NEUROPHYSIOLOGICAL INDICES OF LEXICAL TONE PROCESSING IN CHILDREN WITH DIFFERENT LANGUAGE BACKGROUND

A thesis submitted in partial fulfillment
of requirements for the degree of

MASTER OF SCIENCE

to the faculty of the

DEPARTMENT OF COMMUNICATION SCIENCES AND DISORDERS

of

ST. JOHN'S COLLEGE OF LIBERAL ARTS AND SCIENCES
at
ST. JOHN'S UNIVERSITY
New York
by
Andres Felipe Diaz

Date Submitted May 6, 2025

Date Approved May 7, 2025

____
Andres Felipe Diaz

____
Dr. Yan H. Yu

ABSTRACT

NEUROPHYSIOLOGICAL INDICES OF LEXICAL TONE PROCESSING IN CHILDREN WITH DIFFERENT LANGUAGE BACKGROUND

Andres Felipe Diaz

The development of lexical tone processing in children is shaped by both language experience and acoustic salience. While mismatch responses (MMRs) and late negativity (LN) components of event-related potentials (ERPs) have revealed early sensitivity to lexical tone in monolingual children, the timeline for achieving adult-like neural responses—particularly for acoustically distinct versus subtle tone contrasts—remains unclear. Moreover, little is known about how bilingual language experience, especially with tonal versus non-tonal home languages, modulates this neural processing.

This study employed a passive multi-oddball paradigm using Mandarin Tone 3 (low dipping) as the standard, and Tone 1 (high level) and Tone 2 (rising) as deviants, to examine ERP responses in bilingual Mandarin-English, bilingual Spanish-English, and monolingual English-speaking children aged 5 to 10. Between 100–300 ms, bilingual Mandarin-English children exhibited significantly larger mismatch negativity (MMN) amplitudes at frontal and midline sites (F3, Fz, C3, Cz), indicating enhanced early auditory discrimination shaped by tonal language exposure. Between 300–500 ms, LN responses emerged for both Tone 1 and Tone 2 in the bilingual Mandarin-English group, most robustly at F3 and Fz, while the English and Spanish groups showed LN primarily to the more acoustically salient Tone 1. In contrast, the bilingual Spanish-English group

exhibited a robust LN response F4 and C3 for Tone 1, suggesting differential engagement of attentional or cognitive mechanisms across groups.

Together, these findings highlight both a bilingual language effect and a home language effect, highlighting how early language exposure differentially shapes the neurophysiological mechanisms underlying lexical tone processing during childhood.

DEDICATION

To my Isabella,

for seeing the vision before I did—

and standing with me as I shaped it.

ACKNOWLEDGEMENTS

I would like to extend my deepest gratitude to Dr. Yan Helen Yu for her unwavering support, generosity, and kindness in allowing me to incorporate elements of her NIH-R16 research into this thesis. Her mentorship has been a guiding light throughout this process.

I am also sincerely thankful to Dr. Tyreek Jackson and Dr. Monica Wagner for their valuable contributions to advancing the field.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: LITERATURE REVIEW

1. Language Experience and Speech Processing

Language experience profoundly influences our capacity to interpret and engage with others, forming a dynamic conduit between communication and cognition. The framework of communication and cognition depends on the usage of distinct speech phonemes and the processes that organize those sounds and semantics within the auditory cortex, speech motor cortex, and associated neural networks. At the macro-neuroanatomy level, speech perception begins in the auditory canals and proceeds to the auditory cortex within the temporal lobe. Neuroimaging research (Plakke & Romanski, 2014) has shown that phonemic discrimination activates both dorsal and ventral auditory streams, ultimately engaging neurons in the prefrontal and frontal cortex. These pathways support auditory spatial localization and pitch-based processing, underscoring that speech perception is not confined to sensory input but involves integrative activity between temporal and frontal cortical regions. Thus, speech processing is inherently a function of broader cognitive architecture, especially at the prefrontal and frontal cortices.

A defining characteristic among languages lies in how individuals apply this linguistic framework to convey meaning. Due to the ubiquitous nature of non-tonal languages, such as English and Spanish, contemporary neurophysiological research on speech perception has largely focused on the processing of vowels and consonants (Carreiras & Price, 2008; Shi et al., 2024). However, comparatively limited studies have examined the neural mechanisms underlying lexical tone processing (Kaan et al., 2008; Yu et al., 2017), where tonal languages, such as Mandarin Chinese, rely on specific pitch patterns within syllables to differentiate lexical meaning, a feature similarly observed in languages such as Vietnamese (Chen et al., 2023) and various African linguistic systems (Myers, 2021).

Early in development, infants and children attune to the frequencies, sounds, and patterns of languages they are exposed to. In the modern day, these individuals are increasingly exposed to a diverse range of language backgrounds. This exposure prompts important questions about how this linguistic diversity shapes their phonological processing, and whether this exposure affects their ability to discriminate between linguistically meaningful sound contrasts. Specifically, the extent of a child's language background–whether they developed hearing tonal contrasts or not–may significantly influence their functional framework of auditory processing networks. The differential effects in neural processing of tonal and non-tonal language exposure – and from bilingualism – may provide insights into how tone is encoded in the brain through abstract representations shaped by an individual's unique acoustic experiences.

1.1 Consonants and Vowels

Consonants and vowels are fundamental units in language structure with unique articulatory features (Ladefoged & Johnson, 2011) that are present in a myriad of languages. Research has shown that the segmental features for consonants tend to be perceived more categorically, as listeners schematically group specific sounds to words, while vowels are represented more continuously by directly distinguishing between acoustic differences (Phillips, 2001). These features are often examined for their neurophysiological differences. Cross-linguistic studies using mismatch negativity (MMN) measures have shown that the brain's response to phonemic contrasts—whether consonants (Shafer et al., 2004) or vowels (Näätänen et al., 1997; Yu et al., 2019)—can indicate the efficiency of phonological processing. Magnetoencephalography studies (Altmann et al., 2014) have shown stronger repetition suppression effects for both vowels and consonants in the left superior temporal sulcus and gyrus regions associated with phoneme discrimination. This finding supports the notion of left-hemispheric dominance for language processing, in contrast to the right hemisphere's specialization for prosodic or pitch-related information. Therefore, the perceptual salience of phonemic contrast, especially between native and non-native speakers, is influenced by the magnitude of acoustic variation and prior language experience (Kirmse et al., 2008).

Additionally, although consonants and vowels are traditionally considered the fundamental segmental units of speech, pitch variations—resulting from changes in fundamental frequency (F0)—introduce another dimension of linguistic distinction. For example, pitch is an intrinsic feature of syllables (Ladefoged & Johnson, 2010), and lexical tones represent phonemic contrasts based on F0 differences (Yip, 2002). Yu et al. (2017) demonstrated that Mandarin Chinese native listeners exhibited larger MMN responses during lexical tone processing with a more categorical perception. However, discrimination tests on Thai lexical tones (Abramson, 1977) revealed high perceptual sensitivity but a linear, non-categorical perception pattern, more similar to vowel discrimination. Taken together, these findings emphasize the complexity of tone perception associated with pitch across languages with different phonological compositions, warranting further investigation into the neurophysiological mechanisms underlying lexical tone processing and auditory neuroplasticity.

1.2 Lexical Tone

Lexical tone, a linguistic term, refers to variations in fundamental frequency (F0) that convey lexical or grammatical distinction (Yu et al., 2017). These pitch variations arise from F0 patterns generated during the vocal fold vibrations of speech production (Yan et al., 2017; Yip, 2002). Lexical tone differs from consonants and vowels in both form and function. While consonants and vowels are segmental speech units, lexical tone functions as a non-segmental feature superimposed on the vowel, and as a segmental feature that differentiates word meanings (Burnham, 1986).

Yu et al. (2017) investigated how the brain processes lexical tones by measuring neural responses to tonal contrasts in adults. The researchers showed that language experience enhances the discrimination of speech contrasts at both behavioral and pre-attentive levels, as evidenced by MMN waveforms. Native speakers draw on their linguistic experience to rapidly and accurately discern lexical meanings based on tonal patterns. In contrast, non-native speakers or late second-language learners frequently demonstrate limited perceptual discrimination and identification capabilities (Yu et al., 2017), leading to inaccurate lexical representations (Kaan et al., 2008). These responses often provide insights into the cognitive load and the temporal dynamics (e.g., latency and amplitude) of accessing these phonetic details during speech production (Strange, 2011). A higher amplitude and faster latency are indicative of reduced cognitive effort in retrieving articulation discrimination for MMN. Research on native Mandarin listeners has revealed a consistent pattern in how the brain processes tonal variations. Studies have found that when listeners encounter different tonal categories (between-category differences; native), their brains elicit more robust MMN amplitudes than those within the same tonal category (within-category differences; non-native) (Rivera-Gaxiola et al., 2000a,b; Yu et al., 2017). This effect is particularly pronounced when pitch serves a phonemic function (Ren et al., 2009). In other words, variations in pitch change the meaning of a word, as observed in tonal languages like Mandarin. For example, in Mandarin Chinese, the syllable "ma" can mean "mother" (mā, high-level tone), "hemp" (má, rising tone), "horse" (mǎ, low dipping tone), or "scold" (mà, falling tone), depending solely on pitch contour (Lu et al., 2023). Xi et al. (2010) further strengthened these findings, confirming that between-category tonal contrasts consistently elicit more robust MMN responses. These findings collectively indicate that the brain processes linguistically relevant tonal distinctions more prominently than acoustic variations within the same tonal category.

Mandarin Chinese provides an ideal case study to examine these aspects of lexical tone processing. As a tonal language, it demonstrates how lexical tone combines with other phonetic elements, including vowels, consonants, duration, and intensity, to signal meaning distinctions (Yu et al. 2017). There are five primary lexical tones of Mandarin: Tone 1 (high level), Tone 2 (rising), Tone 3 (low dipping), Tone 4 (falling), and Tone 5 (neutral). Within these tone categories, any type of pitch variation can dramatically change the meaning of a word (Duanma, 2007; Tan et al., 2016). Howie (1976) breaks down these contour differences even further: as shown in Figure 1*,* Tone 1 (T1) has a high and then level F0 contour; Tone 2 (T2) begins with a dip and rises approximately 20% into the vowel duration; Tone 3 (T3) has a dipping start, and then rising F0 contour at 50% syllable pronunciation; and Tone 4 (T4) has a falling F0 contour. Tone 5 (T5) was omitted from the F0 contour in Figure 1 due to its inherently variable nature and lack of a stable pitch target. As delineated by Lee (2003), T5 does not carry a lexical tone contour of its own; rather, its pitch is determined by the preceding tone and prosodic context, making it difficult to represent with a consistent F0 pattern.

In Mandarin, T2 and T3 have very close onset F0 and diverge only after 20% into the syllable, meanwhile T1 differentiates dramatically from T2 and even further from T3. Thus, this formulation yields two distinct salience categories: a hard contrast between Tone 2 and Tone 3, and an easy contrast between Tone 1 and Tone 3. Native listeners process the complete pitch pattern (F0 contour) gradually over time, while non-native listeners rely mainly on the simpler acoustic features such as the pitch onset, pitch offset, or the average pitch level when distinguishing between tones (Gandour & Harshman, 1978). In other words, as neurophysiological indices demonstrate that both native and non-native listeners can detect pitch differences, only individuals with tonal language experience can automatically interpret these acoustic variations as meaningful linguistic distinctions. Alternatively, non-natives may perceive acoustic distinctions without associating them with lexical meaning. This difference in processing reflects the fundamental role of language experience in shaping tone perception.

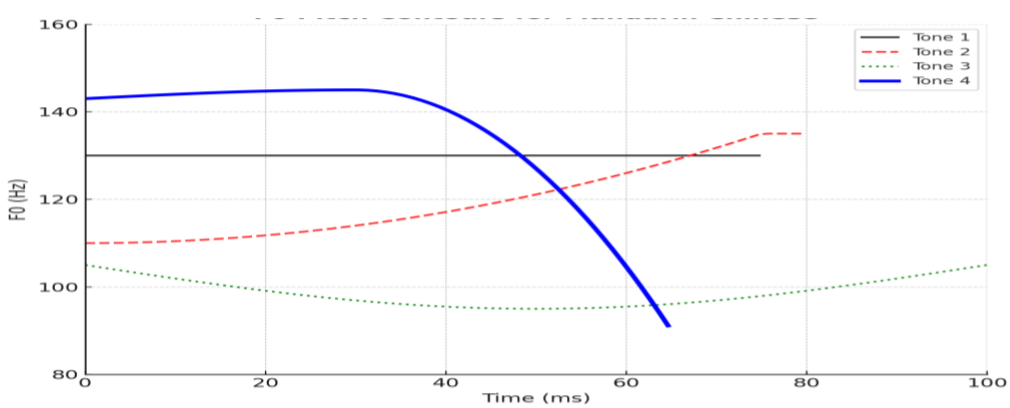

Figure 1: F0 Contour of the Four Mandarin Lexical Tones.

2. Introduction to Neurophysiological Measures and Techniques of Speech Processing

Speech perception plays a crucial role in the early development of language processing capacities. For example, behavioral studies using visually reinforced infant speech discrimination paradigms have demonstrated that infants undergo perceptual reorganization to the phonemic contrasts of their native language within the first year of life (Werker & Tees, 1984). This early tuning allows native phonemic contrasts to be perceived more sharply due to habitual exposure, by categorizing specific speech sounds. Meanwhile, the automatic selective perceptual model has shown through bottom-up process that non-native contrasts become harder to distinguish over time, particularly in late second-language learners (Strange, 2011). Although behavioral paradigms—such as habituation and discrimination tasks—have contributed greatly to understanding the development of phoneme processing, they primarily capture overt behavior, such as conscious recognition and decision-making stages. As a result, these approaches fail to capture the pre-attentive neural mechanisms involved in automatic speech discrimination. They cannot account for the subtle, early cortical responses that occur without conscious awareness, limiting our understanding of the brain's full capacity to detect and process lexical acoustic contrasts.

To address these limitations, modern neurophysiological measures, such as electroencephalography (EEG)—objective and non-invasive tools—have become essential in clinical research and cognitive neuroscience research. EEG measures cortical auditory response with precise temporal but poor spatial activity. When paired with paradigms like the oddball task (*e.g.*, a rare change in a sound pattern), EEG can provide insights into the neural dynamics of auditory representation and tone perception.

Researchers use EEG to capture these cortical auditory responses (Näätänen et al., 2019) through event-related potentials (ERP), which are measurable voltage changes that reflect time-locked neural activity to specific auditory stimuli across subjects. The generated ERP waveforms are created by collecting electroencephalograms (EEGs) from the scalp. When segmented into multiple epochs and averaged ERPs are viewed as waveforms, a series of positive and negative maximum peaks that vary over time in polarity, amplitude, and duration are evoked. EEG records the fluctuations in voltage that constitute ERP waveforms with millisecond-level temporal precision, as electrical potentials propagate through the brain tissue, meninges, skull, and scalp (Kappenman & Luck, 2011).

Scalp-recorded voltage fluctuations underlying ERP waveforms primarily reflect the overall activity of postsynaptic potentials (PSPs) generated by large populations of cortical pyramidal neurons (Kappenman & Luck, 2011). These PSPs occur when large groups of pyramidal neurons are activated simultaneously, creating electrical potentials that produce measurable dipole activity. When these neurons are spatially aligned perpendicular to the cortical surface, they evoke equivalent current dipoles—electrical components that reflect both prior and ongoing mental processes (Kappenman & Luck, 2011)—which are either positive or negative.

Two key negative ERP components arise—mismatch negativity (MMN) and late negativity (LN)—that allow researchers to monitor, for example, phonological differences across languages by indexing automatic auditory discriminations and later cognitive processing demands in speech perception. Several studies have shown that these frontocentral regions (*e.g.,* F3, Fz, F4, C3, Cz, C4) exhibit the largest amplitudes in response to deviant auditory stimuli (Wang et al., 2018; Lindín et al., 2013). These electrodes site are vital for measuring MMN because they capture the brain's response to auditory changes in different cortical regions, providing valuable insights into auditory processing and change detection. For instance, Wang et al. (2018) demonstrated, through scalp topography, a clear evolving negative voltage concentrated over the frontocentral electrodes in response to deviant stimuli. Similarly, Yu et al. (2018) analyzed LN waveforms at the same frontocentral electrodes and recorded strong manifestations of negative voltage in these regions. This pattern further suggests involvement of frontal executive processes, as sustained activity is observed in frontal regions during the evaluation the deviant stimuli, possibly emphasizing access to working memory (Wang et al., 2018). Together, these findings suggest that both MMN and LN reflect not only early sensory discrimination but also higher-order and cognitive processes, extending beyond the auditory cortex into prefrontal and frontal regions.

2.1 Mismatch Responses

Mismatch responses (MMRs) are components of ERPs that reflect the brain’s ability to detects violations in predicted stimulus patterns. Unlike obligatory P1-N1-P2 auditory evoked potentials (AEPs), which indexes the physical properties of a stimulus, MMRs are elicited specifically by contrasts between standard and deviant stimuli (Näätänen et al., 2011). These obligatory AEPs represent early sensory processing in the primary and secondary auditory cortices of the temporal lobe and provide a functional index of neural responsiveness to basic acoustic features. In contrast, the robustness of the MMRs emerge when an incoming sound deviates from an expected pattern, thereby revealing the brain’s capacity for automatic auditory discrimination. The characteristics of the MMR are influenced by several factors, such as the frequency of standard stimulus repetitions, the acoustic distance between the standard and deviant contrasts, their linguistic relevance, and the interstimulus interval (ISI) at which stimuli are presented (Näätänen et al., 2011; Yu et al., 2017). While MMR is often used interchangeably with MMN, MMR is a broader term that encompasses both negative and positive deflections. These components may differ in polarity, amplitude, latency, and scalp distribution depending on variables such as age and gender. For example, Yu et al. (2019) observed a polarity shift in MMRs from robust positive responses in younger children to more adult-like negative responses (nMMRs) as age increased. This shift may reflect enhanced automaticity or increased sensitivity to phonemic differences. Meanwhile, MMN specifically refers to the early negative-going waveforms typically observed in adults.

A typical MMN paradigm presents a frequent standard tone with a repetitive 1000 Hz tone of 50 ms duration, presented at a 1-second ISI, interspersed with infrequent deviant tones of slightly higher frequency while participants passively listen to the stimulus (Näätänen et al., 2019). Both the standard and deviant stimuli elicit the P1-N1-P2 waveform, but MMN reflects the difference between the standard and the deviant—a negative deflection observed in the difference waveform, typically within the 100-300 ms window after the stimulus onset. In children, the MMN peak has been shown to occur at a later latency of 200-400 ms (Datta et al., 2010). The MMN reflects the brain's detection of the deviation from the expected auditory pattern, derived by subtracting the ERPs of the standard stimuli from those of the deviant stimuli across subjects. In the auditory linguistic literature, MMR amplitude is modulated by the linguistic experience of the contrast: phonemic differences elicit early, larger, and more robust MMRs in native speakers, whereas non-phonemic contrasts may yield weaker responses or delayed latencies (Näätänen et al., 1997). In developmental populations, such as infants and young children, since MMN tends to have a broader time window, it may appear with different topography, indicative of maturational differences in auditory cortical structures (Shafer et al., 2000, Yu et al., 2019). The use of MMR waveform provides the platform to understand the brain's automatic sensory mechanism in the cortical region and its ability to continuously update abstract auditory representations towards unexpected sensory cues.

2.2 The Late Negativity Component

The frontocentral late negativity (LN) frequently follows MMN responses in ERP-passive oddball paradigms. This later waveform component typically emerges approximately between 300-500+ ms post-stimulus, and is believed to reflect increased cognitive effort required to accurately differentiate between auditory stimuli. Some researchers refer to the LN as the "second MMN" or "MMN2", due to its potential role in higher-order semantic or phonological processing of the deviant stimuli (see Yu et al., 2017). This is particularly evident in children, as it reflects further processing of deviant sounds due to their limited experiences of language distinctions. Such processing includes attention reorientation, evaluation of the deviant stimulus, and updating memory representations (Korpilahti et al., 2001; Yu et al., 2018). In this context, when a deviant stimulus captures attention, the brain must subsequently redirect focus back to the ongoing task. This is consistent with the broader view that frontocentral negativities represent executive control processes in the frontal cortex that regulate attention to stimuli (Wang et al., 2018).

Further evidence suggests that the LN can also reflect the linguistic and cognitive processing of the auditory stimuli to individuals with multiple language backgrounds. For instance, Yu et al. (2018) found that the LN is larger for deviances in non-native listeners compared to native Mandarin speakers. Additionally, LN responses become more pronounced in bilingual individuals when processing speech contrasts from non-native languages (Ortiz-Mantilla et al., 2010), suggesting that the LN may index the extra cognitive effort or attention required for less familiar phonetic contrasts. These implications suggest that LN amplitudes does offer valuable insights into the efficiency and cognitive demands of auditory discrimination across different linguistic backgrounds. However, serving as an additional index, recent studies have elucidated that the LN can indicate that discrimination has occurred even in the absence of MMN in listeners with weaker phonological skills, challenging the 'second MMN' theory. Notably, Shafer et al. (2005) reported comparable LN amplitudes between children with specific language impairments (SLI) and their typically developing peers. However, interpretations of LN in clinical populations remain mixed, and findings across studies suggest variability depending on the population and the nature of the speech contrasts examined. Thus, LN serves as a valuable neurophysiological marker for understanding how tone perception is processed in children with different language backgrounds.

2.3 The Multi-Oddball Paradigm

Incorporating within-category variation through multiple tokens of both standard and deviant stimuli can elicit distinct MMR and LN response patterns compared to using a single token (Yu et al., 2017). Single-token paradigms tend to promote low-level acoustic discrimination based on surface features like pitch or intensity (Hestvik &

Durvasula, 2016). In contrast, multi-token paradigms introduce natural variability across tokens, such as different speakers or slight acoustic deviations, forcing the brain to abstract away from superficial differences to detect underlying phonemic categories. This requirement for abstraction is crucial, as it more closely mirrors real-world speech perception, where listeners must identify phonological categories across highly variable acoustic input (Yu et al., 2017). Therefore, multi-token designs better engage phonological-level processing and reflect a more ecologically valid measure of linguistic discrimination. Consequently, utilizing the multi-token oddball paradigm for each Mandarin tone category allows researchers to test the robustness of lexical tone contrasts, beyond surface-level acoustic features, as the F0 variations span the entire syllable. (Yu et al., 2017). This method provides enhanced neurophysiological discrimination measures which are crucial for examining the underlying processes of speech discrimination and comparing the processing of native and non-native speech sounds.

This study employed a modified version of the passive ERP oddball paradigm (A1A1A2A1X or A1A2A1A1X) for the oddball task, facilitating a more direct comparison with neurophysiological responses (previous ERP oddball paradigm study: see Burnham et al., 1996; modified ERP oddball paradigm studies: see Yu et al., 2017; Yu et al., 2018; Yu et al., 2019).

3. Theoretical Frameworks on Lexical Tone Processing

Lexical tone processing in children is shaped by multiple overlapping mechanisms rooted in language acquisition, neural development, and cognitive control. This section introduces four major theories that inform the present study: the Bilingual

Advantage Theory, Language Exposure Theory, Home Language Effect, and the Hemispheric Effect.

3.1 Bilingual Advantage Theory

The Bilingual Advantage Theory posits that bilingualism confers cognitive benefits, particularly in executive function and language processing. Specifically in speech perception, early bilinguals tend to better discriminate phonological contrasts compared to late second-language (L2) learners, who exhibit poorer speech perception with non-native sounds acquired after age 11 (Strange, 2011; Gonzales & Lotto, 2013). However, the cognitive advantage may vary depending on language combinations and phonological properties, such as tone. Additionally, studies have reported the MMR polarity shifted from a robust positive response in younger children to a negative mismatch response (nMMR) as age increased (Werwach et al., 2022), particularly in bilingual participants (Yu et al., 2019). This shift may reflect increased attention allocation in bilingual children compared to monolingual children, suggesting enhanced cognitive effort with the auditory stimuli. However, the extent to which these findings generalize to more complex contrasts–such as lexical tone–remain unclear.

3.2 Language Exposure Theory

Closely related is the Language Exposure Theory, which emphasizes the amount, quality, and context of linguistic input as key factors in shaping auditory and phonological development. From an early age, infants and children attune to the frequencies, sound patterns, and prosodic features of the languages in their environment that they are habitual exposed to (Werker & Tee, 1984). Studies suggest that early and sustained exposure to meaningful input is critical for the development of phonological representations and auditory discrimination (Yu et al., 2019). Certain findings suggest that even in the absence of lexical tones in an infant's native phonemic repertoire, there is an initial neural distinction (Liu et al., 2014). For example, Singh et al. (2017), found that infants exposed to two languages could discern words with similar sounds that were distinguished by a vowel contrast, a skill that was not apparent in monolingual speakers. According to this view, it is not just bilingualism alone but the richness of exposure—especially to phonetically rich or typologically different languages—that drives perceptual tuning. Lexical tone processing, being highly dependent on acoustic salience, is particularly sensitive to these experiential differences during early childhood.

3.3 Home Language Effect

Building on these, the Home Language Effect provides a more nuanced understanding of how language environment modulates brain responses. Yu et al. (2019) found that the ratio of English to Spanish spoken at home predicted early mismatch responses (eMMRs) better than bilingual status alone. In their study, Spanish-dominant bilinguals showed reduced eMMR amplitudes compared to monolingual English peers, while balanced or English-dominant bilinguals showed comparable responses. These findings suggest that input quality and dominant home language shape neural sensitivity even within bilingual groups. Interestingly, later ERP components like LN were less influenced by home language balance, pointing to time-specific windows in which home exposure matters most.

3.4 Hemispheric Effect

At a neuroanatomical level, the Hemispheric Effect accounts for the lateralization of language functions across the brain. This study further explores how hemispheric specialization interacts with language experience. It is well established that the left hemisphere, particularly frontal and central regions (e.g., F3, Fz, C3, Cz), is dominant for segmental processing such as vowels and consonants, while the right hemisphere is more responsive to prosodic features like pitch contour and tone (Altmann et al., 2014; Ries et al., 2016; Yu et al., 2018). This lateralization pattern underscores the division of labor in how the brain processes different acoustic features depending on linguistic relevance and language experience. The present study examines whether these lateralization patterns extend to bilingual children processing lexical tones.

4. Purpose of the Present Study

By analyzing MMR responses to lexical tone processing in children across three groups—monolingual English (single non-tonal comparison), bilingual English-Mandarin (tonal and non-tonal comparison), and bilingual English-Spanish (double non-tonal comparison)—this study aims to explore whether the aforementioned theories extend to the domain of lexical tone discrimination. The purpose of this study is twofold: first, to examine the neurophysiological measures for lexical tone processing in bilingual learning children, and second, to examine the interactions between brain maturation and stimulus salience in bilingual speech processing. This study tests that learning multiple languages enhances cognitive skills. While numerous studies have explored this theory in various cognitive domains, such as attention (Paap et al., 2018), working memory (Grundy & Timmer, 2017), and problem-solving (Bialystok, 1999); its applicability to speech processing remains unclear. Specifically, it remains unclear whether the advantages are specific to certain language combinations or if they are a language-general phenomenon.

CHAPTER 2: METHODOLOGY

2.1 Participants

The study tested 20 children ages 5 to 10 (Table 1). They were divided into three groups: 8 bilingual Mandarin-English, 4 Spanish-English, and 8 monolingual mainstream English speakers. There was a total of *N=27* but 7 Spanish participants were excluded to significant noise or experimental errors. A total of 62400 trials were collected from the 20 children stacked into 60 waves using IGOR Pro 9. The participants were paid $50 per session for their voluntary participation. Voluntary informed consents were obtained from all the participants and their parental guardians at the beginning of their participation in the study. The study was approved by the human subject research institutional review board at the National Institute of Health (NIH) and at St. John's University, New York, and was conducted in compliance with the Declaration of Helsinki.

Table 1. Participant Demographics by Language Group

Group	N	Age Range
Monolingual English	8	5-10
Bilingual Mandarin-English	8	5-10
Bilingual Spanish-English	4	5-10
Total	**20**	—

2.2 Linguistic Measures

Before their ERPs were recorded, a hearing screening was administered to ensure that the participants' hearing fell within normal limits. All participants passed a hearing screening and had no history of neurological impairment.

The children were administered the Peabody Picture Vocabulary Test, 5th edition (PPVT-5) to test their receptive language skills and measure their current vocabulary acquisition. Receptive vocabulary knowledge provides a proxy for general linguistic ability in terms of semantic lexical development. Bilingual students were administered an additional PPVT-5 test in each of their spoken languages, *i.e.*, in Spanish and Mandarin.

The participants ages 6 and up were also tested for their non-verbal intelligence using the Test of Non-Verbal Intelligence, 4th Edition (TONI-4). This test was administered in order to gain insight into their aptitude in general level of cognition and processing ability. While the TONI-4 evaluates pattern-based reasoning and problem-solving abilities, if a student does not perform within a normal range on this test, his or her data will be excluded from this study.

These standardized tests ensured that each participant fell within one standard deviation of the mean level of performance in each of these areas: receptive vocabulary and non-verbal intelligence.

2.3 Stimuli and Paradigm

A passive mutli-oddball paradigm was used to elicit ERPs in response to changes in Mandarin lexical tone (adapted from Yu et al., 2017). The auditory stimuli consisted of natural speech tokens produced by a female native speaker of Mandarin and digitized at a 1000 Hz sampling rate. The syllabus followed a bisyllabic nonsense word structure in the form /gu?pa/, with three vowel variants: /gu/, /gi/, and /gy/. Lexical tone variations were applied only to the first syllable, while the second syllable /pa/ was consistently produced with Mandarin Tone 1 (high-level tone).

The primary focus of the ERP analysis was on the /gu?pa/ standard (Tone 3) and its two deviant counterparts: /gūpa/ or /gu1pa/ as Tone 1 and /gúpa/ or /gu2pa/ as Tone 2, forming a difficult contrast (Tone 3-Tone 2) and an easier contrast (Tone 3-Tone 1). Specifically, as shown in Figure 1 of Yu et al. (2018), the diagram of the mechanism indicates that three tokens of /gu3pa/ made up 80.6% of the trials, while /gi3pa/ and /gy3pa/ each accounted for 9.7%. The deviant stimuli consisted of two tokens each of /gu1pa/ (Tone 1) and /gu2pa/ (Tone 2) and accounted for 9.7% of the trials, resulting in 280 total deviant trials per tone category. As shown in Figure 2 of Yu et al. (2017), the F0 contours of the stimuli used in this study are represented by similar pitch contours of Mandarin tones (Figure 1). Two stimulus onset asynchrony (SOA) conditions were tested to examine how the timing between sounds influenced neural responses. In the short ISI condition, the SOA was 900 ms (with an average ISI of 575 ms, ranging from 545–609 ms), while the long ISI condition had an SOA of 3,000 ms (average ISI = 2,675 ms, range: 2,645–2,709 ms) (Yu et al., 2017). These timing differences allow researchers to assess the effects of temporal spacing on auditory discrimination. Overall, this design allowed for the examination of automatic neural response to pitch-based lexical contrasts without requiring active participation from the children.

During the experiment, children were seated comfortably in a sound-attenuated booth wearing headphones, and were instructed to relax while playing a game on an iPad or drawing. This passive listening environment was designed to ensure low task demand while maintaining alertness, allowing for reliable measurement of pre-attentive auditory processing through ERP responses. The procedure took 20 minutes.

2.4 Electroencephalography Recording and Offline Processing

Continuous EEG data were recorded from 64 scalp sites using a Geodesic Sensor Net, referenced online to the vertex electrode (Cz), with sampling rate[1] of 1000 Hz, with and bandwidth[2] of 0.1-100 Hz. Prior to analysis, the EEG data were refiltered offline using a finite impulse response (FIR) band-pass filter from 0.3-30 Hz in NetStation v5.4. The FIR filter was selected for its linear phase response, which minimizes waveform distortion. Artefact rejection and correction were conducted using Brain Electrical Source Analyses (BESA Research 7.1, BESA GmbH, Germany). Eye blinks and horizontal/vertical eye movements were automatically detected and corrected using computed HEOG and VEOG channels. A custom ERP paradigm file (*mand_andres.PD*) was used to define triggers, conditions, epoching parameters, and filters. Experimental conditions were defined for gu1 (Tone 1), gu2 (Tone 2), and gu3 (Tone 3), and were assigned categories 4, 5, and 6, respectively. Trigger attribute values corresponding to the number of trials for each tone: 140 for Tone 1, 140 for Tone 2, 854 for Tone 3. EEG signals were time-locked to the onset of each auditory stimulus and segmented into ±1000 ms epochs. Epochs included a -200 ms pre-stimulus baseline and 1000 ms post-stimulus window for duration, and -100 ms to 800 ms for artefact correction. Part of Yu et al. (2017) design, a bandpass filter cutoff frequencies of 0.30-30 Hz were applied to minimize very slow drift, using slopes of 6 db/oct (high-pass) and 24 db/oct (low-pass). Epochs exceeding a threshold of ± 250 µV or gradient of 75 V/ms were excluded from

---

[1] Sampling rate are data points collected per second from the continuous brain activity signal.
[2] Bandwidth range encompasses the range of brainwave frequencies commonly studied and analyzed: delta (0.1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (13-30 Hz), and gamma (30-100 Hz).

analysis. Consistently noisy channels[3] were interpolated–all participants have at least 70 % of trials after artefact rejection–and averaged using BESA 7.1. Data was re-referenced to the average reference after pre-processing.

ERP analyses focused on the six frontocentral scalp sites (see Figure 2): F3 (channel 12), Fz (channel 6), F4 (channel 60), C3 (channel 16), Cz (channel 65/VREF), and C4 (channel 50). These sites were selected due to their relevance in measuring brain activity related to higher-order cognitive processes and control, such as attention, memory, and re-orientation of attention (Light et al., 2010). Specifically, MMN/MMR was analyzed within the 100-300 ms post-stimulus window as a marker of early automatic auditory discrimination, whereas LN was analyzed within the 400-500+ ms window to index later cognitive processing.

The electrode net was first soaked in a saline solution and then excess solution was removed. The nets were placed on the participant's scalp. The impedances of electrodes were maintained at or below 50 kΩ, which is sufficient for high input impedance amplifiers (200 MΩ).

---

[3] *e.g.*, noisy channels 63, 62, 61, 64, 55, 23, 1, 17, 32, 43, 37, 5, 10

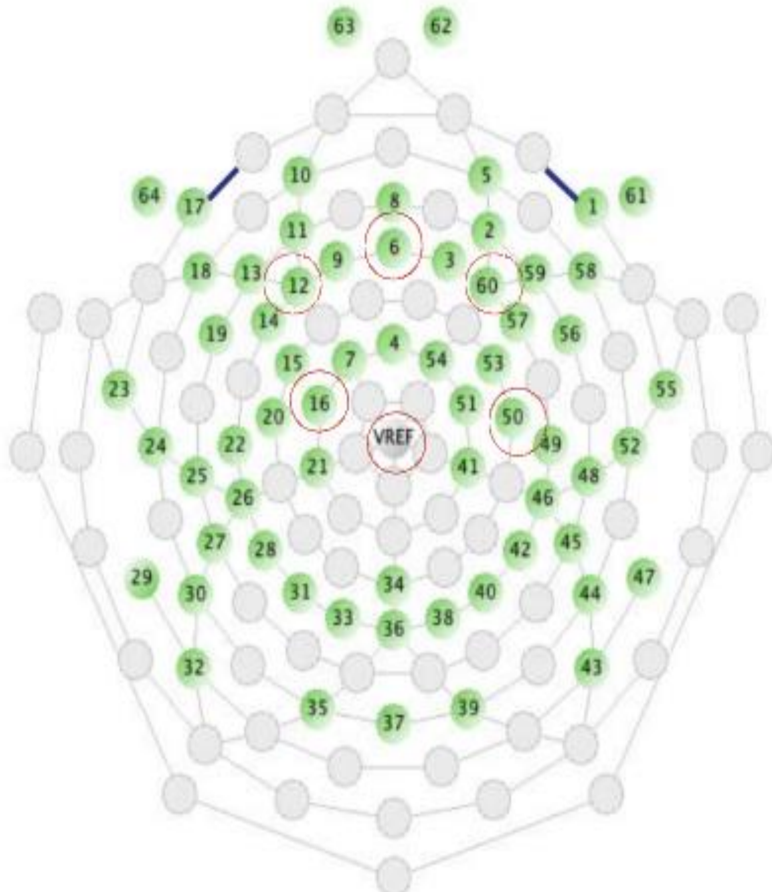

Figure 2: Scalp Electrode Sites of Interest from the Geodesic Tensor Net.

Electrodes used for ERP analysis: F3 (channel 12), Fz (channel 6), F4 (channel 60), C3 (channel 16), Cz (channel 65/VREF), and C4 (channel 50).

2.5 Statistical Analysis

Permutational multivariate analysis of variance (PERMANOVA) and permutation-based pairwise t-tests were used to analyze MMN and LN responses in this study. These non-parametric statistical methods are well-suited for EEG research, particularly when working with high-dimensional data collected across multiple scalp sites and time windows. They do not rely on the assumptions of normality or independence, making them particularly appropriate for small sample sizes and temporally correlated ERP data (Chen et al., 2013; Holt & Sullivan, 2025).

All ERP amplitudes were averaged across ten consecutive 20-ms time bins, capturing the relevant peak window for each ERP component (100–280 ms for MMR and 300–480 ms for LN). This averaging approach both smooths temporal noise and focuses the analysis on periods of maximal signal change.

PERMANOVA was selected as the primary statistical test for comparing neural responses between the three language groups (Mandarin-English, Spanish-English, and monolingual English). It calculates a pseudo-F statistic by computing the ratio of between-group variance to within-group variance, then generates a null distribution by randomly permuting group labels thousands of times. This provides a robust method for evaluating whether observed differences in ERP amplitudes across electrodes are statistically significant.

As a step-down analysis, permutation-based pairwise t-tests were performed for pairwise comparisons between groups (e.g., Mandarin vs. English). These tests reassign group membership randomly over many iterations to simulate the distribution of differences expected by chance. The observed group mean differences are then evaluated against this null distribution to yield permutation-based p-values.

The rationale for using both PERMANOVA and permutation t-tests is that ERP data, especially in developmental studies, often reflect subtle neural modulations influenced by both language background and age. These effects require both statistical sensitivity and robustness, which permutation-based methods afford. By leveraging the inherent correlations in ERP data across time and space, these tests reduce the risk of Type I errors and are less biased by small sample sizes.

All statistical analyses were conducted in RStudio, using the RVAideMemoire package for PERMANOVA and custom permutation scripts for t-tests.

CHAPTER 3: RESULTS

3.1 ERP Results

Table 2 presents individual average ERP amplitudes and standard deviations elicited by the deviant stimuli (Tone 1 and Tone 2) across six electrodes sites F3 (channel 12), Fz (channel 6), F4 (channel 60), C3 (channel 16), Cz (channel 65/VREF), and C4 (channel 50) between 100-300 ms. These values were measured within the MMN in the 100-300 ms window, typically associated with lexical tone discrimination processes. Table 3 presents individual average LN amplitudes within the later 300-500 ms window at the same electrode sites for the same deviant stimuli. This window captures later auditory-cognitive processes potentially linked to reorientation for lexical meaning. The data presented reflects the voltage fluctuations present in the tables are across 20 participants from 100-500 ms in response to the stimulus. These tables provide a quantitative comparison of mean MMN and LN amplitudes in response to deviant tones. The values correspond to the ERP waveforms of the deviant and standard shown in Figure 3 through Figure 5.

Figure 3, Figure 4, and Figure 5 displays the grand average ERPs for each participant per group to compare variability in EEG activities. Bilingual Mandarin-English children (Figure 3) exhibited stronger ERP responses to lexical tone stimuli, with more pronounced earlier peaks between 100–300 ms and reduced later peaks negativity between 400–600 ms ERP amplitudes, while monolingual English children (see Figure 5) demonstrated overall weaker amplitudes across both time windows; however, there MMN and LN responses were intermediate compared to bilingual Spanish-English children (see Chapter 4). Bilingual Spanish-English children (see Figure 4) showed intermediate and highly robust ERP amplitudes, particularly at Cz during the early peak window. Across all grand average waveforms, Cz showed significantly larger negative deflections for the bilingual Mandarin-English children compared to both bilingual Spanish-English and English monolingual children (see Figure 3). These hemispheric effects were most prominent in lateral frontal and central electrodes (F3, F4, C3, C4).

Table 2. Mean Amplitudes (µV) and Standard Deviations (SD) in the 100-300 ms Time Window (MMN) for Deviant Stimuli at Six Electrode Sites.

Note gu3 (/gu3pa/) is excluded as it serves as the standard stimulus.

			MMN											
			gu1 (Tone 3- Tone 1)						gu2 (Tone 3-Tone 2)					
Language Group	Participant ID	Time Window (ms)	F3	Fz	F4	C3	Cz	C4	F3	Fz	F4	C3	Cz	C4
Mandarin-English	R16C08	100-280	0.79	0.38	-0.05	-0.76	0.49	-0.02	0.60	1.28	0.75	-0.34	-0.63	-0.23
		(SD)	*1.05*	*1.29*	*0.70*	*0.89*	*0.46*	*0.42*	*0.29*	*0.78*	*0.41*	*0.68*	*0.31*	*0.39*
	R16C09	100-280	1.86	2.51	2.00	1.80	3.60	0.18	-2.28	-1.35	-1.51	0.69	1.03	-2.04
		(SD)	*1.09*	*0.92*	*1.91*	*1.10*	*1.41*	*0.61*	*0.93*	*0.89*	*0.53*	*1.21*	*0.86*	*0.73*
	R16C11	100-280	-0.63	-0.52	0.06	0.87	0.48	-0.08	-0.24	-1.13	-0.50	1.16	0.25	-0.83
		(SD)	*0.78*	*0.48*	*0.64*	*0.36*	*0.67*	*0.37*	*0.79*	*0.67*	*0.52*	*1.01*	*0.46*	*0.57*
	R16C13	100-280	0.10	-1.37	0.25	-3.27	-1.71	-0.25	0.59	0.54	1.78	0.72	-0.30	0.20
		(SD)	*0.38*	*0.75*	*0.48*	*1.24*	*0.60*	*0.42*	*0.70*	*0.66*	*0.88*	*0.70*	*0.52*	*0.43*
	R16C17	100-280	-0.36	-1.11	-0.45	0.96	1.53	0.22	1.22	-0.17	-0.62	1.91	0.40	0.68
		(SD)	*1.45*	*0.32*	*0.45*	*0.80*	*0.28*	*0.31*	*0.44*	*0.46*	*0.63*	*0.61*	*0.55*	*0.43*
	R16C23	100-280	1.56	1.63	-2.58	1.63	1.23	-1.12	-1.21	-1.08	-0.08	0.60	-0.09	0.66
		(SD)	*1.51*	*2.30*	*1.44*	*0.50*	*1.78*	*1.37*	*1.64*	*1.90*	*1.60*	*0.92*	*1.09*	*0.79*
	R16C28	100-280	0.86	-0.77	-0.25	-0.27	-0.86	-0.51	-0.41	-1.01	0.68	-0.04	-1.15	0.34
		(SD)	*0.90*	*0.56*	*0.61*	*0.71*	*0.68*	*0.73*	*0.57*	*0.52*	*0.61*	*0.93*	*0.90*	*0.76*
	R16C30	100-280	0.32	1.37	-1.09	0.97	2.43	-0.11	-2.19	-0.78	0.82	-0.32	-3.92	-0.54
		(SD)	*1.07*	*0.64*	*1.10*	*0.37*	*1.19*	*0.63*	*1.00*	*1.29*	*0.89*	*0.68*	*1.76*	*0.85*
Spanish-English	R16S01	100-280	0.72	-0.10	0.31	-0.25	-1.85	-0.05	-0.45	1.83	0.02	0.19	0.57	0.58
		(SD)	*0.54*	*0.99*	*0.56*	*0.94*	*1.59*	*0.51*	*0.38*	*1.53*	*0.70*	*1.19*	*1.47*	*0.56*
	R16S02	100-280	0.72	-0.10	0.31	-0.25	-1.85	-0.05	-0.88	1.70	-1.84	-1.14	1.26	-0.05
		(SD)	*0.54*	*0.99*	*0.56*	*0.94*	*1.59*	*0.51*	*0.68*	*0.34*	*0.38*	*0.58*	*0.54*	*0.75*
	R16S03	100-280	4.42	-0.79	1.05	4.31	2.76	1.36	-0.64	-0.98	0.67	1.08	0.59	-0.24
		(SD)	*0.61*	*0.38*	*0.57*	*0.85*	*0.68*	*0.80*	*0.34*	*0.62*	*0.24*	*0.35*	*0.27*	*0.57*
	R16S04	100-280	-4.22	-4.20	-5.54	-4.43	-4.58	-1.97	0.83	0.24	3.13	-5.34	-3.02	-3.21
		(SD)	*0.63*	*0.44*	*0.79*	*1.80*	*1.22*	*0.56*	*0.81*	*0.41*	*0.88*	*3.30*	*1.12*	*0.64*
	R16S06	100-280	0.46	-0.34	-1.17	0.55	0.00	-0.42	1.07	0.18	0.11	-0.26	-0.99	-0.24
		(SD)	*0.79*	*0.51*	*0.86*	*0.63*	*0.49*	*0.38*	*1.14*	*0.90*	*0.71*	*0.74*	*0.80*	*0.65*
	R16S07	100-280	-0.18	-0.01	-1.29	-0.09	-1.05	-0.65	-2.10	1.00	0.14	1.37	0.28	0.89
		(SD)	*0.48*	*0.34*	*0.50*	*0.59*	*1.32*	*0.29*	*0.57*	*0.24*	*0.46*	*0.72*	*1.16*	*0.72*
	R16S08	100-280	-0.06	-1.13	1.07	-0.10	0.19	0.20	2.25	1.01	-0.44	-1.03	0.84	-1.97
		(SD)	*2.07*	*1.76*	*1.89*	*2.25*	*2.46*	*1.09*	*0.76*	*1.21*	*0.66*	*0.54*	*0.72*	*0.22*
	R16S09	100-280	6.05	2.23	0.41	-2.13	0.14	0.03	8.49	3.91	1.38	-2.08	1.18	3.03
		(SD)	*1.51*	*0.23*	*0.81*	*0.32*	*1.06*	*0.82*	*0.80*	*0.29*	*0.88*	*0.44*	*0.78*	*0.45*
English	R16E01	100-280	-3.40	-2.51	-1.94	-2.86	-2.48	-0.14	-1.81	0.24	0.26	-1.22	-1.72	0.34
		(SD)	*0.89*	*0.81*	*0.59*	*0.78*	*0.40*	*1.09*	*0.87*	*0.58*	*0.92*	*0.73*	*0.91*	*0.54*
	R16E02	100-280	0.82	0.52	-1.37	1.50	1.93	1.74	-1.59	-0.03	-1.88	0.19	1.25	0.63
		(SD)	*1.01*	*0.49*	*0.70*	*1.40*	*2.21*	*0.99*	*0.18*	*0.48*	*0.54*	*0.57*	*0.88*	*0.85*
	R16E04	100-280	0.40	0.97	1.40	0.37	-0.71	0.37	0.34	1.07	1.95	-0.62	0.14	1.17
		(SD)	*0.35*	*0.46*	*0.90*	*0.58*	*0.28*	*0.40*	*0.41*	*0.52*	*0.55*	*0.27*	*0.35*	*0.63*
	R16E05	100-280	1.13	-0.45	-0.48	-0.40	-4.00	-0.96	0.44	-0.27	-0.90	-0.70	-1.51	-0.78
		(SD)	*0.75*	*1.61*	*0.89*	*0.51*	*0.72*	*1.05*	*0.43*	*0.59*	*0.76*	*0.39*	*0.52*	*0.63*

Table 3. Mean Amplitudes (μV) and SD in the 300-480 ms Time Window (LN) for Deviant Stimuli at Six Electrode Sites.

Note gu3 (/gu3pa/) is excluded as it served as the standard stimulus.

			LN											
			gu1 (Tone 3- Tone 1)						gu2 (Tone 3-Tone 2)					
Language Group	Participant ID	Time Window (ms)	F3	Fz	F4	C3	Cz	C4	F3	Fz	F4	C3	Cz	C4
Mandarin-English	R16C08	300-480	-1.56	-2.33	-3.40	-2.92	-0.70	1.66	-1.10	-1.35	-0.42	-2.43	-2.52	1.00
		(SD)	*0.61*	*0.48*	*0.89*	*1.26*	*0.77*	*0.17*	*1.63*	*0.85*	*0.99*	*0.76*	*0.43*	*0.19*
	R16C09	300-480	2.72	2.70	4.19	3.71	4.92	2.88	-5.22	-2.35	-0.32	3.40	4.81	1.32
		(SD)	*0.77*	*1.29*	*2.01*	*0.54*	*1.04*	*1.47*	*1.42*	*0.55*	*1.48*	*0.44*	*0.95*	*0.90*
	R16C11	300-480	-1.22	-1.20	-0.18	1.36	0.52	-0.58	-1.00	-0.50	-0.44	1.32	1.57	0.57
		(SD)	*0.59*	*0.48*	*0.80*	*0.53*	*0.69*	*0.60*	*0.77*	*0.75*	*0.77*	*0.52*	*0.36*	*0.63*
	R16C13	300-480	-0.53	-3.63	-1.68	-6.13	-3.81	-0.32	-0.32	-1.58	0.04	-1.49	-2.75	-0.27
		(SD)	*0.52*	*0.37*	*0.50*	*0.58*	*0.85*	*0.52*	*0.82*	*0.80*	*0.67*	*0.97*	*0.62*	*0.50*
	R16C17	300-480	-3.45	-2.50	-0.93	-2.00	-1.23	-1.21	-1.73	-1.28	-1.05	-0.87	-1.90	-0.57
		(SD)	*0.65*	*0.75*	*0.89*	*0.70*	*0.93*	*0.89*	*1.23*	*1.03*	*0.83*	*1.22*	*1.18*	*1.27*
	R16C23	300-480	1.17	1.53	-1.90	2.38	-0.13	-1.32	-1.29	-4.16	-3.06	1.93	-1.75	-0.70
		(SD)	*2.42*	*2.19*	*1.03*	*0.87*	*0.75*	*0.78*	*1.70*	*2.14*	*1.26*	*0.58*	*0.87*	*0.55*
	R16C28	300-480	0.36	-0.74	-0.34	-0.33	-1.60	-1.33	-1.26	-1.16	0.85	-0.36	-2.16	-1.03
		(SD)	*0.51*	*0.61*	*0.50*	*0.53*	*0.53*	*0.50*	*0.60*	*0.36*	*0.62*	*1.36*	*0.96*	*0.90*
	R16C30	300-480	-0.79	1.65	-2.43	0.35	2.43	-0.49	-3.91	-2.31	-0.82	0.44	-1.76	-0.56
		(SD)	*1.38*	*1.29*	*1.28*	*1.18*	*1.35*	*0.96*	*1.02*	*1.30*	*1.64*	*0.94*	*2.95*	*1.07*
Spanish-English	R16S01	300-480	1.21	-0.15	-0.25	1.09	-1.65	-0.86	1.05	2.49	0.51	0.35	-1.08	0.21
		(SD)	*0.70*	*1.51*	*0.49*	*1.19*	*1.19*	*0.46*	*0.41*	*0.82*	*0.39*	*0.45*	*0.38*	*0.50*
	R16S02	300-480	2.76	-0.62	4.53	8.65	12.34	2.84	-0.50	0.61	-0.33	1.96	5.66	1.30
		(SD)	*1.21*	*0.88*	*0.65*	*0.53*	*2.30*	*0.39*	*0.27*	*0.28*	*0.46*	*0.56*	*0.55*	*0.53*
	R16S03	300-480	0.39	4.53	-2.64	7.08	3.17	2.28	-0.50	0.44	0.20	5.03	2.39	2.69
		(SD)	*0.40*	*0.58*	*0.26*	*0.71*	*0.46*	*0.64*	*1.47*	*0.24*	*0.52*	*0.70*	*0.56*	*0.52*
	R16S04	300-480	-7.06	-3.72	-9.75	0.40	5.11	-1.95	-2.66	0.62	2.34	-4.18	3.31	-4.91
		(SD)	*1.21*	*0.64*	*1.11*	*0.52*	*0.84*	*0.60*	*0.65*	*0.76*	*0.45*	*1.20*	*0.79*	*1.16*
	R16S06	300-480	0.08	-0.60	-0.91	0.16	-0.43	-0.19	-1.78	-1.15	-1.00	-0.66	-1.29	-0.68
		(SD)	*0.74*	*0.50*	*0.52*	*0.55*	*0.59*	*0.88*	*1.18*	*0.73*	*0.53*	*1.10*	*0.58*	*0.30*
	R16S07	300-480	1.43	-1.37	-0.54	-0.25	-0.93	-0.88	-0.57	0.02	0.57	1.97	0.54	1.27
		(SD)	*0.50*	*0.99*	*0.67*	*0.91*	*1.95*	*1.38*	*0.82*	*0.50*	*0.40*	*0.46*	*1.16*	*0.83*
	R16S08	300-480	1.30	-3.18	-1.33	-0.83	-0.72	-0.15	2.81	0.58	-1.14	-1.07	0.77	-2.59
		(SD)	*0.36*	*0.95*	*0.85*	*0.71*	*0.62*	*0.52*	*1.00*	*0.94*	*1.42*	*0.49*	*0.52*	*0.69*
	R16S09	300-480	6.85	0.34	0.58	-2.29	1.24	0.00	6.67	2.97	-1.15	-2.16	2.62	1.89
		(SD)	*1.60*	*1.19*	*1.58*	*0.29*	*1.41*	*1.00*	*2.11*	*1.09*	*1.87*	*0.15*	*1.08*	*1.24*
English	R16E01	300-480	0.67	-1.69	0.71	-3.69	-7.75	-0.46	1.86	2.59	2.61	-2.75	-6.72	-0.76
		(SD)	*0.40*	*0.42*	*1.17*	*1.20*	*1.52*	*0.56*	*0.62*	*0.36*	*0.93*	*1.20*	*1.85*	*0.91*
	R16E02	300-480	0.76	0.12	0.99	0.79	1.89	2.27	-0.43	0.28	0.43	0.70	0.87	0.72
		(SD)	*0.29*	*0.90*	*0.82*	*0.92*	*1.62*	*0.94*	*0.25*	*0.38*	*0.61*	*0.55*	*0.72*	*0.71*
	R16E04	300-480	-0.28	1.17	-0.25	0.15	-0.51	-0.36	-0.08	-0.13	-0.41	-0.27	-0.51	0.51
		(SD)	*0.43*	*0.22*	*0.44*	*0.60*	*0.10*	*0.44*	*0.58*	*0.35*	*0.40*	*0.21*	*0.55*	*0.51*
	R16E05	300-480	0.88	1.18	0.41	0.52	-1.22	-0.31	0.07	0.09	0.36	-0.06	0.17	0.29
		(SD)	*1.15*	*1.55*	*1.23*	*1.44*	*1.00*	*1.16*	*1.33*	*1.46*	*1.83*	*1.85*	*2.37*	*2.26*

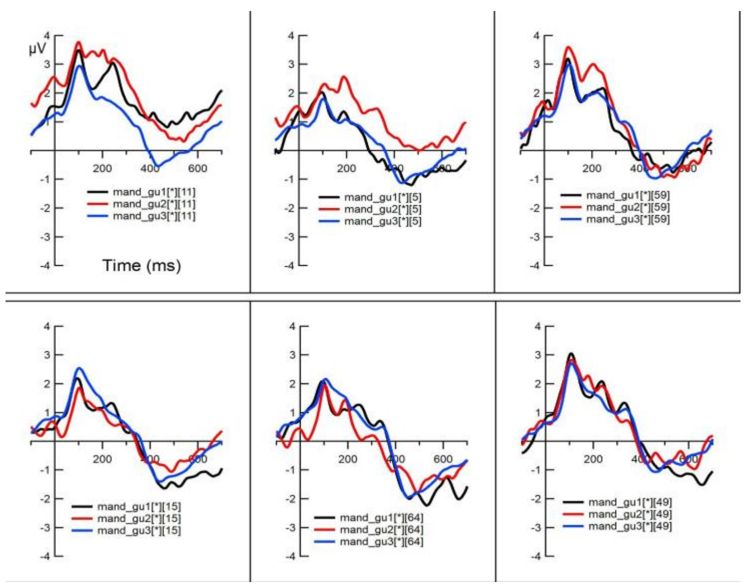

Figure 3. Grand Average ERP Waveforms for Tone 1 (/gu1pa/), Tone 2 (/gu2pa/), and Tone 3 (/gu3pa/) in Bilingual Mandarin-English Children across Six Electrodes.

Electrodes F3, Fz, F4 (top row) and C3, Cz, C4 (bottom row), shown left to right

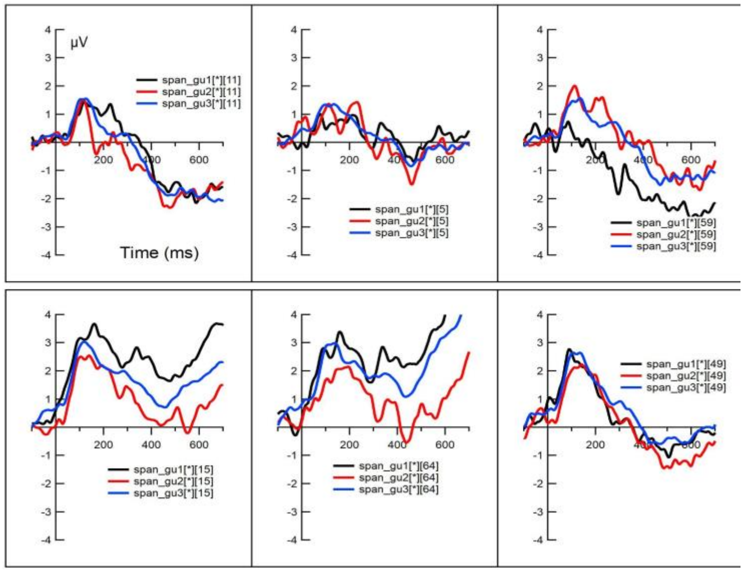

Figure 4. Grand Average ERP Waveforms for Tone 1 (/gu1pa/), Tone 2 (/gu2pa/), and Tone 3 (/gu3pa/) in Bilingual Spanish-English Children across Six Electrodes.

Electrodes F3, Fz, F4 (top row) and C3, Cz, C4 (bottom row), shown left to right

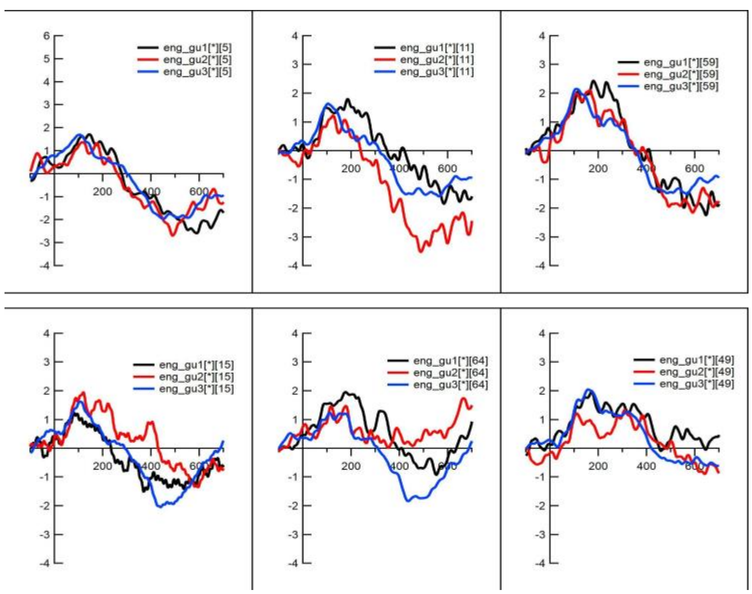

Figure 5. Grand Average ERP Waveforms for Tone 1 (/gu1pa/), Tone 2 (/gu2pa/), and Tone 3 (/gu3pa/) in Monolingual English Children across Six Electrodes.

Electrodes F3, Fz, F4 (top row) and C3, Cz, C4 (bottom row), shown left to right

Table 4. Mean MMN Amplitudes (µV) and SD in the 100-300 ms Time Window for Deviants Stimuli at Six Electrodes

Note: gu3 (/gu3pa/) is excluded as it served as the standard stimulus

GROUP MMN													
		gu1 (Tone 3- Tone 1)						gu2 (Tone 3-Tone 2)					
Language Group	Time Window (ms)	F3	Fz	F4	C3	Cz	C4	F3	Fz	F4	C3	Cz	C4
Mandarin-English	100-280	0.56	0.27	-0.27	0.24	0.90	-0.21	-0.49	-0.46	0.17	0.55	-0.55	-0.22
	(SD)	*0.88*	*1.43*	*1.29*	*1.67*	*1.71*	*0.43*	*1.31*	*0.94*	*1.04*	*0.77*	*1.51*	*0.91*
Spanish-English	100-280	0.99	-0.55	-0.61	-0.30	-0.78	-0.19	1.07	1.11	0.40	-0.90	0.09	-0.15
	(SD)	*3.10*	*1.78*	*2.19*	*2.46*	*2.13*	*0.93*	*3.29*	*1.45*	*1.44*	*2.13*	*1.44*	*1.87*
English	100-280	-0.26	-0.37	-0.60	-0.35	-1.31	0.25	-0.66	0.25	-0.14	-0.59	-0.46	0.34
	(SD)	2.11	1.55	1.46	1.85	2.55	1.14	1.21	0.58	1.64	0.59	1.41	0.82

Table 5. Permutation ANOVA and Two-Sample t-Test for MMN Amplitudes (µV) Across Six Electrodes by Group Difference

Electrode	***Mandarin x English (ANOVA)***	***Mandarin x Spanish (ANOVA)***	***English x Spanish (ANOVA)***	***Mandarin x English (t-test)***	***Mandarin x Spanish (t-test)***	***English x Spanish (t-test)***
F3	**p < .001***	**p = .001***	**p < .001***	**p < .001***	**p = .001***	**p < .001***
Fz	p = .847	**p = .001***	p = .0675	p = .1088	**p = .0182***	p = .0726
F4	p = .1127	p = .3085	p = .6266	p = .1088	p = .3124	p = .6266
C3	**p < .001***	**p < .001***	p = .4866	**p < .001***	**p < .001***	p = .4866
Cz	**p < .001***	p = .1116	**p = .009***	**p < .001***	p = .1116	**p = .009***
C4	**p < .001***	p = .8178	**p = .013***	p = .1204	p = .8178	**p = .0146***

Note. Asterisks indicate p < .05 (statistically significant)

3.2 Statistical Analysis of MMR/MMN Amplitudes

Statistical Analysis of MMN and MMR Amplitudes Permutation-based statistical analysis (ANOVA and pairwise t-tests) were conducted to examine MMR amplitudes in response to Tone 1 (/gu1pa/) and Tone 2 (/gu2pa/) deviants across six frontocentral electrodes (F3, Fz, F4, C3, Cz, C4) within the 100–300 ms time window. PERMANOVA was used as the primary group-level test to identify whether group-level differences (bilingual Mandarin-English, bilingual Spanish-English, monolingual English) existed at each site, with follow-up permutation t-tests serving as step-down pairwise comparisons. In sites where MMN was not reliably evoked—either minimal, absent, or positive—the term MMR was adopted.

At F3, PERMANOVA revealed the strongest group-level differences at this site ($p < .001$), highlighting F3's sensitivity to language background. Permutation t-tests showed that the bilingual Mandarin-English group exhibited significantly more negative MMN amplitudes than both the monolingual English ($p < .001$) and bilingual Spanish-English groups ($p < .001$), respectively. Additionally, the monolingual English group exhibited more negative amplitudes than the bilingual Spanish-English group ($p < .001$), suggesting a graded MMR.

At Fz, PERMANOVA indicated a significant group difference between the bilingual Mandarin-English and bilingual Spanish-English groups ($p = .001$), suggesting midline sensitivity to tonal exposure. Permutation t-tests confirmed this finding ($p = .0182$), with the bilingual Mandarin-English group demonstrating a more robust MMN. No significant differences were observed between the bilingual Mandarin-English and monolingual English groups, nor between the monolingual English and bilingual Spanish-English groups, though the latter approached marginal significance ($p = .0726$), suggestive of a weak MMR.

At F4, PERMANOVA revealed no significant group differences at this site, indicating minimal influence of language background. Permutation t-tests confirmed the absence of group effects (all $p > .1$), with MMR being absent or negligible. In other words, this part of the brain did not respond differently based on language background.

At C3, PERMANOVA revealed significant group differences ($p < .001$), primarily driven by the bilingual Mandarin-English group. Permutation t-tests showed that the bilingual Mandarin-English group had significantly more negative MMN amplitudes than both the monolingual English ($p < .001$) and bilingual Spanish-English groups ($p < .001$). No difference was found between the monolingual English and bilingual Spanish-English groups ($p = .4866$), both of which exhibited attenuated MMRs.

At Cz, PERMANOVA showed that the bilingual Mandarin-English and monolingual English groups were statistically similar; however, both groups differed significantly from the bilingual Spanish-English group, respectively. Specifically, the bilingual Spanish-English group exhibited the weakest (i.e., least negative or more positive) brain responses at Cz to the MMR. Permutation t-tests further clarified these patterns: the bilingual Mandarin-English group had significantly more negative MMN amplitudes than the monolingual English group ($p < .001$), and the monolingual English group showed more negative MMN amplitudes than the bilingual Spanish-English group ($p = .009$). This demonstrates another graded MMN pattern, with bilingual Mandarin-English children showing the strongest response, monolingual English children showing an intermediate MMN, and bilingual Spanish-English children showing a positive MMR.

At C4, PERMANOVA revealed significant group differences at this site ($p = .013$), with the monolingual English group showing stronger MMN responses than the bilingual Spanish-English group, but weaker MMN compared to the bilingual Mandarin-English group. Permutation t-tests indicated that the monolingual English group had significantly more negative MMN amplitudes than the bilingual Spanish-English group ($p = .0146$). No other group comparisons at C4 reached significance, including the comparison between bilingual Mandarin-English and monolingual English groups ($p = .1204$).

This analysis confirms that MMN responses were most robust in the bilingual Mandarin-English group, particularly over left frontal and central sites (F3, C3, Fz, and Cz), aligning with the neural architecture of phonological processing. MMR patterns were more prevalent in the bilingual Spanish-English group, especially at F4 and C4, suggesting limited early sensitivity to pitch contrasts in the absence of tonal language experience.

See Table 4 for the averaged amplitudes of each group per electrode and Table 5 for the results from the PERMANOVA and permutation pairwise t-tests for each comparison. Figure 6 illustrates the difference waves isolating MMN components (in green) for each group and contrast condition.

3.3 Acoustic Saliency Effects on Mismatch Responses (MMN/MMR)

The effects of acoustic saliency were evaluated by examining mismatch responses to two Mandarin tonal deviants—Tone 1 (/gu1pa/, high-level) and Tone 2 (/gu2pa/, rising)—across frontal and central electrode sites. Deviant Tone 2, being acoustically more similar to the standard (Tone 3), and deviant Tone 1, being more distinct, allowed for the assessment of saliency effects. These MMN amplitudes were driven by the ERP waveforms illustrated in Figures 3–5 and are further detailed in Figure 6, which displays frontal (F3, Fz, F4) and central (C3, Cz, C4) scalp distributions.

In the frontal regions, at F3, the bilingual Mandarin-English group exhibited robust MMNs for both tonal deviants, reflecting heightened pre-attentive sensitivity to pitch contrasts. In contrast, the monolingual English and bilingual Spanish-English groups showed attenuated or absent MMRs, indicating minimal mismatch activity at this location. At Fz, the bilingual Mandarin-English group again demonstrated clear MMNs to both tones, while the monolingual English group showed an intermediate MMN—particularly in response to Tone 2—and the bilingual Spanish-English group exhibited a weak MMR, suggestive of reduced midline sensitivity to tonal saliency. This pattern supports a graded frontal midline response that aligns with the extent of tonal language experience. At F4, mismatch activity was generally weak. All groups displayed attenuated or absent MMRs, although the bilingual Spanish-English group showed a relatively robust MMR to Tone 1. This unique finding of right-frontal engagement in the bilingual Spanish-English group will be explored further in the Chapter 4.

In the central regions, the bilingual Mandarin-English group again showed strong MMNs at C3 in response to both tones. By contrast, the monolingual English and bilingual Spanish-English groups demonstrated negligible MMR. This reinforces the idea that the left-central region is particularly sensitive to pitch contrasts in children with tonal language exposure. At Cz, a graded response was observed: the bilingual Mandarin-English group exhibited the most negative MMN amplitudes, followed by a moderate MMN in the monolingual English group, and a weak or more positive MMR in the bilingual Spanish-English group. This midline central trend further supports the influence of both saliency and language experience on mismatch detection. Finally, at C4, mismatch activity was generally attenuated. Only the monolingual English group showed a distinct MMN to Tone 2, while the Mandarin-English and Spanish-English groups displayed weak or absent MMRs, reflecting limited discrimination at this right-central site.

Table 6. Mean LN Amplitudes (µV) and SD in the 300-480 ms Time Window for Deviants Stimuli at Six Electrodes

Note: gu3 (/gu3pa/) is excluded as it served as the standard stimulus

GROUP LN													
		gu1 (Tone 3- Tone 1)						gu2 (Tone 3-Tone 2)					
Language Group	Time Window (ms)	F3	Fz	F4	C3	Cz	C4	F3	Fz	F4	C3	Cz	C4
Mandarin-English	300-480	-0.41	-0.56	-0.83	-0.45	0.05	-0.09	-1.98	-1.83	-0.65	0.24	-0.81	-0.03
	(SD)	*1.86*	*2.29*	*2.30*	*3.16*	*2.66*	*1.54*	*1.68*	*1.12*	*1.13*	*1.92*	*2.64*	*0.87*
Spanish-English	300-480	0.87	-0.60	-1.29	1.75	2.27	0.13	0.57	0.82	0.00	0.16	1.61	-0.10
	(SD)	*3.85*	*2.52*	*4.01*	*3.93*	*4.68*	*1.63*	*2.98*	*1.32*	*1.18*	*2.84*	*2.34*	*2.54*
English	300-480	0.51	0.19	0.47	-0.56	-1.90	0.29	0.35	0.71	0.75	-0.59	-1.55	0.19
	(SD)	0.53	1.35	0.53	2.11	4.12	1.33	1.02	1.27	1.30	1.50	3.49	0.66

Table 7. Permutation ANOVA and Two-Sample t-Test for LN Amplitudes (µV) Across Six Electrodes by Group

Electrode	Mandarin x English (ANOVA)	Mandarin x Spanish (ANOVA)	English x Spanish (ANOVA)	Mandarin x English (t-test)	Mandarin x Spanish (t-test)	English x Spanish (t-test)
F3	**p < .001***	**p < .001***	p = .3951	**p < .001***	**p = .0002**	p = .3948
Fz	**p < .001***	**p < .001***	**p = .0108**	**p < .001***	**p = .0002**	**p = .0092**
F4	**p < .001***	p = .7351	**p = .0002***	**p < .001***	p = .7238	**p = .0002**
C3	p = .1549	**p = .0016**	**p < .001***	p = .1618	**p = .0024**	**p = .0002**
Cz	**p = .0012**	**p < .001***	**p < .001***	**p = .0014**	**p = .0002**	**p = .0002**
C4	p = .1242	p = .709	p = .4964	p = .1204	p = .703	p = .4862

Note. Asterisks indicate p < .05 (statistically significant).

3.3 Statistical Analysis of LN Amplitude

Permutation-based statistical analysis (ANOVA and pairwise T-tests) were conducted to examine late negativity (LN) amplitudes in response to Tone 1 (/gu1pa/) and Tone 2 (/gu2pa/) deviants across six frontocentral electrodes (F3, Fz, F4, C3, Cz, C4) within the 300–480 ms time window. PERMANOVA served as the primary method to detect overall group differences across bilingual Mandarin-English, bilingual Spanish-English, and monolingual English children, with follow-up permutation t-tests used for pairwise comparisons.

At F3, PERMANOVA revealed significant group differences ($p < .001$), with the bilingual Mandarin-English group significantly differing from both the monolingual English and bilingual Spanish-English groups, respectively. Permutation t-tests confirmed that LN amplitudes were significantly more negative in the Mandarin-English group than in the monolingual English ($p < .001$) and bilingual Spanish-English ($p = .0002$) groups, respectively. No significant difference was observed between the monolingual English and bilingual Spanish-English groups ($p = .3948$).

At Fz, PERMANOVA indicated the strongest overall group-level difference ($p < .001$). Pairwise permutation t-tests revealed that the bilingual Mandarin-English group showed more negative LN amplitudes than both the monolingual English ($p < .001$) and bilingual Spanish-English ($p < .001$) groups, respectively. Additionally, the monolingual English group exhibited significantly greater LN negativity than the bilingual Spanish-English group ($p = .0092$), indicating a graded pattern in LN response strength across groups.

At F4, PERMANOVA revealed a significant group effect ($p < .001$), with the monolingual English group significantly differing from both bilingual groups, respectively. No difference was observed between the bilingual Mandarin-English and bilingual Spanish-English groups ($p = .7351$). Pairwise permutation t-tests showed that the monolingual English group exhibited significantly more negative LN amplitudes than both the bilingual Mandarin-English ($p < .001$) and bilingual Spanish-English ($p = .0002$) groups, respectively.

At C3, PERMANOVA showed a significant group difference ($p = .0016$), primarily driven by the bilingual Spanish-English group differing from both the bilingual Mandarin-English and monolingual English groups, respectively. Permutation t-tests confirmed that both the bilingual Mandarin-English ($p = .0024$) and monolingual English ($p = .0002$) groups had significantly larger LN amplitudes than the bilingual Spanish-English group, respectively.

At Cz, PERMANOVA indicated robust group differences ($p = .0012$). A clear rank order of LN amplitude was observed across groups: bilingual Mandarin-English > monolingual English > bilingual Spanish-English. Permutation t-tests supported this pattern: the bilingual Mandarin-English group had significantly more negative LN amplitudes than both the monolingual English ($p = .0014$) and bilingual Spanish-English ($p = .0002$) groups, respectively, and the monolingual English group also differed significantly from the bilingual Spanish-English group ($p = .0002$).

At C4, neither PERMANOVA (all $p > .05$) nor permutation t-tests showed any significant group differences. LN amplitudes at this right-central site were comparable across all three language groups, suggesting limited involvement in later-stage lexical tone processing.

See Table 6 for the averaged amplitudes of each group per electrode and Table 7 for the results from the PERMANOVA and permutation pairwise t-tests for each comparison. Figure 6 illustrates the difference waves isolating LN components (in purple) for each group and contrast condition.

3.4 Acoustic Saliency Effects on Late Negativity Responses (LN)

The LN amplitudes associated with Tone 1 (/gu1pa/) and Tone 2 (/gu2pa/) deviants were examined across frontal and central electrode sites to evaluate the effects of acoustic saliency. These amplitude patterns were driven by the ERP waveforms shown in Figures 3–5, with topographic distribution and statistical comparisons summarized in Figure 6 (Panels: F3, Fz, F4, C3, Cz, C4) and detailed in the statistical analysis section.

In the frontal regions, at F3, the bilingual Mandarin-English group exhibited the strongest LN amplitudes to both tones. In contrast, the monolingual English and bilingual Spanish-English groups showed attenuated responses. At Fz, the bilingual Mandarin-English group again demonstrated strong LN responses to both tones. The monolingual English group exhibited moderate LN amplitudes, particularly to Tone 2, while the bilingual Spanish-English group displayed the weakest response to both tones. These findings reflect a graded LN saliency pattern across groups, supported by the statistical comparisons. At F4, the monolingual English group uniquely exhibited the strongest LN response to Tone 1, differing from both bilingual groups—a statistically significant pattern indicating possible lateralized processing during late-stage attention or evaluation.

In the central region, at C3, both the bilingual Mandarin-English and monolingual English groups displayed larger LN amplitudes than the bilingual Spanish-English group, which again showed the weakest and most positive-going response to tone 1. At Cz, the most robust group differentiation was observed to Tone 2: Mandarin-English children evoked the most negative LN, followed by the monolingual English group, while the Spanish-English group exhibited minimal LN activity. These differences were statistically validated in the analysis section. Finally, at C4, all groups displayed similarly weak responses, and no statistically significant differences were observed, indicating a lack of late-stage engagement with acoustic saliency at this site regardless of language background.

Late Negativity (LN) responses to acoustic saliency varied across language groups and scalp regions. Mandarin-English bilingual children consistently showed stronger LN amplitudes, especially at left and midline sites, reflecting greater late-stage neural engagement with pitch contrasts. In contrast, monolingual English and bilingual Spanish-English children exhibited weaker responses, with the Spanish-English group showing the most attenuated LN. Right-hemisphere sites showed minimal activity across all groups. These effects were observed for both the more salient (Tone 1) and less salient (Tone 2) contrasts and reflect previously identified statistical differences in LN amplitudes, highlighting the influence of language background and hemispheric processing.

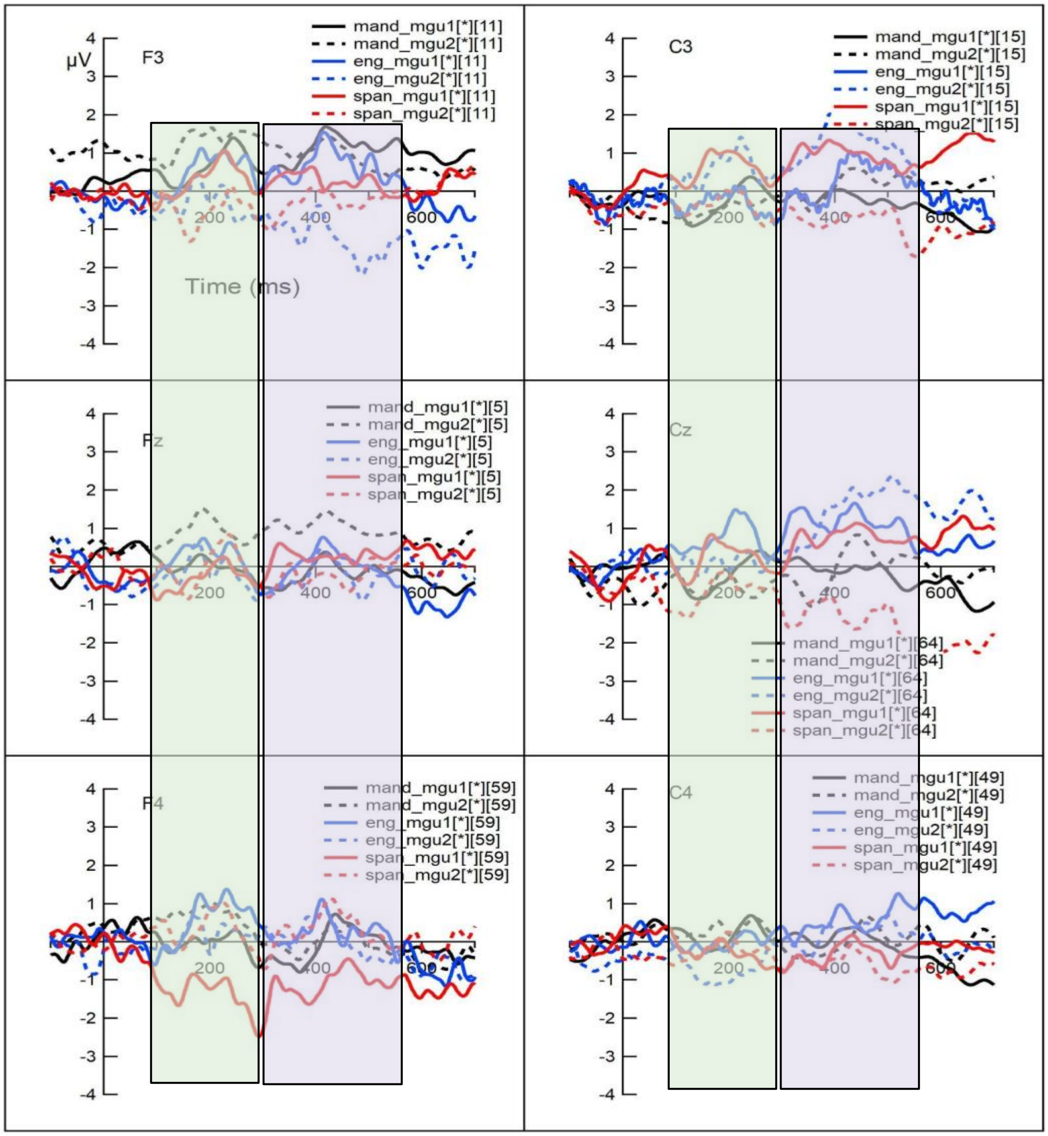

Figure 6. MMN and LN wave forms for Tone 1 (gu3pa-gu1pa) and Tone 2 (gu3pa-gu2pa) Deviants Across Language Groups.

See Figures 3-5 for Standard and Deviant Waveforms Used for the MMN and LN. Solid lines represent Tone3-Tone 1 (easy contrast) difference waves; dashed lines represent Tone 3-Tone 2 (hard contrast) difference waves.

Panels: Electrodes: (Left, Top to Bottom) F3, Fz, F4, (Right, Top to Bottom) C3, Cz, and C4

CHAPTER 4: DISCUSSION

The current study was designed to evaluate neural correlate modulation of MMR and LN to Mandarin lexical tones in children 5 to 10 years of age by the factor of bilingual or monolingual language experience. The acoustic-phonetic discriminations seen in both ERPs demonstrated that language experience is a significant determinant of lexical tone processing. Statistical comparisons, supported by permutation-based multivariate analyses (PERMANOVA) and follow-up permutation two sample t-tests, revealed robust group differences in neural responses to tonal contrasts. As hypothesized, across both early (MMN) and later (LN) negative components, bilingual Mandarin-English children evoked more negative amplitudes to the deviant tones (see Figure 6), elucidating enhanced language experience effects on lexical tone processing. In contrast, bilingual Spanish-English children and monolingual English children—whose home languages lacked lexical tonal contrasts—showed attenuated MMR responses and reduced LN amplitudes. Hemispheric analysis further revealed greater left-dominant activation for bilingual Mandarin-English children during early discrimination, reflecting the combined influence of bilingual language experience and home language phonological properties on auditory processing of lexical representations.

1. MMR: Early Auditory Discrimination of Lexical Tone

The statistical analysis revealed that bilingual Mandarin-English children showed significantly larger MMN amplitudes than both monolingual English and bilingual Spanish-English groups at left and middle frontocentral electrodes (F3, Fz, C3, Cz). This finding indicates that bilingual Mandarin-English children demonstrated enhanced pre-attentive discrimination of pitch patterns—a key feature driven by their home tonal language experience. Early exposure to Mandarin tones—even in a bilingual environment with English—appears to heighten cortical sensitivity to pitch contrasts, as reflected in the enhanced early auditory processing seen in this study. These robust mismatch responses were predominantly lateralized to the left and middle scalp distributions, with weaker responses at right hemisphere electrodes (F4 and C4), consistent with left-hemispheric dominance for phonological and linguistic processing (Ries et al., 2016).

The observed hemispheric patterns align with prior evidence that speech sound discrimination, including non-segmental features like pitch, engages left-dominant auditory networks in individuals with native language experience (Preisig & Meyer, 2025). The reduced MMN amplitudes at right hemisphere sites among bilingual Mandarin-English children support the interpretation that lexical tone is processed linguistically, rather than purely acoustically, in children with tonal language exposure.

Tone-specific saliency patterns also emerged. Tone 1 (high level)—the acoustically easier contrast—consistently elicited stronger MMNs across groups, while Tone 2 (rising), being more similar to the standard Tone 3 (low dipping), evoked weaker or absent MMR. This contrast in discrimination strength reinforces prior evidence of acoustic saliency effects on early tone perception (Yu et al., 2017). In bilingual Spanish-English and monolingual English children, particularly, MMRs were often attenuated or absent for Tone 2, supporting the view that pitch-based contrasts are not automatically registered without prior exposure to tonal lexical systems.

These neurophysiological indices provide evidence that MMN amplitude reflects a home language effect: children with tonal language experience display more robust automatic auditory change detection to pitch contrasts than children exposed only to non-tonal languages. In practical terms, bilingual Mandarin-English children's brains prioritize pitch pattern changes as salient linguistic signals, even without conscious effort. In contrast, monolingual English children's brains—lacking an expectation for pitch to signal lexical meaning—show weaker automatic discrimination. Although Spanish is not a tonal language, it utilizes pitch for intonational pragmatics (Face, 2002). This reasoning might explain the small, though inconsistent, MMRs observed in bilingual Spanish-English children. However, significantly stronger MMN amplitudes were observed at Cz in monolingual English children compared to bilingual Spanish-English children. This implies that tonal language exposure, rather than bilingualism alone, plays a more important role in shaping auditory discrimination of lexical tones.

Further, when compared to adult MMN patterns reported in the literature, the robust and sharply localized mismatch responses observed in bilingual Mandarin-English children likely reflect early cortical specialization driven by language experience and early neuroplasticity, rather than differences in neural processing speed. Although adult MMNs are often broader and more distributed, consistent with less phonological commitment to tone contrasts (Shafer et al., 2004), the children's responses suggest that frequent early exposure to lexical tones facilitates the tuning of auditory cortical pathways for automatic phonemic discrimination.

2. LN: Later-Stage Cognitive Processing of Lexical Tone

The LN findings offer further insight into higher-order cognitive processing of tonal contrasts. Bilingual Mandarin-English children exhibited more negative LN amplitudes than both comparison groups at key frontal and midline electrodes (F3, Fz, Cz), for both Tone 1 and Tone 2 deviants. This suggests that bilingual Mandarin-English children may engage in more automatic processing of pitch differences due to their home language exposure (Strange, 2011), or alternatively, they may rely on the offset of the F0 contour to complete tone discrimination through a synthesized evaluation process (Burnham et al., 1986). These responses were most robust at Fz, consistent with enhanced engagement of top-down attentional networks for deviant detection in linguistically relevant pitch contrasts (Frith & Friston, 1996; Debener et al., 2003).

Monolingual English children showed intermediate LN responses, particularly at right frontal (F4) and central sites (C3, Cz), suggesting that while their early MMN discrimination was weaker, they recruited additional attentional and executive control mechanisms during later stages. Enhanced LN negativity at F4 among monolingual English children supports the interpretation that right-frontal attentional reorientation was engaged to process the deviant pitch patterns. This pattern suggests a right-hemispheric bias in non-tonal language experience children, likely reflecting recruitment of auditory neural systems in the right superior temporal and right pre-frontal cortices, that are typically involved in the evaluation of pitch and melodic changes that are not tied to linguistic meaning (Zatorre et al., 1994). Thus, in the absence of tonal language experience, pitch variations may be processed more similarly to prosodic information rather than lexical contrasts.

In contrast, bilingual Spanish-English children consistently showed the smallest LN amplitudes, with nearly attenuated waveforms across electrode sites and slightly positive deflections at C3 and F4, indicating minimal late-stage cognitive engagement with tonal contrasts. This pattern suggests limited evidence for a bilingual advantage in the absence of tonal input. An important limitation should be acknowledged. The observed variability in LN amplitudes at F4 and C3 for Spanish-English participants may reflect the small sample size (N = 4), and thus interpretations regarding the absence of a bilingual advantage should be made cautiously. While no bilingual advantage was observed among bilingual Spanish-English children relative to monolingual English children, the limited sample constrains the generalizability of these findings regarding bilingual language experience effects. Larger sample sizes are necessary to robustly assess whether bilingualism without tonal exposure can enhance auditory discrimination mechanisms.

Importantly, tone-specific patterns persisted into the LN time window. Across all groups, Tone 1 deviants elicited stronger LN responses than Tone 2 deviants, likely due to Tone 1's greater acoustic distance from the standard. This finding parallels the MMN results and reinforces the role of stimulus saliency in shaping neural responses at later stages of processing. The LN findings also point to bilingual language experience effects. Bilingual Mandarin-English children, exposed to both tonal and non-tonal languages, exhibited enhanced automatic discrimination (MMN) and greater later cognitive engagement (LN), whereas bilingual Spanish-English children, exposed to two non-tonal languages, did not exhibit enhanced neural responses. Thus, the specific type of bilingual language experience influences the development of auditory neural networks for tone processing.

When compared to adult LN patterns, which are often reduced, absent, or highly task-dependent (Yu et al., 2019), the LN observed here was also broader and more distributed. This consistent with previous findings that later-stage auditory discrimination and executive control processes remain under development across childhood due to cortical immaturity (Korpilahti et al., 2001; Shafer et al., 2005). These findings collectively suggest a developmental trade-off: tonal-exposed children engage both early automatic and late cognitive mechanisms more effectively, while non-tonal children show weaker early responses and more reliance on later additional attention processes, particularly when discriminating less salient Tone 2 contrasts.

CHAPTER 5: CONCLUSION

Overall, the data suggest that bilingual exposure to a tonal language facilitates both early automatic (MMN) and later-stage cognitive (LN) neurophysiological processing of pitch-based lexical contrasts in children. These effects were most pronounced at left and midline electrode sites (F3, Fz, Cz), highlighting stronger language-driven sensitivity in cortical areas typically associated with phonemic processing. This lateralized pattern suggests a neuroplastic adaptation to tonal input, where the auditory cortex of bilingual Mandarin-English children becomes attuned to pitch as a linguistically meaningful discrimination. The robust MMN and LN responses observed in bilingual Mandarin-English children delineate enhanced phonemic discrimination and greater sustained attention to tone changes, respectively, reflecting the impact of a home tonal language exposure.

In contrast, monolingual English children demonstrated intermediate neural responses, with overall attenuated MMR amplitudes and increased LN activity at right-frontal regions (e.g., F4), suggesting greater engagement of domain-general attentional networks during later auditory evaluation. This right-hemispheric engagement may reflect pitch-based analysis that is more prosodic than lexical, consistent with non-tonal language exposure.

Bilingual Spanish-English children, however, showed the weakest MMR and LN amplitudes, especially at C3 and F4, with LN responses being minimal or even positive in some cases. These findings suggest that there is no bilingual advantage without tonal language input.

Collectively, these findings reinforce three key theoretical insights: (1) the home language effect plays a critical role in shaping neural responses to speech, (2) the type of bilingualism exposure—driven by language experience—significantly impacts how children process lexical tone, and (3) the strongest group effects and acoustic salient effects occurred at left and midline electrodes, highlighting the influence of language experience over right-lateralized, non-linguistic processing routes.

Despite revealing clear neurophysiological markers of tone discrimination, the study employed a passive oddball paradigm and did not include any behavioral task of tone perception. As a result, interpretations regarding the relationship between neural indices and conscious perceptual abilities remain limited. Future studies should adopt multimodal approaches that combine electrophysiological measures with active behavioral tasks. For instance, integrating an active oddball paradigm or using a passive somatosensory MMN paradigm paired with behavioral responses—such as piloerector muscle activation combined with a mouse click to signal deviant detection—could offer a more comprehensive evaluation of both automatic and conscious tone processing. Additionally, further research should investigate tone perception in African tonal languages. These languages possess unique phonetic and articulatory properties that may provide critical insights into the evolutionary development of tonal systems and how the brain adapts to linguistically diverse pitch contrasts.

REFERENCES

Abramson, A. (1977). Noncategorical perception of tone categories in Thai | Request PDF. *The Journal of the Acoustical Society of America*, *61*(S1). https://doi.org/10.1121/1.2015837

Altmann, C. F., Uesaki, M., Ono, K., Matsuhashi, M., Mima, T., & Fukuyama, H. (2014). Categorical speech perception during active discrimination of consonants and vowels. *Neuropsychologia*, *64*, 13–23. https://doi.org/10.1016/j.neuropsychologia.2014.09.006

Bialystok, E. (1999). Cognitive Complexity and Attentional Control in the Bilingual Mind. *Child Development*, *70*(3), 636–644. https://doi.org/10.1111/1467-8624.00046

Burnham, D. K. (1986). Developmental loss of speech perception: Exposure to and experience with a first language. *Applied Psycholinguistics*, *7*(3), 207–239. https://doi.org/10.1017/S0142716400007542

Carreiras, M., & Price, C. J. (2008). Brain activation for consonants and vowels. *Cerebral Cortex (New York, N.Y.: 1991)*, *18*(7), 1727–1735. https://doi.org/10.1093/cercor/bhm202

Chen, C., Witte, M., Heemsbergen, W., & Herk, M. van. (2013). Multiple comparisons permutation test for image based data mining in radiotherapy. *Radiation Oncology*, *8*(1), 293. https://doi.org/10.1186/1748-717X-8-293

Chen, J., Antoniou, M., & Best, C. T. (2023). Phonological and phonetic contributions to perception of non-native lexical tones by tone language listeners: Effects of memory load and stimulus variability. *Journal of Phonetics*, *96*, 101199. https://doi.org/10.1016/j.wocn.2022.101199

Datta, H., Shafer, V. L., Morr, M. L., Kurtzberg, D., & Schwartz, R. G. (2010). Electrophysiological Indices of Discrimination of Long-Duration, Phonetically Similar Vowels in Children With Typical and Atypical Language Development. *Journal of Speech, Language, and Hearing Research*, *53*(3), 757–777. https://doi.org/10.1044/1092-4388(2009/08-0123)

Debener, S., Herrmann, C. S., Kranczioch, C., Gembris, D., & Engel, A. K. (2003). Top-down attentional processing enhances auditory evoked gamma band activity. *NeuroReport*, *14*(5), 683.

Duanma, S. (2007). *The Phonology of Standard Chinese (The Phonology of the World's Languages) - Duanmu, San; Duanma, San: 9780199215782 - AbeBooks*. https://www.abebooks.co.uk/9780199215782/Phonology-Standard-Chinese-Worlds-Languages-0199215782/plp

Face, T. (2002). Spanish evidence for pitch-accent structure. *Linguistics*, *40(2)*, 319–345. https://doi.org/10.1515/ling.2002.014

Frith, C. D., & Friston, K. J. (1996). The Role of the Thalamus in "Top Down" Modulation of Attention to Sound. *NeuroImage*, *4*(3), 210–215. https://doi.org/10.1006/nimg.1996.0072

Gandour, J. T., & Harshman, R. A. (1978). Crosslanguage differences in tone perception: A multidimensional scaling investigation. *Language and Speech*, *21*(1), 1–33. https://doi.org/10.1177/002383097802100101

Gonzales, K., & Lotto, A. J. (2013). A bafri, un pafri: Bilinguals' pseudoword identifications support language-specific phonetic systems. *Psychological Science*, *24*(11), 2135–2142. https://doi.org/10.1177/0956797613486485

Grundy, J. G., & Timmer, K. (2017). Bilingualism and Working Memory Capacity: A Comprehensive Meta-Analysis. *Second Language Research*, *33*(3), 325–340. https://doi.org/10.1177/0267658316678286

Hestvik, A., & Durvasula, K. (2016). Neurobiological evidence for voicing underspecification in English. *Brain and Language*, *152*, 28–43. https://doi.org/10.1016/j.bandl.2015.10.007

Holt, C. A., & Sullivan, S. P. (2023). Permutation tests for experimental data. *Experimental Economics*, *26*(4), 775–812. https://doi.org/10.1007/s10683-023-09799-6

Howie, J. M. (1976). *Acoustical Studies of Mandarin Vowels and Tones*. Cambridge University Press.

Kaan, E., Barkley, C. M., Bao, M., & Wayland, R. (2008). Thai lexical tone perception in native speakers of Thai, English and Mandarin Chinese: An event-related potentials training study. *BMC Neuroscience*, *9*(1), 53. https://doi.org/10.1186/1471-2202-9-53

Kappenman, E. S., & Luck, S. J. (2011). *ERP Components: The Ups and Downs of Brainwave Recordings*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0014

Kirmse, U., Ylinen, S., Tervaniemi, M., Vainio, M., Schröger, E., & Jacobsen, T. (2008). Modulation of the 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*(2), 131–143. https://doi.org/10.1016/j.ijpsycho.2007.10.012

Korpilahti, P., Krause, C. M., Holopainen, I., & Lang, A. H. (2001). Early and Late Mismatch Negativity Elicited by Words and Speech-Like Stimuli in Children. *Brain and Language*, *76*(3), 332–339. https://doi.org/10.1006/brln.2000.2426

Kuhl, B. A., Dudukovic, N. M., Kahn, I., & Wagner, A. D. (2007). Decreased demands on cognitive control reveal the neural processing benefits of forgetting. *Nature Neuroscience*, *10*(7), 908–914. https://doi.org/10.1038/nn1918

Kwok, V. P. Y., Dan, G., Yakpo, K., Matthews, S., & Tan, L. H. (2016). Neural systems for auditory perception of lexical tones. *Journal of Neurolinguistics*, *37*, 34–40. https://doi.org/10.1016/j.jneuroling.2015.08.003

Ladefoged, P., & Johnson, K. (2011). A course in phonetics (6th ed.). Wadsworth/Cengage Learning. https://theswissbay.ch/pdf/Books/Linguistics/A%20Course%20in%20Phonetics%206th%20Edition%20-%20Peter%20Ladefoged,%20Keith%20Johnson.pdf

Lee, W.-S. (2003). A Phonetic Study of the Neutral Tone in Beijing Mandarin. *International Congress of Phonetic Sciences*, 1121–1124.

Light, G. A., Williams, L. E., Minow, F., Sprock, J., Rissling, A., Sharp, R., Swerdlow, N. R., & Braff, D. L. (2010). Electroencephalography (EEG) and event-related potentials (ERPs) with human participants. *Current Protocols in Neuroscience*, *Chapter 6*, Unit 6.25.1-24. https://doi.org/10.1002/0471142301.ns0625s52

Lindín, M., Correa, K., Zurrón, M., & Díaz, F. (2013). Mismatch negativity (MMN) amplitude as a biomarker of sensory memory deficit in amnestic mild cognitive impairment. *Frontiers in Aging Neuroscience*, *5*, 79. https://doi.org/10.3389/fnagi.2013.00079

Liu, H.-M., Chen, Y., & Tsao, F.-M. (2014). Developmental Changes in Mismatch Responses to Mandarin Consonants and Lexical Tones from Early to Middle Childhood. *PLOS ONE*, *9*(4), e95587. https://doi.org/10.1371/journal.pone.0095587

Lu, J., Li, Y., Zhao, Z., Liu, Y., Zhu, Y., Mao, Y., Wu, J., & Chang, E. F. (2023). Neural control of lexical tone production in human laryngeal motor cortex. *Nature Communications*, *14*, 6917. https://doi.org/10.1038/s41467-023-42175-9

Myers, S. (2021). (PDF) The intonation of yes–no questions in Luganda. *Journal of the International Phonetic Association*, *53*(1), 1–35. https://doi.org/10.1017/S0025100321000025

Näätänen, R., Kujala, T., & Light, G. (2019). *Mismatch Negativity: A Window to the Brain*. Oxford University Press.

Näätänen, R., Kujala, T., & Winkler, I. (2011). Auditory processing that leads to conscious perception: A unique window to central auditory processing opened by the mismatch negativity and related responses. *Psychophysiology*, *48*(1), 4–22.

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*(6615), 432–434. https://doi.org/10.1038/385432a0

Olulade, O. A., Jamal, N. I., Koo, D. S., Perfetti, C. A., LaSasso, C., & Eden, G. F. (2016). Neuroanatomical Evidence in Support of the Bilingual Advantage Theory. *Cerebral Cortex (New York, N.Y.: 1991)*, *26*(7), 3196–3204. https://doi.org/10.1093/cercor/bhv152

Ortiz-Mantilla, S., Choudhury, N., Alvarez, B., & Benasich, A. 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. https://doi.org/10.1016/j.brainres.2010.09.031

Paap, K. R., Anders-Jefferson, R., Mason, L., Alvaradk, K., & Zimiga, B. (2018, August 14). *Frontiers | Bilingual Advantages in Inhibition or Selective Attention: More Challenges*. Frontiers. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2018.01409/full

Phillips, C. (2001). Levels of representation in the electrophysiology of speech perception. *Cognitive Science*, *25*(5), 711–731. https://doi.org/10.1207/s15516709cog2505_5

Plakke, B., & Romanski, L. M. (2014). Auditory connections and functions of prefrontal cortex. *Frontiers in Neuroscience*, *8*. https://doi.org/10.3389/fnins.2014.00199

Preisig, B. C., & Meyer, M. (2025). Predictive coding and dimension-selective attention enhance the lateralization of spoken language processing. *Neuroscience & Biobehavioral Reviews*, *172*, 106111. https://doi.org/10.1016/j.neubiorev.2025.106111

Ren, G.-Q., Yang, Y., & Li, X. (2009). Early cortical processing of linguistic pitch patterns as revealed by the mismatch negativity. *Neuroscience*, *162*(1), 87–95. https://doi.org/10.1016/j.neuroscience.2009.04.021

Ries, S. K., Dronkers, N. F., & Knight, R. T. (2016). Choosing words: Left hemisphere, right hemisphere, or both? Perspective on the lateralization of word retrieval. *Annals of the New York Academy of Sciences*, *1369*(1), 111–131. https://doi.org/10.1111/nyas.12993

Rivera-Gaxiola, M., Csibra, G., Johnson, M. H., & Karmiloff-Smith, A. (2000). Electrophysiological correlates of cross-linguistic speech perception in native English speakers. *Behavioural Brain Research*, *111*(1), 13–23. https://doi.org/10.1016/S0166-4328(00)00139-X

Rivera-Gaxiola, M., Johnson, M. H., Csibra, G., & Karmiloff-Smith, A. (2000). Electrophysiological correlates of category goodness. *Behavioural Brain Research*, *112*(1), 1–11. https://doi.org/10.1016/S0166-4328(00)00218-7

Shafer, V. L., Morr, M. L., Datta, H., Kurtzberg, D., & Schwartz, R. G. (2005). Neurophysiological Indexes of Speech Processing Deficits in Children with Specific Language Impairment. *Journal of Cognitive Neuroscience*, *17*(7), 1168–1180. https://doi.org/10.1162/0898929054475217

Shafer, V. L., Morr, M. L., Kreuzer, J. A., & Kurtzberg, D. (2000). Maturation of Mismatch Negativity in School-Age Children. *Ear and Hearing*, *21*(3), 242.

Shafer, V. L., Schwartz, R. G., & Kurtzberg, D. (2004). Language-specific memory traces of consonants in the brain. *Cognitive Brain Research*, *18*(3), 242–254. https://doi.org/10.1016/j.cogbrainres.2003.10.007

Shi, B., Zhou, Q., & Li, J. (2024). (PDF) Influence of Dialect on the English Consonants Perception by English L2 learners: Evidence from Jiangsu Province for Chinese Native Speakers. *Communications in Humanities Research*, *30*(1), 29–36.

Singh, L., Loh, D., & Xiao, N. G. (2017). Bilingual Infants Demonstrate Perceptual Flexibility in Phoneme Discrimination but Perceptual Constraint in Face Discrimination. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01563

Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech: A working model. *Journal of Phonetics*, *39*(4), 456–466. https://doi.org/10.1016/j.wocn.2010.09.001

Tan, J., Dowell, R., & Vogel, A. (2016). Mandarin Lexical Tone Acquisition in Cochlear Implant Users With Prelingual Deafness: A Review. *American Journal of Audiology*, *25*(3), 246–256. https://doi.org/10.1044/2016_AJA-15-0069

Wang, X., Fu, R., Xia, X., Chen, X., Wu, H., Landi, N., Pugh, K., He, J., & Cong, F. (2018). Spatial Properties of Mismatch Negativity in Patients with Disorders of Consciousness. *Neuroscience Bulletin*, *34*(4), 700–708. https://doi.org/10.1007/s12264-018-0260-4

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development*, *7*(1), 49–63. https://doi.org/10.1016/S0163-6383(84)80022-3

Werwach, A., Männel, C., Obrig, H., Friederici, A. D., & Schaadt, G. (2022). Longitudinal trajectories of electrophysiological mismatch responses in infant speech discrimination differ across speech features. *Developmental Cognitive Neuroscience*, *56*, 101127. https://doi.org/10.1016/j.dcn.2022.101127

Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience*, *170*(1), 223–231. https://doi.org/10.1016/j.neuroscience.2010.06.077

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, *25*(1), 61–83. https://doi.org/10.1006/jpho.1996.0034

Yip, M. (2002). *Tone*. Cambridge University Press. https://doi.org/10.1017/CBO9781139164559

Yu, Y. H., Shafer, V. L., & Sussman, E. S. (2017a). Neurophysiological and Behavioral Responses of Mandarin Lexical Tone Processing. *Frontiers in Neuroscience*, *11*. https://doi.org/10.3389/fnins.2017.00095

Yu, Y. H., Shafer, V. L., & Sussman, E. S. (2017b). Neurophysiological and Behavioral Responses of Mandarin Lexical Tone Processing. *Frontiers in Neuroscience*, *11*. https://doi.org/10.3389/fnins.2017.00095

Yu, Y. H., Shafer, V. L., & Sussman, E. S. (2018). The Duration of Auditory Sensory Memory for Vowel Processing: Neurophysiological and Behavioral Measures. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00335

Yu, Y. H., Tessel, C., Han, X., Campanelli, L., Vidal, N., Gerometta, J., Garrido-Nag, K., Datta, H., & Shafer, V. L. (2019). Neural Indices of Vowel Discrimination in Monolingual and Bilingual Infants and Children. *Ear and Hearing*, *40*(6), 1376–1390. https://doi.org/10.1097/AUD.0000000000000726

Zatorre, R. J., Evans, A. C., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience*, *14*(4), 1908–1919. https://doi.org/10.1523/JNEUROSCI.14-04-01908.1994

Vita

Name	*Andres Felipe Diaz*
Baccalaureate Degree	*Bachelor of Arts, CUNY Queens College, New York City, Major: Biology*
Date Graduated	*September, 2021*