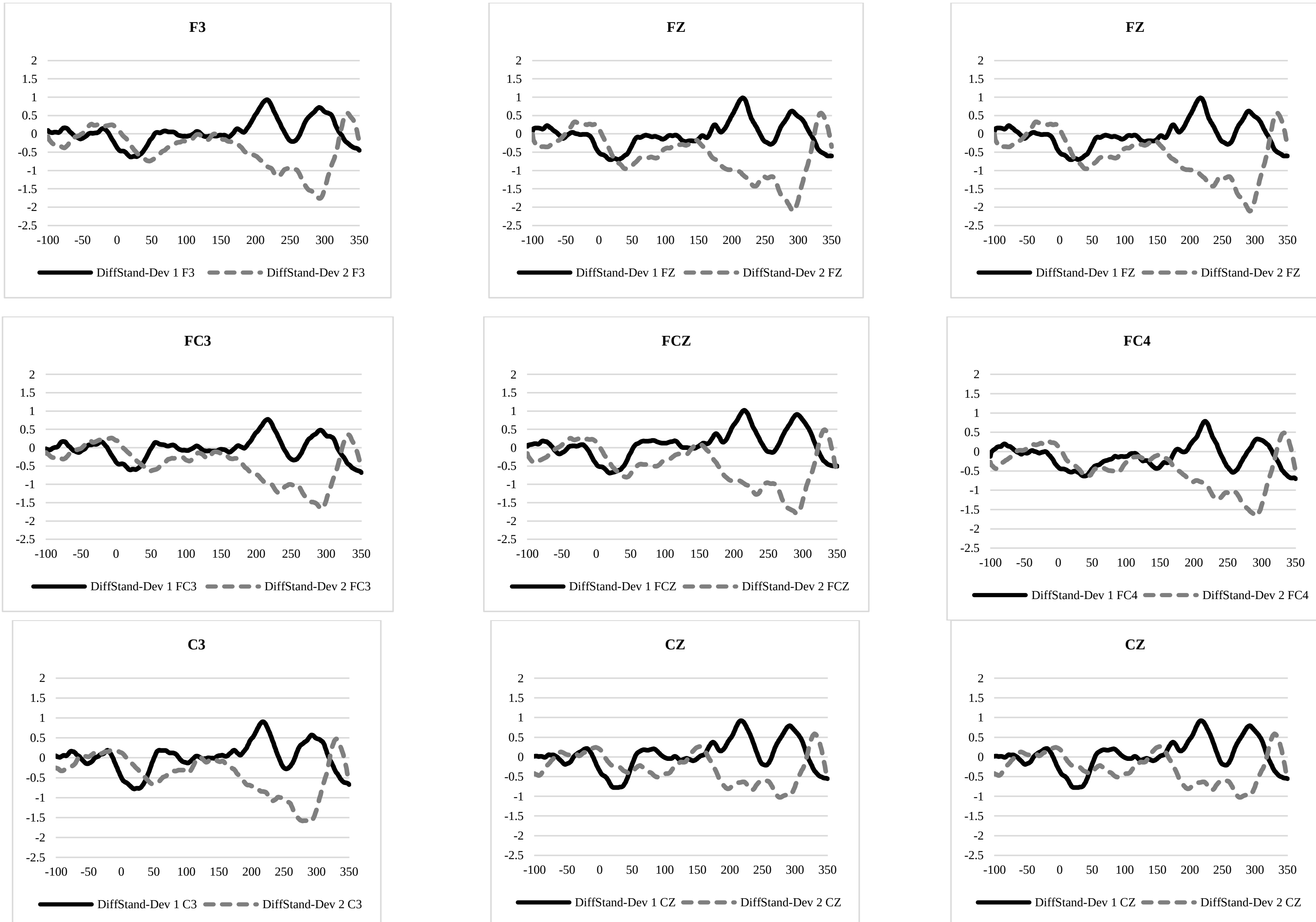

Figure 22. Native Spanish speakers, Session 1, manipulated MMN paradigm.

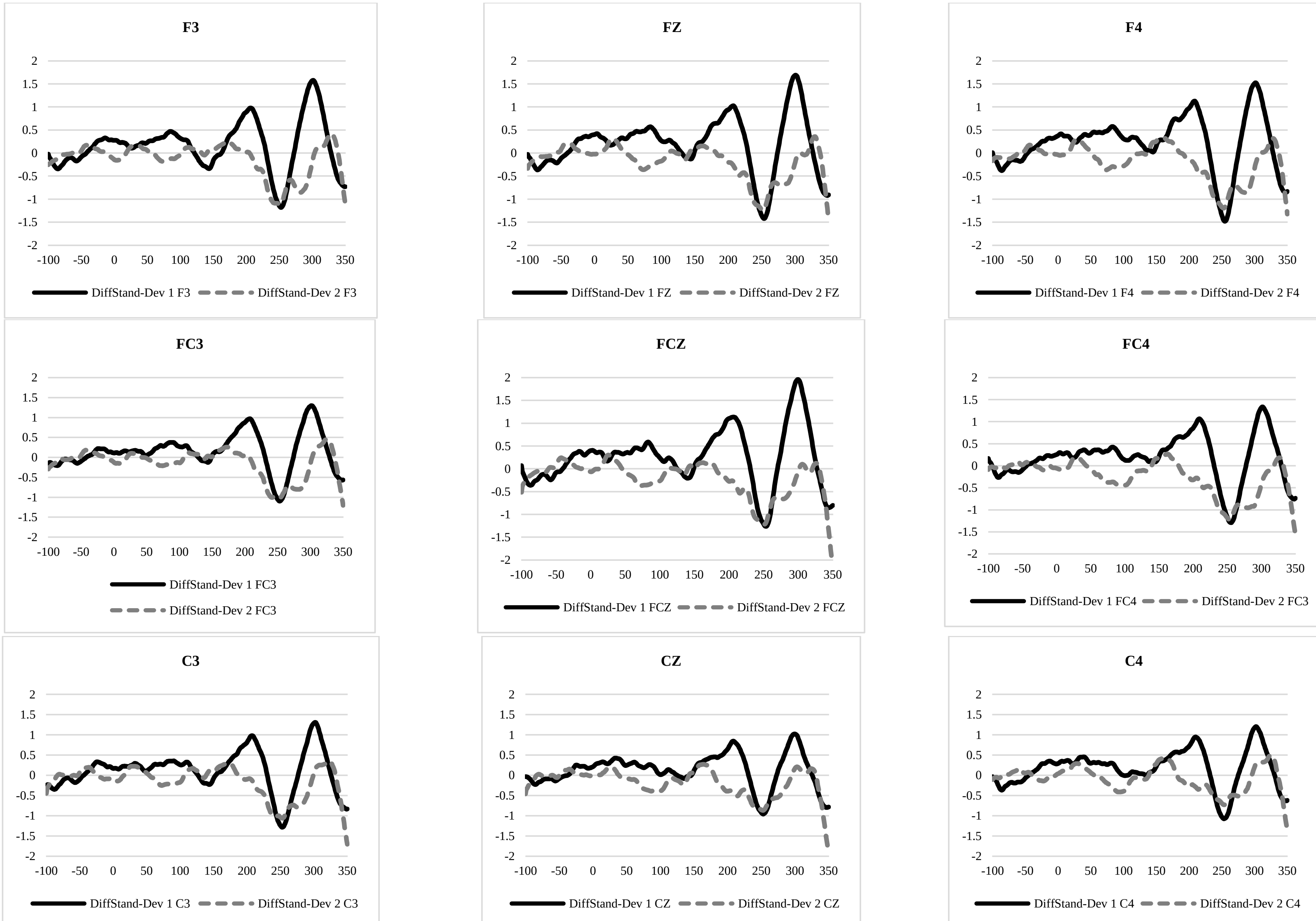

Figure 23. Native Spanish speakers, Session 2, traditional MMN paradigm.

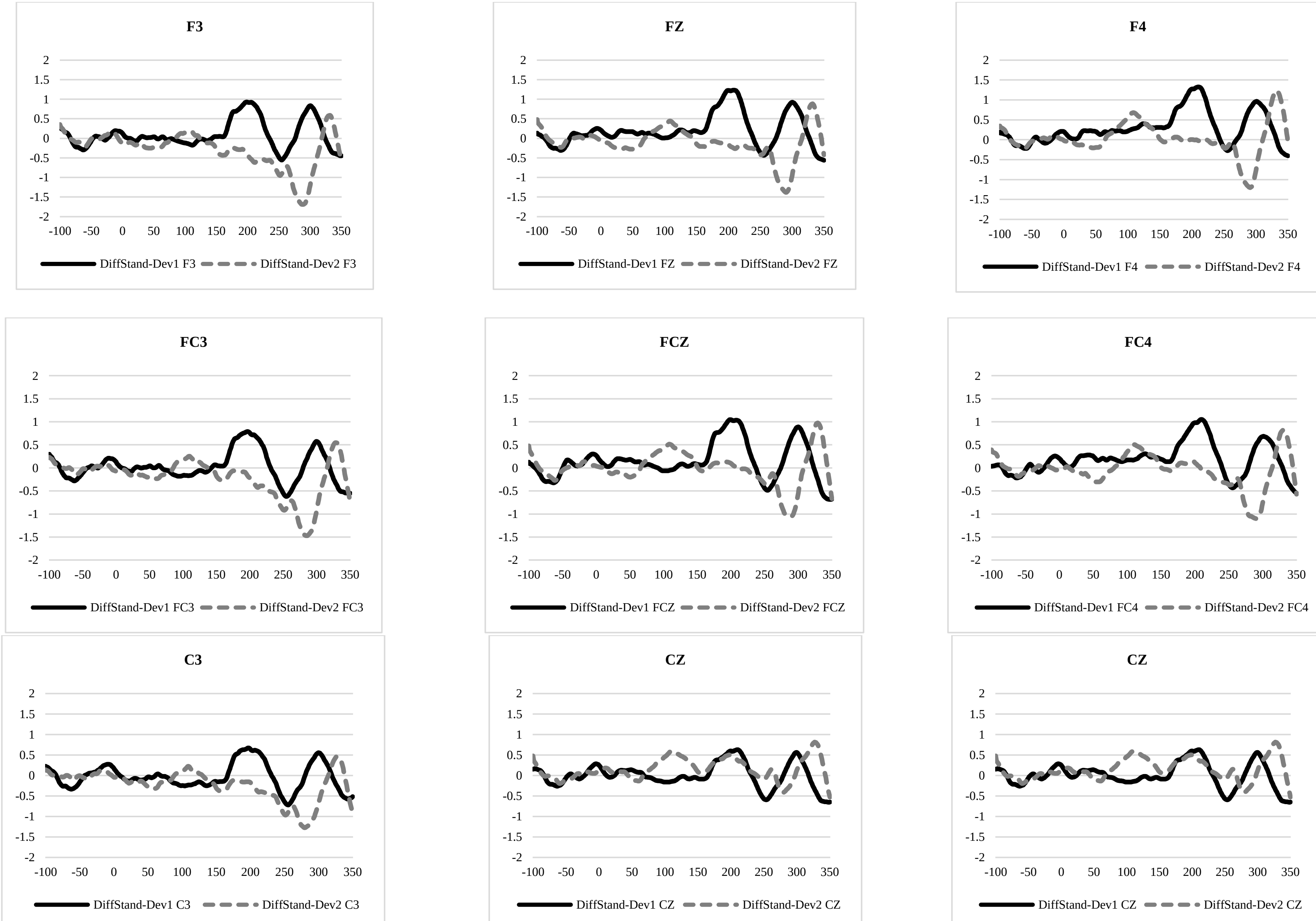

Figure 24. Native Spanish speakers, Session 2, manipulated MMN paradigm.

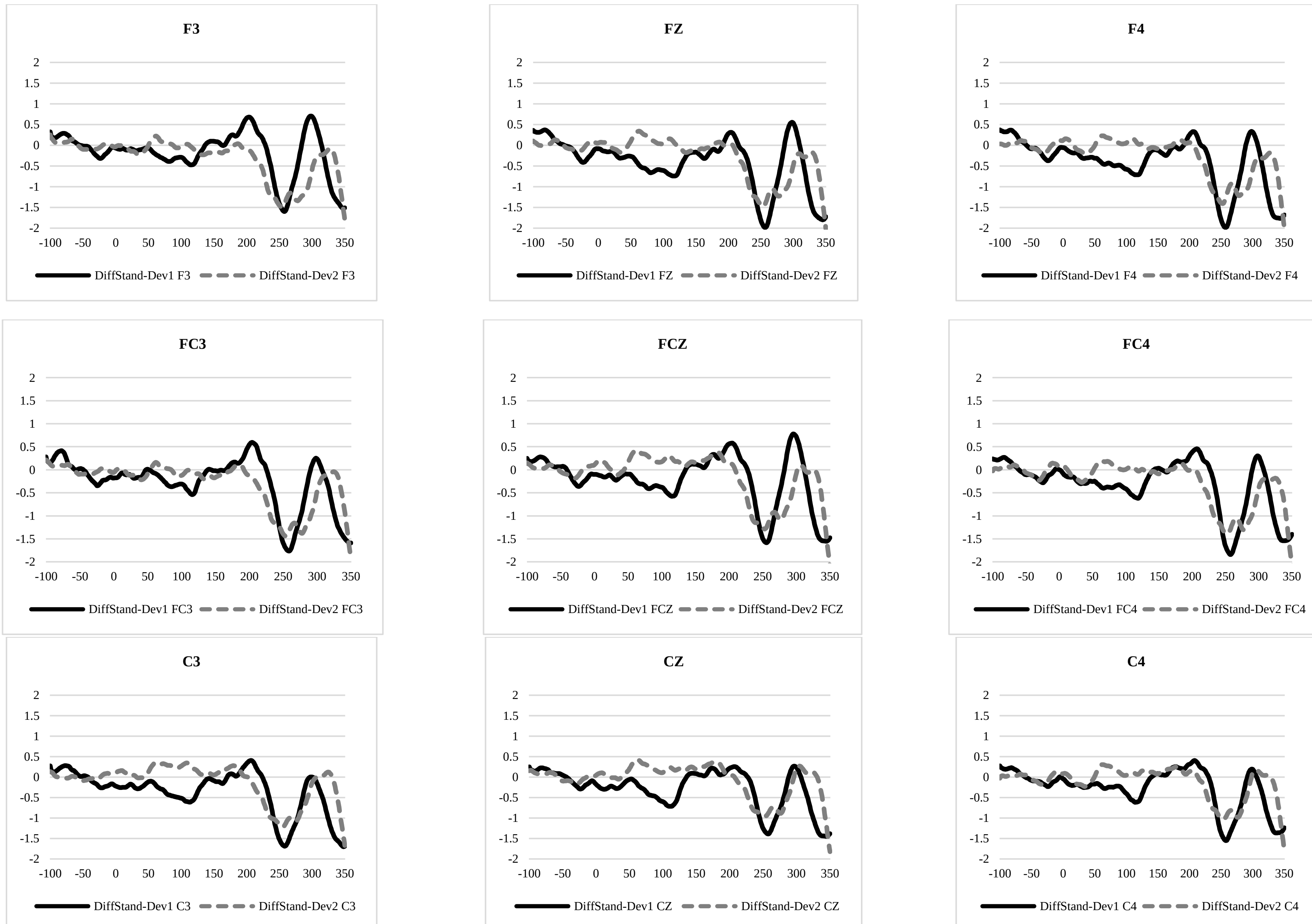

Figure 25. Native Spanish speakers, Session 3, traditional MMN paradigm.

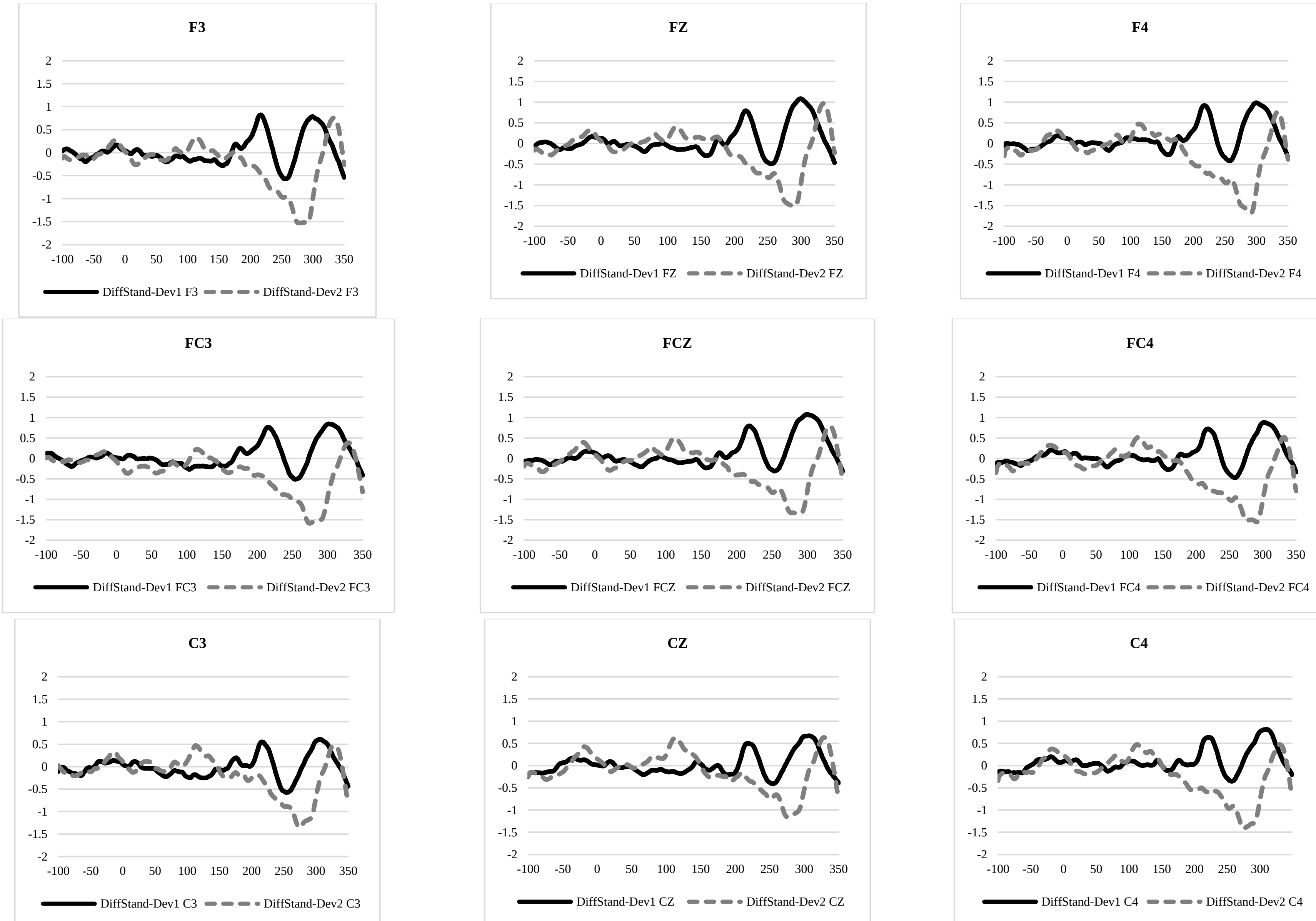

Figure 26. Native Spanish speakers, Session 3, manipulated MMN paradigm.

4.0 DISCUSSION

The present study compared how the first language can influence the perception and acquisition of nonnative phonetic contrasts; specifically, native English and native Spanish speakers were trained to perceive differences in Hindi phonetic contrasts that were present or not present in their first language. The phonetic contrast tested here, /p/ and /p^h/, is a minimal pair distinction in Hindi, such that words or phrases that differ in only this contrast have distinct meanings. However, native English speakers categorize both the /p/ and /p^h/ phones under the /p/ category, i.e., there is no minimal pair distinction, and native Spanish speakers do not have the fully aspirate /p^h/ phone in their phonetic repertoire at all. Native English and native Spanish speakers underwent three days of training with an AXB task to learn to perceive the differences between these phonemes as meaningfully distinct. Behavioral and ERP measures during training were taken to examine the baseline ERP response to the contrasts, ERP responses during the training, and ERP responses after the training was completed.

Results showed an overall effect of difference type for the native English and native Spanish speakers compared to native Hindi speakers. Both learner groups showed a significant difference in the MMN response to the /p/-/p^h/ contrast and the /p/-/b/ contrast, such that there was a more negative response to the across-category phonetic contrast in both language groups, /p-b/, compared to the /p-p^h/ contrast. Native Hindi speakers did not show a difference between contrasts types. These results suggest that the traditional MMN response was elicited successfully by the paradigm within the first session. Furthermore, the learner groups were expected to show differences between sessions in their MMN response to the contrasts. Native Spanish speakers showed a significant difference between contrast responses from session 1 to session 2 and from session 1 to session 3, suggesting that the the response changed during training; however, there was a significant difference between the native Hindi speakers and the native Spanish speakers indicating that the difference between the responses to the phonetic contrasts did not match those of the native speaker group. The same held true for native English speakers; the difference in responses between phonetic contrasts held throughout the session as opposed to lessening like the native Hindi speaker responses. Generally, both learner groups showed a greater MMN response to the contrast that was already present in their native language (/b/ - /p/) and the response did not change significantly over the course of the training.

Behavioral results showed that participants performed better on the training with feedback compared to no feedback, confirming previous work showing that feedback is beneficial to accurate perception of phonetic contrasts. Contrary to the hypothesis that attentional feedback could provide an additional benefit because it gives more information, there were no behavioral differences in performance between standard and attentional feedback. The lack of difference between the standard and the attentional feedback conditions could be due to the fact that the stimuli were kept constant across all linguistic features except aspiration and therefore, participants, if they were receiving feedback and able to notice the difference between the sounds, only had one difference to which they had to pay attention. That is, regardless of the type of feedback participants received, there was only ever one difference between the unaspirated and aspirated /p/; participants in the standard feedback condition would be paying attention to the same feature as those participants in the attentional feedback condition, minimizing the differences in performance between conditions.

Previous research has examined how the MMN response changes over the course of training with speakers of a single language and specific phonetic contrasts (Kraus et al., 1995; Kraus et al., 1997). The current study extends this to examining MMN responses across language groups to the same phonetic contrast. Results showed that better performance during the training resulted in more negative /p^{h}a/-/pa/ discrimination in native English speakers, suggesting that the training transferred to phonetic perception for the within-category contrast that was intended to become a between-category contrast. However, because there was still an overall difference between the deviant MMN responses in Session 3, the training did not improve their discrimination enough to show comparable responses to the native Hindi speakers (who showed no difference between deviant responses). It could be that with more training, this effect could strengthen such that more training would increase the MMN response to the within-category deviant enough to have it become a "between-category" deviant. Native English speakers have decades of exposure to their L1 and reversing the within- and between-category phonetic representations may take more than three one-hour sessions of training.

There was a converse effect for native Spanish speakers, such that better performance during the training seemed to consolidate the between-category /ba/-/pa/ difference as opposed to heightening the response to the unique aspirated /p/. Although the /ba/ stimulus used as the deviant was clearly in the VOT range of a /b/ for native Spanish speakers, it could be that the training highlighted the /ba/-/pa/ difference as opposed to the /p^{h}a/-/pa/ difference which is not present at all in Spanish. Thus, this suggests that forming a new phonetic category requires more exposure to the phonemes than was provided in the current manipulation. Finally, the amount of training may have influenced the MMN results for the between-category deviant because training on a phonetic contrast causes a heightened MMN response due to increased coherence between neurons for that stimulus (Kraus et al., 1995). That is, training consolidates the response to the phonetic contrasts and increases the number of neurons that fire for that particular stimulus. Here, native Spanish speakers show a heightened MMN response to the between-category deviant throughout training, including in the last EEG session where they received no training.

Importantly, this underscores the importance of cue availability and reliability in the environment when forming new phonetic categories. The UCM predicts that based on the cue validity (availability and reliability), a unique condition can either be processed more or less efficiently. The lack of MMN response for the /p^ha/-/pa/ after training for the native Spanish speakers suggests that the cue for /p^h/ was not as available or reliable as it could have been; in this case, it is important to note that not only did participants receive cues during training, but cues from the environment would also play a role in perception. Although the native Spanish speakers all resided within the United States at the time of testing and therefore exposed every day to English, many of them were new to the country and still used Spanish in their everyday lives with family and friends. A longer training period may have affected their MMN responses in the long term. However, because native English speakers have had much more reliable and available cues for aspiration due to exposure to English from birth, the importance of cues for native English speakers are confounded with native language exposure. Speakers weight features in language that help them discriminate relevant speech sounds from others; for example, native English speakers do not have consonant or vowel duration as a relevant cue (i.e., minimal pair distinctions) and therefore do not weight the cue of duration highly during processing. Cue re-weighting can overcome these differences, such that intensive training can lead to native English speakers using duration as a relevant cue (e.g., Ylinen et al., 2010). In this case, it is hard to overcome the lack of weighting for aspiration in English to form a new aspirated category and our training may not have been enough to do this. Therefore, there were unreliable cues for aspiration for native Spanish speakers due to the mixing of English and Spanish in their daily lives, and not enough cues during training to help native English speakers reweight them to be salient.

Although the nonnative phonetic contrasts either did or did not exist in the participants' L1 (the aspirated /p/ exists in English, but not in Spanish), the continued significant difference between the two by the end of training highlights the endurance of the L1 phonetic categories. Native Hindi speakers show an MMN response to both deviant-standard conditions because they are both between-category phonemes; however, native Spanish and native English speakers still showed a significant difference between the two, even though the amplitude changed from Session 1 to Session 3. The NLM posits the idea of neurally committed wiring for phonetic representations within a language once a listener reaches adulthood and the brain's plasticity decreases (Kuhl & Rivera-Gaxiola, 2008); the results here suggest that the neural hardwiring may be too strong to overcome for native English and native Spanish speakers even when the contrast to be learned exists in the L1. However, this is tempered by the interaction with training – the better performance on training led to increased MMN responses for the aspirated /p/ contrast, indicating that more effective or more exposure to training may, in fact, reverse the hardwiring to allow for the same native-like responses to the nonnative contrast. For native Spanish speakers, forming a completely new category over the course of three days may not be enough time to affect the neural commitment of the brain's hardwiring to the phonetic perceptual representations.

Importantly, the results here indicate that the same kind of response to both within and between category deviants can be elicited with a traditional MMN oddball paradigm and with one in which the standard category is varied within a phonetic category. Due to the inherently variable nature of speech, examining how the brain responds to natural speech variation for phonetic categorization allows us to more confidently examine how the MMN response indexes the underlying phonetic representations in the L1 and how those change with exposure to L2 phonetic representations. Previous work has shown that the MMN can be elicited with varying nonspeech stimuli in the standard presentation (Winkler et al., 1990), but to our knowledge, varying the standard presentation within a phonetic category of natural speech has been shown here for the first time and suggests that the MMN is a clear marker of phonetic status of the incoming stimuli as it is elicited with a variable standard presentation. Further analysis should examine peak amplitude differences between the traditional and the manipulated oddball paradigm blocks to uncover more information on the exact physiological differences between these two presentations. Importantly, the similarity between the MMNs elicited in both paradigms speaks to the success of using this type of paradigm to provide more evidence that the MMN captures phonetic differences between sounds, and not just acoustic. Although the MMN was more variable in the manipulated paradigm, there were significant effects for both contrasts when the standards were varied. Importantly, the native English speakers showed an MMN to the /p^h/-/p/ contrast in the manipulated paradigm much earlier compared to the traditional paradigm. This suggests that the inherent variability within the standard presentations helped native English listeners to more accurately encode aspiration as a relevant feature.

Previous research examining variability during training and learning of an L2 has shown that variability leads to better learning for words (Perry, Samuelson, Malloy & Schiffer, 2010), grammar (Gómez, 2002), morphosyntax (Eidsvag, Austad, Plante, & Asbjornson, 2015), and reading (Apfelbaum, Hazeltine, & McMurray, 2013). Importantly, input variability has been extensively studied for speech perception and production and previous research has demonstrated that input variability can lead to more accurate perception of novel phonemes in the long term, though there may be short term costs (Logan, Lively, & Pisoni, 1991, 1993; Nygaard & Pisoni, 1998; Pisoni, 1993). In this case, it may be that simply being exposed to the stimuli in the manipulated MMN paradigm in tandem with the training allowed participants to better shift their perceptual representation to consider the aspirated /p/ as a separate category. However, previous research has also shown that this variability can transfer to other features, such as voices, gender, talker identity, etc. In this case, it is interesting to note that exposure to the manipulated MMN paradigm did not seem to enhance MMN responses in the traditional MMN paradigm, though due to the mixing of blocks and overall exposure to training, it is difficult to examine if there was transfer between more and less variable paradigms. Finally, this same response was not seen for native Spanish speakers, who showed comparable responses between the manipulated and traditional MMN paradigm across sessions. This difference between native English and native Spanish speakers may arise from differences in similarity between the native Hindi phonemes and the respective phonemes in English and Spanish. That is, because native English speakers have aspiration as a systematic feature in their native language, they may be better able to use the variability in the manipulated paradigm and the training to shift their perceptual representation, as opposed to native Spanish speakers who must build up a new category.

The similarity between the language groups was manipulated to examine how cross-language similarity between languages would influence the neural responses. According to the

PAM, the aspirated /p/ could be categorized as a native phonemic segment for native English speakers because it does exist in English under the consonant category of /p/; for native Spanish speakers, the aspirated /p/ would most likely be categorized as "uncategorized", such that it is not a native language phoneme. Within the PAM categorizations, the categorized exemplars are then rated on goodness-of-fit to examine how "good" or "bad" they are in contrast to the L1 category. The present study did not specifically ask native English speakers to rate how well the Hindi aspirated /p/ fits into the English /p/ category, although the identification task does suggest that they had no issues categorizing it as a /p/. The PAM would most likely suggest that the /p^ha/-/pa/ contrast would be categorized as Single Category assimilation, wherein two nonnative phones are perceived as belonging to the same native category; in this case, both the Hindi aspirated and unaspirated /p/ belong in the larger English /p/ category. In accordance with the PAM predictions, native English speakers do not show an MMN response to this contrast pre-training because it is assimilated to one category.

Accordingly, the acquisition of the aspirated /p/ would depend on whether or not the phone is categorized as a good or bad exemplar of the native category (Best & Tyler, 2007). In this case, we do not know if native English participants considered the aspirated /p/ a good or bad exemplar of the English /p/ category; the PAM predicts that speakers would have to learn a new phonetic category for this phone before being able to successfully use it at the phonological level. Based on the MMN responses seen here, it seems native English speakers started to build the new phonetic category by the end of training, but due to the training effects that enhance MMN responses in general (Kraus et al., 1995), the difference between the /ba/-/pa/ category and the /p^ha/-/pa/ category remained significant.

Alternatively, the presence of the aspirated /p/ in English phonetics may have caused competition in attempting to perceive and acquire it as a separate phone, according to the competition framework set forth by the UCM. The increase in MMN response to the aspirated /p/ contrast for native English speakers suggests that the similarity between the two phones contributed to more successful phonetic perception here, as measured by the MMN response difference from Session 1 to Sessions 2 and 3. The UCM would also predict this based on positive transfer between constructs that are similar between languages. In sum, the competition between the existing English /p^ha/ and the nonnative Hindi /p^ha/ was not strong enough to overrule effects of similarity between the languages. Examining these results in conjunction with the identification task, it seems that the difference between the /p^ha/-/pa/ contrast was only seen in the MMN responses and not in the overt behavioral responses (i.e., native English speakers did not show a difference in categorization of the /b/-/p/ continuum pre- to post-training). That is, participants showed a change in the MMN response from pre- to post-training for the /p^ha/-/pa/ contrast but this did not change their categorization of the entire continuum.

For native Spanish speakers, the perception of the uncategorized aspirated /p/ would depend on the similarity to the L1 perceptual category (SLM; Flege, 1995) and the relationship between that category and other possible similar phones (Best & Tyler, 2007). The aspirated /p/ for native Spanish speakers is uncategorized because it does not have a similar phone in the L1 (in general, Spanish does not use aspiration as a contrastive or even acoustic cue). This would suggest that it may be easy to learn perceptually because it does not have any competing or pre-existing phonemes that would hinder perception or acquisition. However, the PAM also suggests that the ease of forming a new perceptual category depends on the relationship between the to-be-learned phone and other existing phonetic representations (i.e., “its comparative relationships within the interlanguage phonological system", Best & Tyler, 2007). This suggests that other L1 phones that may be similar will also affect the perception and acquisition of the category. In this case, because Spanish does not use aspiration as a functionally contrastive cue, the Hindi aspirated /p/ would remain uncategorized with no similar phonemes in Spanish, suggesting that it should be fairly easy to differentiate from the unaspirated /p/. Our MMN results show that native Spanish speakers showed a significant difference throughout training for both contrasts; for the /p^{h}a/-/pa/ contrast, the difference between the phones lessened from session 1 to session 2 (M = .289) and from session 1 to session 3 (M = .038). This suggests that although the effect was overshadowed by the L1 between-category response to /b/-/p/, speakers did show a change in MMN to /p^{h}a/-/pa/. Interestingly, native Spanish speakers did seem to show a difference from pre- to post-training on their continuum categorization, such that post-training, they showed a cross-category boundary more similar to that of native Hindi speakers. This may indicate that the behavioral training affected their categorization of the /b/-/p/ continuum, though this result should be interpreted cautiously because the training performance and ERP results were not significantly different from the native English speakers. However, examining what may have led to this shift in perceptual category boundary would elucidate how categories can be shifted with exposure to nonnative phonetic contrasts. It may be that the training emphasized already-existing phonetic contrasts and allowed the perceptual boundary to shift and mirror the native Hindi speaker boundary; however, this only happened for the /ba/-/pa/ difference, as seen in the correlational analyses. Therefore, the shift in boundary was linked to an over-emphasis on the /ba/-/pa/ contrast and not due to the shifting of the /p^{h}a/-/pa/ contrast.

Alternatively, the training may have highlighted the /b/-/p/ difference for native Spanish speakers because it is possible that the /b/-/p/ distinction is not as frequent as it is in English.

That is, it is a salient cue because it is a minimal pair, but it may not occur as frequently as it does in English, leading to an increased response during training as participants hear it continuously (and in the MMN paradigm as well, leading to increased exposure). An initial search for exhaustive lists of /b/-/p/ minimal pairs does not appear to show that there is such a large difference between Spanish and English in /b/-/p/ minimal pairs, but a more in-depth search using corpus data would help answer this question.

In sum, the results of the present study show that the similarity to the first language can influence the MMN response, such that the more similar contrast (i.e., the between-category contrast) showed a clear and definite MMN response before training; the /pha/-/pa/ contrast showed a more clearly defined MMN response after training, suggesting that our training worked even though it also enhanced the MMN response for the between-category contrast. There were no differences between language groups across any of the sessions; both speaker groups showed a larger MMN response for the between-category deviant /ba/-/pa/ across all sessions. However, the correlations show that the language groups showed differential results, such that the performance on the training task correlated with mean amplitudes for different contrasts. This suggests that depending on the native language, training augmented perception of either the within-category deviant (English) or the between-category deviant (Spanish). A possible explanation for this could be that the training highlights already-existing contrasts in the languages: the native English speakers have the aspirated /p/ as part of their perceptual category and the training successfully emphasizes the perceptual boundary that they are asked to make. Because native Spanish speakers do not have the aspirated /p/ as part of their initial perceptual category, the training only serves to highlight the already existing contrast of /b/-/p/, such as seen in training studies in which the existing MMN has been augmented (Kraus et al., 1995).

Importantly, the linguistic feature (aspiration) used to perceive the difference between the critical contrast /p^h/-/p/ was perceived at different time points during acquisition; that is, native Spanish speakers, who do not use aspiration as a systematic linguistic feature, were able to begin to perceive the difference between /p^h/ and /p/ earlier during training (by Session 3) compared to native English speakers, who did not show increased negativity for the /p^h/-/p/ contrast by the last day of training. This suggests that the similarity to the L1 can influence how the MMN is elicited during acquisition of new nonnative contrasts. Although similarity between languages (in this case, /b/-/p/ between all three languages) overshadowed the results seen for the critical /p^h/-/p/ contrast, it is possible that the competition between the English /p^h/ and Hindi /p^h/ did not allow native English speakers to perceive the /p^h/ as a separate phoneme before the native Spanish speakers, who just had to learn to use the aspiration as a relevant linguistic cue.

It is important to note that native Spanish speakers already had another language that they were exposed to that may have influenced their responses – that is, due to the fact that they were immersed in their L2 language and culture, native Spanish speakers were already juggling two language systems at the time of testing. This may have contributed to the lack of difference between native English and native Spanish speakers in ERP mean amplitudes for the different contrast pairs because native Spanish speakers were exposed to English during their daily lives. This exposure may have already started to influence the perceptual representation of the phonetic categories. Although this is a possibility, due to the pre-attentive and early nature of the MMN response, it is unclear how much influence exposure to English may have had. That is, because the MMN response taps into the earliest phonetic representation of the incoming stimulus that does not require attention or conscious access, it seems unlikely that native Spanish speakers would use their second language as the mediator for perceiving nonnative phonetic contrasts.

Previous research has shown that learning a third language (L3) can be mediated through the L2 instead of the L1, depending on the similarity between the L1/L2 to the L3; learners can use their L2 to learn their L3 if it is more typologically similar to the L3 (Cenoz, 2001; Hammarberg, 2001; Ringbom, 1987). However, much of this prior research has examined grammar and vocabulary learning, or morphosyntactic and syntactic processing (Cenoz, 2001) or case studies of one participant with a varied linguistic background (Hammarberg, 2001). For speech perception and production, this idea remains to be explored in depth in future research.

The current study also provided new information about how individual differences can affect L2 speech perception. These results may be spurious, but the frontal and frontocentral electrodes show significant correlations with Raven's that carry over from one session to another, suggesting that it may not be due to chance. Our results show significant correlations with performance on Raven's matrices across language groups, sessions, and electrode sites. Importantly, the correlations correspond to each language group's sensitivity to the difference responses; that is, native English speakers showed correlations between Raven's and ERP responses for the p^h-deviant condition across the three sessions, but native Spanish speakers showed correlations between Raven's and ERP responses for the b-deviant condition. Native English speakers showed less correlations than native Spanish speakers; only two correlations out of a total of 36 possible correlations were seen in each session for native English speakers and we cannot discount the possibility of spurious relationships. But native Spanish speakers showed six significant correlations in Session 1, before training occurred, suggesting that their baseline perception was most correlated with Raven's, gradually diminishing throughout training. Previous work has shown that there is a positive correlation between performance on Raven's matrices and tone awareness in learning to read Chinese in third, fourth, and fifth gradeage children (Huang & Hanley, 1995; Siok & Fletcher, 2001). However, the authors partial out performance on Raven's to account for effects of IQ and do not further speculate on the mechanisms behind the correlation.

Importantly, there were no significant correlations between performance on Raven's and training, suggesting that a separate mechanism may be behind the relationship between nonspatial intelligence and speech perception. Previous research has examined nonverbal intelligence using Raven's in conjunction with linguistic tasks; Chung, McBride-Chang, Cheung, and Wong (2013) showed that intelligence measures do not have a significant effect on speech perception in English-learning Chinese children, but they used Raven's as a way to ensure that there were no differences between experimental groups and so kept it controlled in regression analyses as opposed to examining it as a theoretically important variable. Other studies have also used Raven's in speech perception and production tasks, but have kept it controlled in analyses or as a way to ensure that there were no differences between groups (e.g., Jakoby et al., 2011; Soroli, Szenkovits, & Ramus, 2010). Although our aim was also to ensure that there were no differences between groups, the significant difference between native English and native Spanish speakers on the Raven's test seems to show important effects on the ERP responses. It is important to note that native English speakers only showed significant effects of Raven's on the perception of the within-category contrast (/p^h/-/p/) across all sessions but native Spanish speakers showed a shift such that better Raven's performance correlated with more negative responses for the /p^h/-/p/ contrasts only in Session 1. After Session 1, more accurate Raven's performance correlated with the between category difference, /b/-/p/. This suggests that the training did, in fact, overemphasize the existing contrasts for native Spanish speakers in general and it possibly influenced participants with higher scores on Raven's more than participants with lower scores on the Raven's test. Future research should examine how nonspatial and nonverbal intelligence measures can affect speech perception, especially at a pre-attentive level.

In conclusion, the present study elaborated on how models of L2 speech perception can use the MMN response as an objective measure of phonetic processing in L2 learners. Importantly, due to the fact that our manipulation with varying standards showed much the same results as the traditional oddball paradigm, future researchers should make sure to examine how varying standards within an oddball paradigm can more effectively mimic the natural variations within the speech signal. Although it is more difficult to control for acoustic effects, employing varying standards illustrates that the MMN is an effective objective measure of phonetic differences between categories. Previous research has controlled for acoustic distance between auditory stimuli to capture the linguistically meaningful phonetic differences; however, it is still a very controlled and unnatural exploration of the phonetic variation within a language. Our study demonstrates that the same type of MMN response can be elicited with natural speech variation, at least for contrasts that vary with aspiration and voicing. This is important for models and theories of L2 speech perception that search for an objective measure of similarity between contrasts. The MMN response, modulated by similarity (Kirmse et al., 2008) but not by natural speech variation within a phonetic category (Best & Tyler, 2007; Winkler et al., 1990, Näätänen, 2001), pre-attentive and objective, represents the objective measure of similarity that models of L2 speech perception have examined with participant ratings.

Furthermore, the present study demonstrated that the MMN response to within-category deviants can change over the course of training and that better performance on training elicits a larger change in the ERP amplitude for the contrasts. Models of L2 speech perception that rely on neural hardwiring for the L1 phonetic categories should ensure that they allow for exposure to the L2 contrasts to affect the neuronal patterns for perceptual categorization, much like the NLM (Kuhl & Rivera-Gaxiola, 2008). The experiment presented here supports the idea that L1 phonetic categories can be shifted with enough training, but the L1 phonetic categories themselves are still the most salient during perception and processing, at least with the current training parameters.

Our results also show that participants performed better on training with feedback, though the type of feedback did not matter. This demonstrates that, in accordance with previous studies, feedback is helpful for phonetic perception, especially when attempting to actively shift or form perceptual categories, but more information is not necessarily better. In language learning classrooms, intensive exposure for a short time to dissimilar L2 speech categories with simple feedback may ensure that learners can start the neuronal re-hardwiring to more effectively acquire the L2 speech categories.

Finally, it is important to note that our study used three language groups to examine the influence of cross-linguistic similarity, improving on studies that have only used two languages to examine how the perceptual categorization of one language affects acquisition of another language. In this case, we were able to highlight how similarity between languages can influence behavioral and neural results, by comparing two languages (English and Spanish) to each other in acquiring phonetic contrasts in another language (Hindi), and also comparing those languages to the learned language, as has been done in previous research.

The present results contribute to our understanding of L2 speech perception both practically, in terms of learning with feedback, and theoretically, to understand how L2 speech categories can be described and measured. Furthermore, we are able to investigate these areas with an objective, independent method that captures the natural variation of speech and the potential differences between languages that influence the acquisition and perception of L2 speech.

APPENDIX A

NORMING STUDY

To ensure that our stimuli had been manipulated correctly before we used them in the training and ERP paradigm, we had ten native English and ten native Spanish speakers listen to our stimuli and categorize them. Native English speakers were recruited from the undergraduate psychology pool and received one credit for their participation. Native Spanish speakers were paid $7/hour for their time. They followed the same procedure as the participants in our training experiment: they were presented with a sound and asked to respond if they believed it began with a /p/ or a /b/ sound. Each participant was presented with 10 instances of each stimulus in a pseudorandom order such that there were no repetitions. Participants showed the appropriate category boundaries for native Spanish and native English speakers, such that native Spanish speakers categorized stimuli from 0 ms VOT and higher as /p/ and native English speakers categorized the stimuli from 20 ms VOT and higher as /p/ (Lisker & Abramson, 1964). This was the same categorization our participants in the critical task showed during the current study pre-training. See Figures 25 and 26 for categorization continuums for these speakers.

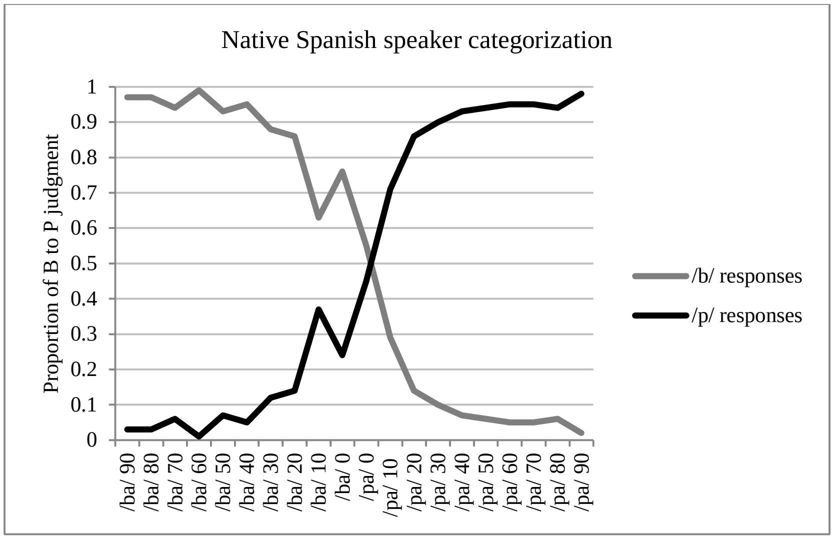

Figure 27. Native Spanish speaker categorization, norming study.

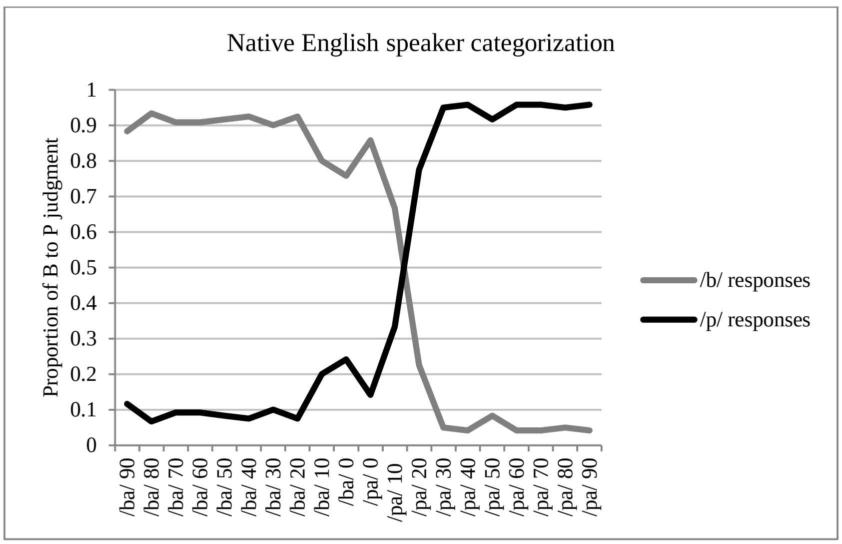

Figure 28. Native English speaker categorization, norming study.

APPENDIX B

CORRELATIONS

Tables 3 and 4 show correlations between training accuracy, individual difference measures and MMN amplitude across all electrode sites for each of the Deviant – Standard responses.

Table 3. Native English speakers, correlations between electrode sites, training, and individual difference measures.

Difference Response by Electrode Site		Individual Difference Measures			
		Training Accuracy	*Operation Span*	*Flankers*	*Ravens*
Session 1					
Diff Stand - Dev 1 Block 1F3	*r*	.14	.30	-.24	.32
	p-value	.57	.22	.32	.18
Diff Stand - Dev 2 Block 1F3	*r*	-.08	.12	-.11	.15
	p-value	.74	.64	.67	.55
Diff Stand - Dev 1 Block 3F3	*r*	-.19	-.21	-.07	.05
	p-value	.44	.38	.79	.85
Diff Stand - Dev 2 Block 3F3	*r*	-.42	-.06	.22	-.17
	p-value	.07	.81	.37	.49
Diff Stand - Dev 1 Block 1FZ	*r*	.20	.40	-.36	.43
	p-value	.42	.09	.14	.07
Diff Stand - Dev 2 Block 1FZ	*r*	-.13	.12	-.25	.20
	p-value	.59	.63	.31	.42
Diff Stand - Dev 1 Block 3FZ	*r*	-.06	.07	-.25	.23
	p-value	.80	.78	.30	.35

Diff Stand - Dev 2 Block 3FZ	*r*	**-.48***	-.13	.08	-.12
	p-value	**.04**	.59	.73	.61
Diff Stand - Dev 1 Block 1F4	*r*	-.05	.21	-.44	.35
	p-value	.84	.38	.06	.15
Diff Stand - Dev 2 Block 1F4	*r*	-.43	.12	-.41	.12
	p-value	.07	.64	.08	.62
Diff Stand - Dev 1 Block 3F4	*r*	.02	.14	-.26	.12
	p-value	.92	.58	.29	.62
Diff Stand - Dev 2 Block 3F4	*r*	**-.57***	-.13	-.04	-.24
	p-value	**.01**	.61	.89	.32
Diff Stand - Dev 1 Block 1FC3	*r*	.23	.33	.00	.23
	p-value	.35	.17	1.00	.35
Diff Stand - Dev 2 Block 1FC3	*r*	.08	.28	-.11	.23
	p-value	.74	.24	.65	.34
Diff Stand - Dev 1 Block 1FCZ	*r*	.19	.28	-.19	.41
	p-value	.43	.25	.43	.09
Diff Stand - Dev 2 Block 1FCZ	*r*	-.13	.19	-.18	.21
	p-value	.59	.43	.46	.40
Diff Stand - Dev 1 Block 3FCZ	*r*	.04	.01	-.23	.19
	p-value	.86	.97	.35	.43
Diff Stand - Dev 2 Block 3FCZ	*r*	-.32	-.17	.08	-.17
	p-value	.18	.50	.75	.49
Diff Stand - Dev 1 Block 1FC4	*r*	.12	.36	-.45	.42
	p-value	.62	.13	.05	.07
Diff Stand - Dev 2 Block 1FC4	*r*	-.22	.19	-.30	.24
	p-value	.38	.44	.21	.32
Diff Stand - Dev 1 Block 3FC4	*r*	.07	.05	-.19	.21
	p-value	.79	.86	.44	.39
Diff Stand - Dev 2 Block 3FC4	*r*	-.37	-.23	.11	-.25
	p-value	.12	.35	.65	.30
Diff Stand - Dev 1 Block 1C3	*r*	.35	.28	-.05	.42
	p-value	.15	.25	.84	.08
Diff Stand - Dev 2 Block 1C3	*r*	-.09	.35	-.24	.23
	p-value	.71	.15	.33	.35
Diff Stand - Dev 1 Block 3C3	*r*	.04	-.27	.07	.04
	p-value	.89	.26	.77	.88
Diff Stand - Dev 2 Block 3C3	*r*	-.13	.01	-.10	-.13
	p-value	.60	.96	.69	.59
Diff Stand - Dev 1 Block 1CZ	*r*	.29	.31	-.21	**.49***
	p-value	.23	.20	.38	**.03**

Diff Stand - Dev 2 Block 1CZ	*r*	-.12	.25	-.17	.10
	p-value	.61	.29	.49	.70
Diff Stand - Dev 1 Block 3CZ	*r*	.06	-.18	-.13	.09
	p-value	.82	.45	.61	.71
Diff Stand - Dev 2 Block 3CZ	*r*	-.13	-.25	.04	-.11
	p-value	.60	.30	.87	.65
Diff Stand - Dev 1 Block 1C4	*r*	.15	.28	-.44	**.51***
	p-value	.54	.24	.06	**.03**
Diff Stand - Dev 2 Block 1C4	*r*	-.35	.12	-.22	.01
	p-value	.14	.63	.37	.99
Diff Stand - Dev 1 Block 3C4	*r*	.16	-.12	-.17	.21
	p-value	.51	.61	.50	.40
Diff Stand - Dev 2 Block 3C4	*r*	-.21	-.20	.02	-.23
	p-value	.40	.41	.94	.34
Session 2					
Diff Stand - Dev 1 Block 1F3	*r*	.06	-.18	-.15	.43
	p-value	.82	.45	.55	.07
Diff Stand - Dev 2 Block 1F3	*r*	-.22	.20	-.15	.18
	p-value	.37	.41	.55	.47
Diff Stand - Dev 1 Block 3F3	*r*	.37	.30	-.20	.06
	p-value	.12	.21	.42	.81
Diff Stand - Dev 2 Block 3F3	*r*	-.14	-.14	-.05	-.26
	p-value	.57	.57	.85	.28
Diff Stand - Dev 1 Block 1FZ	*r*	.18	.00	-.28	.42
	p-value	.47	.99	.26	.07
Diff Stand - Dev 2 Block 1FZ	*r*	-.08	.06	-.35	.34
	p-value	.75	.80	.14	.16
Diff Stand - Dev 1 Block 3FZ	*r*	.28	.17	-.19	.18
	p-value	.25	.48	.43	.46
Diff Stand - Dev 2 Block 3FZ	*r*	-.23	-.20	.03	-.32
	p-value	.34	.42	.91	.19
Diff Stand - Dev 1 Block 1F4	*r*	.12	.11	-.44	.19
	p-value	.62	.65	.06	.45
Diff Stand - Dev 2 Block 1F4	*r*	-.08	.23	-.23	.28
	p-value	.76	.34	.34	.24
Diff Stand - Dev 1 Block 3F4	*r*	.30	.05	-.15	.10
	p-value	.22	.83	.55	.69
Diff Stand - Dev 2 Block 3F4	*r*	-.25	-.23	.09	-.23
	p-value	.30	.35	.73	.35
Diff Stand - Dev 1 Block 1FC3	*r*	.02	-.37	.09	.12

	p-value	.95	.12	.71	.62
Diff Stand - Dev 2 Block 1FC3	*r*	-.09	.20	-.17	.21
	p-value	.70	.41	.48	.39
Diff Stand - Dev 1 Block 3FC3	*r*	.36	.24	.06	.04
	p-value	.13	.33	.82	.88
Diff Stand - Dev 2 Block 3FC3	*r*	-.23	-.26	-.06	-.34
	p-value	.35	.29	.79	.16
Diff Stand - Dev 1 Block 1FCZ	*r*	.14	.04	-.32	**.48***
	p-value	.56	.88	.19	**.04**
Diff Stand - Dev 2 Block 1FCZ	*r*	-.09	.00	-.13	.21
	p-value	.71	1.00	.59	.39
Diff Stand - Dev 1 Block 3FCZ	*r*	.37	.17	-.10	.12
	p-value	.12	.50	.68	.61
Diff Stand - Dev 2 Block 3FCZ	*r*	-.32	-.18	-.08	-.30
	p-value	.18	.47	.76	.22
Diff Stand - Dev 1 Block 1FC4	*r*	.19	-.14	-.33	.24
	p-value	.45	.58	.17	.33
Diff Stand - Dev 2 Block 1FC4	*r*	.04	.07	-.37	.34
	p-value	.87	.76	.11	.16
Diff Stand - Dev 1 Block 3FC4	*r*	.42	.07	.17	.12
	p-value	.07	.77	.49	.64
Diff Stand - Dev 2 Block 3FC4	*r*	-.31	-.20	-.05	-.18
	p-value	.19	.42	.83	.47
Diff Stand - Dev 1 Block 1C3	*r*	.14	-.18	.00	.42
	p-value	.56	.46	1.00	.07
Diff Stand - Dev 2 Block 1C3	*r*	-.01	.04	-.06	.15
	p-value	.97	.87	.82	.54
Diff Stand - Dev 1 Block 3C3	*r*	**.50***	.05	.15	-.07
	p-value	**.03**	.84	.54	.79
Diff Stand - Dev 2 Block 3C3	*r*	-.23	-.24	-.06	-.26
	p-value	.35	.32	.80	.28
Diff Stand - Dev 1 Block 1CZ	*r*	.21	.04	-.30	**.51***
	p-value	.39	.87	.21	**.03**
Diff Stand - Dev 2 Block 1CZ	*r*	.01	.29	-.20	.30
	p-value	.95	.24	.42	.21
Diff Stand - Dev 1 Block 3CZ	*r*	.26	-.07	.15	-.14
	p-value	.29	.78	.55	.57
Diff Stand - Dev 2 Block 3CZ	*r*	-.19	-.14	-.15	-.24
	p-value	.45	.56	.54	.32
Diff Stand - Dev 1 Block 1C4	*r*	.11	-.16	-.35	.22

	p-value	.66	.52	.15	.37
Diff Stand - Dev 2 Block 1C4	*r*	.15	.12	-.29	.29
	p-value	.54	.63	.23	.22
Diff Stand - Dev 1 Block 3C4	*r*	.45	-.08	.29	.00
	p-value	.05	.76	.24	1.00
Diff Stand - Dev 2 Block 3C4	*r*	-.32	-.15	-.06	-.19
	p-value	.18	.54	.81	.45
Session 3					
Diff Stand - Dev 1 Block 1F3	*r*	.32	.04	.07	.32
	p-value	.19	.87	.77	.18
Diff Stand - Dev 2 Block 1F3	*r*	-.13	-.01	-.22	.09
	p-value	.59	.99	.37	.71
Diff Stand - Dev 1 Block 3F3	*r*	.18	.29	-.02	.09
	p-value	.47	.23	.94	.73
Diff Stand - Dev 2 Block 3F3	*r*	-.14	.12	**-.49***	.07
	p-value	.57	.62	**.03**	.78
Diff Stand - Dev 1 Block 1FZ	*r*	.33	-.03	.06	.28
	p-value	.17	.90	.81	.25
Diff Stand - Dev 2 Block 1FZ	*r*	-.14	-.05	-.27	.22
	p-value	.57	.83	.27	.37
Diff Stand - Dev 1 Block 3FZ	*r*	**.52***	.15	.21	.04
	p-value	**.02**	.54	.39	.87
Diff Stand - Dev 2 Block 3FZ	*r*	.06	.05	-.26	.07
	p-value	.80	.84	.28	.77
Diff Stand - Dev 1 Block 1F4	*r*	.16	-.42	.29	.29
	p-value	.53	.07	.23	.23
Diff Stand - Dev 2 Block 1F4	*r*	-.04	.19	-.33	.17
	p-value	.88	.43	.18	.48
Diff Stand - Dev 1 Block 3F4	*r*	.31	.22	.26	-.19
	p-value	.21	.36	.28	.43
Diff Stand - Dev 2 Block 3F4	*r*	.12	.01	-.25	.02
	p-value	.62	.96	.31	.94
Diff Stand - Dev 1 Block 1FC3	*r*	.14	-.07	-.21	**.50***
	p-value	.58	.78	.38	**.03**
Diff Stand - Dev 2 Block 1FC3	*r*	-.08	.04	**-.48***	.24
	p-value	.73	.87	**.04**	.33
Diff Stand - Dev 1 Block 3FC3	*r*	.28	.14	.08	-.04
	p-value	.25	.57	.73	.88
Diff Stand - Dev 2 Block 3FC3	*r*	.00	.06	-.25	.04
	p-value	.99	.81	.30	.88

Diff Stand - Dev 1 Block 1FCZ	*r*	.04	-.26	.05	.32
	p-value	.86	.29	.83	.19
Diff Stand - Dev 2 Block 1FCZ	*r*	-.08	.12	**-.46***	.19
	p-value	.75	.63	**.05**	.45
Diff Stand - Dev 1 Block 3FCZ	*r*	**.46***	.21	.23	.03
	p-value	**.05**	.39	.34	.89
Diff Stand - Dev 2 Block 3FCZ	*r*	.31	.11	-.16	-.06
	p-value	.20	.66	.52	.82
Diff Stand - Dev 1 Block 1FC4	*r*	.20	-.25	-.04	.39
	p-value	.42	.30	.86	.10
Diff Stand - Dev 2 Block 1FC4	*r*	-.09	.06	**-.52***	**.51***
	p-value	.72	.82	**.02**	**.03**
Diff Stand - Dev 1 Block 3FC4	*r*	.25	.12	.25	-.26
	p-value	.31	.63	.31	.29
Diff Stand - Dev 2 Block 3FC4	*r*	.25	.14	-.05	-.12
	p-value	.31	.58	.84	.64
Diff Stand - Dev 1 Block 1C3	*r*	.13	.17	-.02	.31
	p-value	.59	.48	.95	.19
Diff Stand - Dev 2 Block 1C3	*r*	-.11	.09	**-.52***	.34
	p-value	.65	.72	**.02**	.16
Diff Stand - Dev 1 Block 3C3	*r*	.18	.19	-.12	.08
	p-value	.45	.43	.63	.76
Diff Stand - Dev 2 Block 3C3	*r*	.01	.03	-.20	.14
	p-value	.97	.92	.42	.58
Diff Stand - Dev 1 Block 1CZ	*r*	.03	.02	-.10	.31
	p-value	.90	.94	.68	.20
Diff Stand - Dev 2 Block 1CZ	*r*	-.23	.11	**-.57***	.35
	p-value	.34	.64	**.01**	.14
Diff Stand - Dev 1 Block 3CZ	*r*	.31	.07	.11	.09
	p-value	.19	.78	.65	.72
Diff Stand - Dev 2 Block 3CZ	*r*	.14	-.06	-.04	.06
	p-value	.56	.80	.87	.81
Diff Stand - Dev 1 Block 1C4	*r*	.08	-.11	-.10	.33
	p-value	.76	.65	.67	.17
Diff Stand - Dev 2 Block 1C4	*r*	-.23	.12	-.44	.33
	p-value	.34	.62	.06	.17
Diff Stand - Dev 1 Block 3C4	*r*	.29	.16	.22	-.16
	p-value	.24	.51	.36	.52
Diff Stand - Dev 2 Block 3C4	*r*	.34	-.04	.09	-.08
	p-value	.16	.86	.71	.75

Notes. * .05 alpha value ** .01 alpha value

Table 4. Native Spanish speakers, correlations between electrode sites, training, and individual difference measures.

Difference Response by Electrode Site		Individual Difference Measures *Training Accuracy*	*Operation Span*	*Flankers*	*Ravens*
Session 1					
Diff Stand - Dev 1 Block 1F3	*r*	.13	.31	.14	.33
	p-value	.67	.31	.66	.28
Diff Stand - Dev 2 Block 1F3	*r*	.38	.02	.47	-.03
	p-value	.20	.96	.11	.93
Diff Stand - Dev 1 Block 3F3	*r*	**-.57***	-.34	-.15	-.51
	p-value	**.04**	.25	.63	.08
Diff Stand - Dev 2 Block 3F3	*r*	.00	.03	.50	.26
	p-value	1.00	.93	.09	.39
Diff Stand - Dev 1 Block 1FZ	*r*	.03	.31	.00	.07
	p-value	.91	.30	1.00	.81
Diff Stand - Dev 2 Block 1FZ	*r*	.44	-.09	.30	-.37
	p-value	.13	.76	.31	.22
Diff Stand - Dev 1 Block 3FZ	*r*	-.44	-.39	-.29	**-.65***
	p-value	.14	.19	.33	**.02**
Diff Stand - Dev 2 Block 3FZ	*r*	-.09	-.13	.11	.11
	p-value	.76	.67	.71	.72
Diff Stand - Dev 1 Block 1F4	*r*	-.15	.06	.18	-.10
	p-value	.63	.85	.55	.76
Diff Stand - Dev 2 Block 1F4	*r*	.25	-.06	.06	-.35
	p-value	.42	.85	.85	.24
Diff Stand - Dev 1 Block 3F4	*r*	-.27	-.46	-.24	**-.85****
	p-value	.38	.12	.43	**.00**
Diff Stand - Dev 2 Block 3F4	*r*	-.14	-.08	.03	.22
	p-value	.65	.79	.92	.47
Diff Stand - Dev 1 Block 1FC3	*r*	.13	.19	.09	.09
	p-value	.68	.53	.76	.76
Diff Stand - Dev 2 Block 1FC3	*r*	.29	-.10	.39	-.31
	p-value	.34	.76	.20	.30
Diff Stand - Dev 1 Block 1FCZ	*r*	.02	.25	.10	.06
	p-value	.94	.40	.74	.86
Diff Stand - Dev 2 Block 1FCZ	*r*	.52	-.04	.49	-.18
	p-value	.07	.91	.09	.55

Diff Stand - Dev 1 Block 3FCZ	*r*	-.36	-.50	-.20	**-.73****
	p-value	.23	.08	.52	**.01**
Diff Stand - Dev 2 Block 3FCZ	*r*	-.21	-.14	.18	.14
	p-value	.49	.65	.56	.64
Diff Stand - Dev 1 Block 1FC4	*r*	-.11	.20	.01	.12
	p-value	.72	.52	.98	.70
Diff Stand - Dev 2 Block 1FC4	*r*	.40	-.03	.35	-.19
	p-value	.17	.93	.24	.54
Diff Stand - Dev 1 Block 3FC4	*r*	-.26	-.30	-.08	**-.57***
	p-value	.39	.32	.80	**.04**
Diff Stand - Dev 2 Block 3FC4	*r*	-.11	-.14	-.04	.12
	p-value	.72	.64	.90	.69
Diff Stand - Dev 1 Block 1C3	*r*	-.07	.21	.08	.18
	p-value	.83	.50	.79	.57
Diff Stand - Dev 2 Block 1C3	*r*	.44	-.11	.41	-.29
	p-value	.13	.73	.17	.34
Diff Stand - Dev 1 Block 3C3	*r*	**-.61***	-.24	-.27	-.28
	p-value	**.03**	.42	.38	.36
Diff Stand - Dev 2 Block 3C3	*r*	-.01	-.04	.50	.30
	p-value	.98	.89	.08	.32
Diff Stand - Dev 1 Block 1CZ	*r*	-.06	.17	.08	-.03
	p-value	.84	.58	.80	.91
Diff Stand - Dev 2 Block 1CZ	*r*	.44	.21	.31	-.07
	p-value	.13	.48	.31	.83
Diff Stand - Dev 1 Block 3CZ	*r*	-.47	-.42	-.45	**-.56***
	p-value	.10	.15	.13	**.05**
Diff Stand - Dev 2 Block 3CZ	*r*	-.08	.09	.24	.39
	p-value	.79	.77	.43	.19
Diff Stand - Dev 1 Block 1C4	*r*	.06	.03	.05	-.03
	p-value	.85	.92	.88	.92
Diff Stand - Dev 2 Block 1C4	*r*	.22	-.11	.22	-.20
	p-value	.47	.73	.48	.51
Diff Stand - Dev 1 Block 3C4	*r*	-.15	-.23	-.25	**-.57***
	p-value	.63	.44	.40	**.04**
Diff Stand - Dev 2 Block 3C4	*r*	-.09	-.14	.42	.30
	p-value	.77	.66	.15	.32
Session 2					
Diff Stand - Dev 1 Block 1F3	*r*	.34	.39	-.13	-.10
	p-value	.25	.18	.67	.75
Diff Stand - Dev 2 Block 1F3	*r*	-.29	-.43	-.09	-.53

	p-value	.34	.15	.76	.07
Diff Stand - Dev 1 Block 3F3	*r*	.07	-.08	.31	.48
	p-value	.81	.80	.31	.10
Diff Stand - Dev 2 Block 3F3	*r*	-.15	-.26	-.16	.12
	p-value	.62	.39	.61	.69
Diff Stand - Dev 1 Block 1FZ	*r*	.29	.29	-.01	.26
	p-value	.34	.34	.97	.39
Diff Stand - Dev 2 Block 1FZ	*r*	-.07	-.38	-.08	**-.60***
	p-value	.83	.20	.79	**.03**
Diff Stand - Dev 1 Block 3FZ	*r*	-.12	-.06	.25	.26
	p-value	.71	.84	.42	.39
Diff Stand - Dev 2 Block 3FZ	*r*	-.29	-.02	-.14	.37
	p-value	.34	.94	.65	.22
Diff Stand - Dev 1 Block 1F4	*r*	.10	-.18	-.05	-.14
	p-value	.76	.55	.87	.66
Diff Stand - Dev 2 Block 1F4	*r*	.23	-.29	.01	-.43
	p-value	.46	.34	.98	.15
Diff Stand - Dev 1 Block 3F4	*r*	.14	.06	.31	.53
	p-value	.64	.84	.31	.06
Diff Stand - Dev 2 Block 3F4	*r*	-.24	-.16	.00	.24
	p-value	.44	.61	.99	.43
Diff Stand - Dev 1 Block 1FC3	*r*	.05	.33	-.28	-.02
	p-value	.88	.28	.36	.95
Diff Stand - Dev 2 Block 1FC3	*r*	-.12	-.33	-.05	**-.57***
	p-value	.70	.27	.87	**.04**
Diff Stand - Dev 1 Block 3FC3	*r*	.23	-.04	.29	.53
	p-value	.45	.91	.34	.06
Diff Stand - Dev 2 Block 3FC3	*r*	-.15	-.25	-.16	.17
	p-value	.63	.40	.60	.59
Diff Stand - Dev 1 Block 1FCZ	*r*	.26	.08	-.10	-.04
	p-value	.38	.79	.75	.90
Diff Stand - Dev 2 Block 1FCZ	*r*	-.31	-.30	-.13	-.39
	p-value	.31	.32	.66	.19
Diff Stand - Dev 1 Block 3FCZ	*r*	-.01	.01	.25	.45
	p-value	.97	.98	.41	.12
Diff Stand - Dev 2 Block 3FCZ	*r*	-.40	-.03	-.16	.21
	p-value	.17	.92	.60	.49
Diff Stand - Dev 1 Block 1FC4	*r*	-.06	-.07	-.24	-.22
	p-value	.85	.82	.44	.47
Diff Stand - Dev 2 Block 1FC4	*r*	.02	-.20	-.10	-.55

	p-value	.94	.52	.74	.05
Diff Stand - Dev 1 Block 3FC4	*r*	.01	-.03	.26	.52
	p-value	.98	.92	.40	.07
Diff Stand - Dev 2 Block 3FC4	*r*	-.37	-.10	.03	.23
	p-value	.22	.75	.93	.44
Diff Stand - Dev 1 Block 1C3	*r*	.21	.17	-.06	-.13
	p-value	.50	.58	.85	.68
Diff Stand - Dev 2 Block 1C3	*r*	-.11	-.36	-.12	**-.58***
	p-value	.73	.23	.69	**.04**
Diff Stand - Dev 1 Block 3C3	*r*	-.05	-.15	.30	.42
	p-value	.88	.64	.33	.16
Diff Stand - Dev 2 Block 3C3	*r*	-.24	-.13	-.16	.09
	p-value	.43	.67	.60	.76
Diff Stand - Dev 1 Block 1CZ	*r*	.02	-.04	-.32	-.16
	p-value	.95	.91	.29	.60
Diff Stand - Dev 2 Block 1CZ	*r*	-.06	-.32	-.10	-.55
	p-value	.84	.29	.75	.05
Diff Stand - Dev 1 Block 3CZ	*r*	-.25	-.12	.12	.41
	p-value	.42	.69	.69	.17
Diff Stand - Dev 2 Block 3CZ	*r*	-.34	-.09	-.17	.07
	p-value	.25	.78	.57	.82
Diff Stand - Dev 1 Block 1C4	*r*	-.16	-.30	-.17	-.19
	p-value	.61	.33	.57	.53
Diff Stand - Dev 2 Block 1C4	*r*	.10	-.29	.05	-.44
	p-value	.74	.34	.88	.13
Diff Stand - Dev 1 Block 3C4	*r*	-.15	-.08	.21	.49
	p-value	.64	.79	.49	.09
Diff Stand - Dev 2 Block 3C4	*r*	-.36	-.05	-.06	.17
	p-value	.22	.87	.84	.58
Session 3					
Diff Stand - Dev 1 Block 1F3	*r*	-.11	-.15	-.33	.33
	p-value	.72	.62	.27	.28
Diff Stand - Dev 2 Block 1F3	*r*	**-.66***	-.40	-.33	-.42
	p-value	**.02**	.17	.27	.15
Diff Stand - Dev 1 Block 3F3	*r*	.49	-.20	.22	-.15
	p-value	.09	.51	.48	.61
Diff Stand - Dev 2 Block 3F3	*r*	.14	-.12	.44	.50
	p-value	.64	.69	.14	.08
Diff Stand - Dev 1 Block 1FZ	*r*	-.22	-.36	-.10	.21
	p-value	.46	.23	.74	.50

Diff Stand - Dev 2 Block 1FZ	*r*	**-.65***	-.14	-.34	-.32
	p-value	**.02**	.65	.26	.28
Diff Stand - Dev 1 Block 3FZ	*r*	.51	-.21	.11	-.26
	p-value	.07	.50	.73	.39
Diff Stand - Dev 2 Block 3FZ	*r*	-.09	-.16	.35	.09
	p-value	.78	.60	.24	.76
Diff Stand - Dev 1 Block 1F4	*r*	-.15	-.28	-.12	.40
	p-value	.62	.36	.70	.17
Diff Stand - Dev 2 Block 1F4	*r*	**-.58***	-.06	-.37	-.12
	p-value	**.04**	.86	.21	.70
Diff Stand - Dev 1 Block 3F4	*r*	.54	-.17	.05	-.27
	p-value	.06	.58	.89	.38
Diff Stand - Dev 2 Block 3F4	*r*	-.10	-.08	.04	.26
	p-value	.75	.80	.90	.40
Diff Stand - Dev 1 Block 1FC3	*r*	.12	-.07	-.31	.43
	p-value	.70	.82	.30	.15
Diff Stand - Dev 2 Block 1FC3	*r*	-.52	-.32	-.29	-.29
	p-value	.07	.29	.34	.34
Diff Stand - Dev 1 Block 3FC3	*r*	.39	-.04	.12	-.16
	p-value	.19	.91	.70	.60
Diff Stand - Dev 2 Block 3FC3	*r*	-.04	-.28	.41	.25
	p-value	.91	.35	.17	.41
Diff Stand - Dev 1 Block 1FCZ	*r*	-.22	-.49	-.05	-.01
	p-value	.48	.09	.89	.99
Diff Stand - Dev 2 Block 1FCZ	*r*	-.54	-.31	-.17	-.46
	p-value	.06	.31	.59	.11
Diff Stand - Dev 1 Block 3FCZ	*r*	.54	-.07	.10	-.17
	p-value	.06	.83	.76	.57
Diff Stand - Dev 2 Block 3FCZ	*r*	.08	-.30	.49	.01
	p-value	.80	.33	.09	.97
Diff Stand - Dev 1 Block 1FC4	*r*	-.04	-.20	.04	.45
	p-value	.91	.52	.90	.12
Diff Stand - Dev 2 Block 1FC4	*r*	-.32	-.18	-.16	-.22
	p-value	.29	.55	.60	.47
Diff Stand - Dev 1 Block 3FC4	*r*	**.56***	-.06	.07	-.17
	p-value	**.05**	.85	.82	.59
Diff Stand - Dev 2 Block 3FC4	*r*	.01	-.04	.21	.20
	p-value	.98	.90	.50	.52
Diff Stand - Dev 1 Block 1C3	*r*	-.09	-.16	-.24	.30
	p-value	.78	.60	.42	.32

Diff Stand - Dev 2 Block 1C3	*r*	-.34	**-.63***	-.10	**-.73****
	p-value	.25	**.02**	.75	**.01**
Diff Stand - Dev 1 Block 3C3	*r*	.48	-.07	.19	-.10
	p-value	.10	.82	.55	.76
Diff Stand - Dev 2 Block 3C3	*r*	-.03	-.45	**.57***	.06
	p-value	.93	.13	**.04**	.86
Diff Stand - Dev 1 Block 1CZ	*r*	.09	-.30	.02	.29
	p-value	.77	.31	.95	.33
Diff Stand - Dev 2 Block 1CZ	*r*	-.35	-.50	-.01	**-.61***
	p-value	.25	.08	.97	**.03**
Diff Stand - Dev 1 Block 3CZ	*r*	.51	-.05	.03	-.18
	p-value	.08	.87	.93	.56
Diff Stand - Dev 2 Block 3CZ	*r*	.20	-.27	.51	.08
	p-value	.52	.37	.08	.79
Diff Stand - Dev 1 Block 1C4	*r*	-.02	-.29	.10	.30
	p-value	.95	.34	.75	.32
Diff Stand - Dev 2 Block 1C4	*r*	-.38	-.41	-.14	-.50
	p-value	.20	.17	.65	.08
Diff Stand - Dev 1 Block 3C4	*r*	.54	.04	.03	-.16
	p-value	.06	.91	.93	.61
Diff Stand - Dev 2 Block 3C4	*r*	-.02	-.21	.21	.02
	p-value	.96	.48	.50	.95

Notes. * .05 alpha value ** .01 alpha value.

Correlations between the individual difference measures yielded several significant effects. First, there was a significant correlation between operation set size span and Flankers reaction time, r = - .47, p = .003, showing that participants with higher set size had faster reaction times on the Flankers reaction time for correct trials. This suggests that participants with higher working memory scores showed less interference during the Flanker's task. Within the learner groups, native English speakers showed this same significant effect, r = -.59, p = .004, but the native Spanish speakers did not show this, suggesting that the main correlational effect is driven by the native English speakers. There was also a positive correlation between accuracy on the training and accuracy on the Raven's Matrices, $r = .39$, $p = .02$, suggesting that participants with higher accuracy on a nonspatial intelligence test had better accuracy on the training. See Tables 5 and 6 for a summary of correlations between training and individual difference measures for each language group.

Table 5. Native English speaker correlations between training and individual difference measures.

	Individual Difference Measure		
Training	*Operation Span*	*Flanker's*	*Raven's*
Day 1	.07	.13	.10
Day 2	.21	-.12	.33
Day 3	.25	-.12	.45 *
Combined Accuracy	.20	-.01	.33
Individual Difference Measures			
Operation Span	--	--	--
Flanker's	-.59 **	--	--
Raven's	.34	-.39	--

Notes. * .05 significance level ** .01 significance level

Table 6. Native Spanish speaker correlations between training and individual difference measures.

	Individual Difference Measure		
Training	*Operation Span*	*Flanker's*	*Raven's*
Day 1	.09	.28	.47
Day 2	-.58 *	.19	.18
Day 3	.34	.04	-.11
Combined Accuracy	-.05	.26	.29
Individual Difference Measures			
Operation Span	--	--	--
Flanker's	-.22	--	--
Raven's	.45	.10	--

Notes. * .05 significance level

BIBLIOGRAPHY

Apfelbaum, K., Hazeltine, E., & McMurray, B. (2013). Statistical learning in reading: variability in irrelevant letters helps children learn phonics skills. *Developmental Psychology, 49*, 1348-1365.

Best, C. T. (1990). Adult perception of nonnative contrasts differing in assimilation to native phonological categories. *Journal of the Acoustical Society of America, 88*, S177.

Best, C. T. (1991). The emergence of native-language phonological influences in infants: A perceptual assimilation model. *Haskins Laboratories Status Report on Speech Research, SR-107/108*, 1-30.

Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Ed.), Speech perception and linguistic experience: Issues in cross-language research (pp. 171-204). Timonium, MD: York Press.

Best, C. T., McRoberts, G., & Sithole, N. (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, 14*, 345-360.

Best, C. T., Roberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America, 109*, 775-794.

Best, C.T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Haskins Laboratories Status Report on Speech Research, SR-109/110*, 89-108.

Best, C.T., Traill, A., Carter, A., Harrison, K.D., & Faber, A. (2003). !Xóõ click perception by English, Isizulu, and Sesotho listeners. *Proceedings of the 15th International Congress of phonetic Sciences*, Universitat Autònoma de Barcelona, Barcelona, Spain.

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

Boersma, P. & Weenink, D. (2013). Praat: doing phonetics by computer [Computer program]. Version 5.3.60, retrieved 8 December 2013 from http://www.praat.org/

Bradlow, A., Akahane-Yamada, R., Pisoni, D., & Tokhura, Y. (1999). Training Japanese speakers to identify English /r/ and /l/: long-term retention of learning in speech perception and production. *Perception & Psychophysics, 61*, 977-985.

Bradlow, A., Pisoni, D., Akahane-Yamada, R., & Tokhura, Y. (1997). Training the Japanese listener to identify English /r/ and /l/: some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America, 101*, 2299-2310.

Brandmeyer, A., Desain, P., & McQueen, J. (2012). Effects of native language of perceptual sensitivity to phonetic cues. *NeuroReport, 23*, 653-657.

Cenoz, J. (2001). The effect of linguistic distance, L2 status and age on cross-linguistic influence in third language acquisition. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives,* (pp. 8-20): Multilingual Matters, Clevedon.

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*, *1(5):* 351-353.

Chung, K., McBride-Chang, C., Cheung, H., & Wong, S. (2013). General auditory processing, speech perception and phonological awareness skills in Chinese-English biliteracy. *Journal of Research in Reading, 36*, 202-222.

Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport, 8*, 919-924.

DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition,22*, 499-533.

Díaz, B., Baus, C., Escera, C., Costa, A., & Sebastián-Gallés, N. (2008). Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language. *Proceedings of the National Academy of Sciences, 105*, 16083-16088.

Eidsvag, S., Austad, M., Plante, E., & Asbjornsen, A. (2015). Input variability facilitates unguided category learning in adults. *Journal of Speech Language Hearing Research, 58*, 826-839.

Eriksen, B., & Eriksen, C. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics, 16*, 143-149.

Flanagan, J.L. (1972). Speech analysis: synthesis and perception. (2nd Ed.) Oxford, England: Springer-Verlag.

Flege, J.E. (1989). Chinese subjects' perception of the word-final English /t/-/d/ contrast: Performance before and after training. *Journal of the Acoustical Society of America, 86*, 1684-1697.

Flege, J.E. (1995). Second language speech learning: theory, findings, and problems. In W. Strange (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, Timonium, MD: York Press.

Flege, J.E., Munro, M.J., & MacKay, I. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, 97*, 3125-3134.

Flege, J.E., Yeni-Komshian, G., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, *41*, 78-104.

Francis, A.L., & Nusbaum, H. (2002). Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance, 28*, 349-366.

Gerrits, E., & Schouten, M. (2005). Categorical perception depends on the discrimination task. *Perception & Psychophysics, 66*, 363-376.

Gómez, R. (2002). Variability and detection of invariant structure. *Psychological Science, 13*, 431-436.

Hammarberg, B. (2001). Roles of L1 and L2 in L3 production and acquisition. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives*, (pp. 21-41): Multilingual Matters, Clevedon.

Huang, H. S., & Hanley, J. R. (1995). Phonological awareness and visual skills in learning to read Chinese and English. *Cognition*, *54*(1), 73-98.

Jakoby, H., Goldstein, A., & Faust, M. (2011). Electrophysiological correlates of speech perception mechanisms and individual differences in second language attainment. *Psychophysiology*, 48, 1516-1530.

Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*, 60-99.

Keil, A., Debener, S., Gratton, G., Junghöfer, M., Kappenman, E., Luck, S., Luu, P., Miller, G., & Yee, C. (2014). Committee Report: Publication guidelines and recommendation for studies using electroencephalography and magnetoencephalography. *Psychophysiology, 51*, 1-21.

Kirmse, U., Ylinen, S., Tervaniemi, M., Vainio, M., Schröger, E., & Jacobsen, T. (2008). Modulation of mismatch negativity (MMN) to vowel duration changes in native speakers of Finnish and German as a result of language experience. *International Journal of Psychophysiology, 67*, 131-143.

Kraus, N., McGee, T., Carrell, T., King, C., Tremblay, K., & Nicol, T. (1995). Central auditory system plasticity associated with speech discrimination training. *Journal of Cognitive Neuroscience, 7*, 25-32.

Kuhl, P. (1993). Innate predispositions and the effects of experience in speech perception: the native language magnet theory. In B. Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, & J. Morton (Eds.), *Developmental Neurocognition: Speech and Face Processing in the First Year of Life.* (pp. 249-274): Springer Netherlands.

Kuhl, P., & Rivera-Gaxiola, M. (2008). Neural substrates of language acquisition. *Annual Review Neuroscience, 31*, 511-534. doi:10.1146/annurev.neuro.30.051606.094321

Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255*, 606-608.

LaPointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 1118–1133.

Lenneberg, E. (1967). *Biological Foundations of Language.* New York: Wiley.

Lisker, L. (1985). How is the aspiration of English /p,t,k/ predictable? *Language & Speech, 27*, 391-394.

Lisker, L., & Abramson, A. S. (1967). Some effects of context on voice onset time in English stops. *Language and Speech*, *10*, 1-28.

Liu, R., & Holt, L. (2011). Neural changes associated with nonspeech auditory category learning parallel those of speech category acquisition. *Journal of Cognitive Neuroscience*, 23, 683-698.

Logan, J., Lively, S., & Pisoni, D. (1991). Training Japanese listeners to identify /r/ and /l/: a first report. *Journal of the Acoustical Society of America, 89*, 874-886.

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

MacKain, K., Best, C., & Strange, W. (1981). Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics, 2*, 269-290.

MacWhinney, B. (2005). A unified model of language acquisition. In J.F. Kroll & A.M.B DeGroot (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 49-67): Oxford University Press.

MacWhinney, B. (2012). The logic of the unified model. In S. Gass & A. Mackey (Eds.), *Handbook of Second Language Acquisition* (pp. 211-227): Routledge.

MacWhinney, B. & Pleh, C. (1988). The processing of restrictive relative clauses in Hungarian. *Cognition*, *29*, 95-141.

McCandliss, B., Fiez, J., Protopapas, A., Conway, M., & McClelland, J. (2002). Success and failure in teaching the [r]-[l] contrast to Japanese adults: tests of a Hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective, & Behavioral Neuroscience*, *2*, 89-108.

McClelland, J., Fiez, J., & McCandliss, B. (2002). Teaching the /r-l/ discrimination to Japanese adults: behavioral and neural aspects. *physiology & Behavior, 77,* 657-662.

Mota, M.B. (2003). Working memory capacity and fluency, accuracy, complexity, and lexical density in L2 speech production. *Fragmentos: Revista de Língua e Literatura Estrangeiras*, *24*, 69-104.

Mueller, J. (2005). Electrophysiological correlates of second language processing. *Second Language Research*, *21*, 152-174.

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

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R.J., Luuk, A., Allik, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, *385*, 432-434.

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

Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*, *125*, 826-859.

Nenonen, S., Shestakova, A., Huotilainen, M., & Näätänen, R. (2003). Linguistic relevance of duration within the native language determines the accuracy of speech-sound duration processing. *Cognitive Brain Research, 16,* 492-495.

Nenonen, S., Shestakova, A., Huotilainen, M., & Näätänen, R. (2005). Speech-sound duration processing in a second language is specific to phonetic categories. *Brain and Language, 92,* 26-32.

Newport, E. (1990). Maturational constraints on language learning. *Cognitive Science, 14*, 11-28.

Norris, J.M. & Ortega, L. (2000). Effectiveness of L2 instruction: a research synthesis and quantitative meta-analysis. *Language Learning, 50*, 417-528.

Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & Psychophysics*, *60*(3), 355-376.

Ortiz-Mantilla, S., Choudhury, N., Alvarez, B., & Benasich, A. (2010). Involuntary switching of attention mediates differences in event-related responses to complex tones between early and late Spanish-English bilinguals. *Brain Research*, 1362, 78-92.

Oyama, S. (1975). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research, 5*, 261-283.

Pallier, C., Bosch, L., & Sebastian-Gallés, N. (1997a). A limit on behavioral plasticity in vowel acquisition. Cognition, 64, B9- B17.

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*, 25-28.

Perry, L., Samuelson, L., Malloy, L., & Schiffer, R. (2010). Learn locally, think globally. Exemplar variability supports higher-order generalization and word learning. *Psychological Science, 21*, 1894-1902.

Piske, T., MacKay, I., & Flege, J.M. (2001). Factors affecting degree of foreign accent in an L2: a review. *Journal of phonetics, 29,* 191-215.

Pisoni, D. (1993). Long-term memory in speech perception: Some new findings on talker variability, speaking rate and perceptual learning. *Speech Communications, 13,* 109-125.

Polka, L. (1995). Linguistic influences in adult perception of nonnative vowel contrasts. *Journal of the Acoustical Society of America, 97*, 1286-1296.

Raven, J. (1981). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Research supplement no. 1: The 1979 British Standardisation of the Standard Progressive Matrices and Mill Hill Vocabulary Scales, together with comparative data from earlier studies in the UK, US, Canada, Germany, and Ireland.* Oxford, England: Oxford Psychologists Press/San Antonio, TX: The Psychological Corporation.

Ringbom, H. (1987). *The role of the first language in foreign language learning*. Clevedon: Multilingual Matters

Rivera-Gaxiola, M., Csibra, G., Johnson, M., & Karmiloff-Smith, A. (2000). Electrophysiological correlates of cross-linguistic speech perception in native English speakers. *Behavioural Brain Research, 111*, 13-23.

Savela, J., Kujala, T., Tuomainen, J., Ek, M., Aaltonen, O., & Näätänen, R. (2003). The mismatch negativity and reaction time as indices of the perceptual distance between the corresponding vowels of two related languages. *Cognitive Brain Research, 16*, 250-256.

Sams, M., Alho, R., & Näätänen, R. (1984). Short-term habituation and dishabituation of the mismatch negativity of the ERP. *Psychophysiology, 21*, 434-441.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129-158.

Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics, 13,* 206-226.

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

Seliger, H., S. Krashen and P. Ladefoged. 1975. Maturational constraints in the acquisition of second language accent. Language Sciences 36: 20-22.

Sharma, A. & Dorman, M.F. (2000). Neurophysiologic correlates of cross-language phonetic perception. *Journal of the Acoustical Society of America, 107*, 2697-2703.

Siok, W., & Fletcher, P. (2001). The role of phonological awareness and visual-orthographic skills in Chinese reading acquisition. *Developmental psychology*, *37*(6), 886.

Soroli, E., Szenkovits, G., & Ramus, F. (2010). Exploring dyslexics' phonological deficit III: foreign speech perception and production. *Dyslexia*, *16*, 318-340.

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

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643-662.

Tahta, S., Wood, S., & Loewenthal, K. (1981). Age changes in the ability to replicate foreign pronunciation and intonation. *Language and Speech, 24,* 363-372.

Tees, R., & Werker, J. (1984). Perceptual flexibility: Maintenance or recovery of the ability to discriminate non-native speech sounds. *Canadian Journal of Psychology, 38*, 579-590.

Tolentino, L., & Tokowicz, N. (2014). Cross-language similarity modulates effectiveness of second language grammar instruction. *Language Learning, 64,* 279-309.

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. Studies in Second Language Acquisition, 27, 173-204.

Tokowicz, N., Michael, E. B., & Kroll, J. F. (2004). The roles of study-abroad experience and working-memory capacity in the types of errors made during translation. *Bilingualism: Language and Cognition, 7*, 255–272.

Tremblay, K., Kraus, N., Carrell, T., & McGee, T. (1997). Central auditory system plasticity: generalization to novel stimuli following listening training. *Journal of the Acoustical Society of America*, *102*, 3762-3773.

Tremblay, K., Kraus, N., & McGee, T. (1998). The time course of auditory perceptual learning: neurophysiological changes during speech-sound training. *NeuroReport*, *9*, 3557-3560.

Tricomi, E., Delgado, M., McCandliss, B., McClelland, J., & Fiez, J. (2006). Performance feedback drives caudate activation in a phonological learning task. *Journal of Cognitive Neuroscience, 18*, 1029-1043.

Weissheimer, J. & Mota, M. (2009). Individual differences in working memory capacity and the development of L2 speech production. *Issues in Applied Linguistics, 17*, 94-112.

Werker, J. (1989). Becoming a native listener. *American Scientist*, *77*, 54-59.

Werker, J. & Tees, R.C. (1999). Influences on infant speech perception: Toward a new synthesis. *Annual Review of Psychology, 50*, 509-535.

Werker, J., & Logan, J.S. (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics, 37*, 35-44.

Werker, J. & Tees, R.C. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7,* 49-63.

Winkler, I., Paavilainen, P., Alho, K., Reinikainen, K., Sams, M., & Näätänen, R. (1990). The effect of small variation of the frequent auditory stimulus on the event-related brain potential to the infrequent stimulus. *Psychophysiology, 27*, 228-235.

Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., Cziglar, I., Csepe, V., Ilmoniemi, R.J., & Näätänen, R. (1999). Brain responses reveal the learning of foreign language phonemes. *Psychophysiology, 36*, 638-642.

Ylinen, S., Uther, M., Latvala, A., Vepsäläinen, S., Iverson, P., Akahane-Yamada, R., & Näätänen, R. (2010). Training the brain to weight speech cues differently: A study of Finnish second-language users of English. *Journal of Cognitive Neuroscience*, *22,* 1319-1332.