2019

Neural Indices of Vowel Discrimination in Monolingual and Bilingual Infants and Children

Yan H. Yu
CUNY Graduate Center, St. John’s University

Carol Tessel
CUNY Graduate Center, Florida Atlantic University

Henry Han
Fordham University

Luca Campanelli
CUNY Graduate Center

Nancy Vidal
CUNY Graduate Center, Iona College

See next page for additional authors

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

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

Part of the Developmental Neuroscience Commons, Developmental Psychology Commons, First and Second Language Acquisition Commons, Phonetics and Phonology Commons, and the Speech and Hearing Science Commons

Recommended Citation

This Article is brought to you by CUNY Academic Works. It has been accepted for inclusion in Publications and Research by an authorized administrator of CUNY Academic Works. For more information, please contact AcademicWorks@gc.cuny.edu.
Neural Indices of Vowel Discrimination in Monolingual and Bilingual Infants and Children

Yan H. Yu1, Carol Tessel2, Henry Han3, Luca Campanelli4,5, Nancy Vidal6, Jennifer Gerometta6, Karen Garrido-Nag7, Hia Datta8, and Valerie L. Shafer4

1Department of Communication Sciences and Disorders, St. John’s University, Queens, New York, USA; 2Department of Communication Sciences and Disorders, Florida Atlantic University, Boca Raton, Florida, USA; 3Department of Computer and Information Science, Fordham University, New York, New York, USA; 4Department of Speech-Language-Hearing Sciences, The Graduate Center, City University of New York, New York, New York, USA; 5Haskins Laboratories, 300 George Street, New Haven, Connecticut, USA; 6Speech Communication Studies, Iona College, New Rochelle, New York, USA; 7Hearing, Speech, Language Sciences, Gallaudet University, Washington DC, USA; 8Speech-Language Pathology, Molloy College, Rockville Centre, New York, USA.

Abstract

Objectives: To examine maturation of neural discriminative responses to an English vowel contrast from infancy to 4 years of age and to determine how biological factors (age and sex) and an experiential factor (amount of Spanish versus English input) modulate neural discrimination of speech.

Design: Event-related potential (ERP) mismatch responses (MMRs) were used as indices of discrimination of the American English vowels [ε] versus [I] in infants and children between 3 months and 47 months of age. A total of 168 longitudinal and cross-sectional data sets were collected from 98 children (Bilingual Spanish–English: 47 male and 31 female sessions; Monolingual English: 48 male and 42 female sessions). Language exposure and other language measures were collected. ERP responses were examined in an early time window (160 to 360 msec, early MMR [eMMR]) and late time window (400 to 600 msec, late MMR).

Results: The eMMR became more negative with increasing age. Language experience and sex also influenced the amplitude of the eMMR. Specifically, bilingual children, especially bilingual females, showed more negative eMMR compared with monolingual children and with males. However, the subset of bilingual children with more exposure to English than Spanish compared with those with more exposure to Spanish than English (as reported by caretakers) showed similar amplitude of the eMMR to their monolingual peers. Age was the only factor that influenced the amplitude of the late MMR. More negative late MMR was observed in older children with no difference found between bilingual and monolingual groups.

Conclusions: Consistent with previous studies, our findings revealed that biological factors (age and sex) and language experience modulated the amplitude of the eMMR in young children. The early negative MMR is likely to be the mismatch negativity found in older children and adults. In contrast, the late MMR amplitude was influenced only by age and may be equivalent to the Nc in infants and to the late negativity observed in some auditory passive oddball designs.

Keywords: Bilingual, Event-related potentials, Infants, Late negativity, Mismatch negativity, Mismatch response, Speech development
Introduction

Considerable research has tracked the time course of first-language speech perception development and has revealed that experience with a language influences how an infant categorizes speech from birth (and even in the womb) through early childhood, typically measured via behavioral preference and novelty tasks (Best et al. 1982; Kuhl et al. 1992; DeCasper et al. 1994; Kuhl 2004). By one year of age, children have learned which phonological patterns are contrastive (phonemes) in their native language and they have already shown sensitivity to language-specific phonological information (e.g., Werker & Tees 1983; Jusczyk & Luce 1994). However, phonological development is incomplete well into the grade-school years (Nitttrouer & Miller 1997). For example, one-year-olds will fail to discriminate difficult phoneme contrasts (e.g., /ba/ and /da/ in more challenging tasks such as lexical learning) until 18 months of age, even though they demonstrate discrimination in a simpler task (e.g., Werker et al. 2002).

Both experience and auditory cortical maturation are likely to be necessary to attain an adult-like phonological system. Studies of speech perception development have largely focused on the former variable. More recent models have acknowledged the importance of considering the developmental level of a child when interpreting speech perception results (Werker & Curtin 2005; Curtin et al. 2011). The goal of the current study is to examine both experience and maturation in the development of speech discrimination.

Studies of neural maturation of the auditory system have revealed considerable cortical immaturity before 6 months of age (e.g., Kostović & Jovanov-Milosević 2006; Moore & Linthicum 2007). Considerable structural maturation (e.g., number of axon collaterals and amount of myelination) and functional maturation (e.g., latency and topography of obligatory, auditory components) occur during the first year of life, with the rate of maturation slowing down through the preschool and into the grade-school years (Moore & Lithicum 2007). Thus internal, biological factors constrain the rate of auditory cortical maturation in neonates and young infants (Leppänen et al. 2004; Kostović & Jovanov-Milosević 2006; Moore & Lithicum 2007). Studies of children with congenital deafness followed by cochlear implant indicate that early input also plays a role in maturation of the auditory system (Sharma et al. 2002). What is less clear is the degree to which the characteristics of the input (e.g., monolingual versus bilingual speech input) modulate auditory development.

An additional source of evidence for biological factors constraining maturation comes from studies of sex differences. Sex differences during prenatal and early infancy periods have been observed (for basic sensory difference: see Alexander & Wilcox 2012; for a review, Oral-motor function: Miller et al. 2006; Nagy et al. 2007). Male and female infants show different rates of cortical maturation driven by differing levels of sex hormones (e.g., Shucard et al. 1981; Shucard & Shucard 1990; Friederici et al. 2008). In particular, female infants initially showed more rapid maturation of electrophysiological responses to tones over left compared with right hemisphere sites and male infants showed the opposite pattern (Shucard & Shucard 1990); furthermore, one study found that female infants generally showed more linguistic rule-based learning than male infants (Mueller et al. 2012). Irrespective of the exact underlying sources of these sex differences related to cortical maturation, variations in cortical maturation can be inferred from these findings of sex differences.

Cross linguistic studies of speech perception have shown that listeners often have difficulty discriminating and categorizing non-native phonological contrasts not found in their
first language (L1) (e.g., Werker & Tees 1984; Best 1985). In the case of second-language (L2) learning, late learners of an L2 (after approximately 11 years of age) may show poor L2 speech perception (e.g., Strange 2011). In contrast, early learning of an L2 results in better behavioral perception of L2 phonological contrasts (e.g., Flege 1995; Sundara & Polka 2008; Gonzales & Lotto 2013). Some studies, however, suggest that even early learning of an L2 (before 5 years of age) may result in less robust brain discrimination of L2 speech contrasts, when compared with monolingual listeners for some (e.g., Hisagi et al. 2014), but not all early bilinguals (Datta et al. 2019). Understanding the factors that influence L2 speech perception in adult bilinguals who had learned the L2 early in life is a challenging task. Previous studies have shown that amount of input in L1 versus L2 is likely to be a major factor (Flege 1995; Flege & Mackay 2004). Thus, examining L2 speech perception in the time period when amount of input can be better evaluated is desirable (Datta et al. 2019).

Behavioral studies of bilingual language development in infants have shown that speech processing in one or both of a bilingual’s languages can be delayed if the two languages are difficult to separate (e.g., prosodically similar languages such as Spanish and Catalan) (Bosch & Sebastián-Gallés 2003; Sebastián-Gallés & Bosch 2009). In cases that the languages (e.g., English and French) are easier to separate, the time course of speech perception development for both languages appears to be somewhat similar for bilinguals and monolinguals (Sundara et al. 2008). However, in a previous event-related potential (ERP) study, we found a small but significant effect of bilingual exposure on vowel speech processing in children between 3 months and 36 months for two languages (Spanish and English) that should be relatively easy to separate (Shafer et al. 2011, 2012). Therefore, the relationship between language similarity and bilingual development may be less straightforward.

In addition to language similarity, the amount of language exposure in each language also affects speech and language development in young children (Conboy & Mills 2006; Hoff et al. 2012). Hoff et al. (2012) reported that the rate of development of each language in bilingual children is a function of their relative amount of language exposure. Conboy and Mills (2006) also found a large effect of language experience on electrophysiological measures to words in the dominant versus nondominant languages of Spanish–English toddlers. Recently, Cattani et al. (2014) reported that 60% or more English exposure will lead to monolingual-like performance on English language tests in bilingual 6-year-old children.

There has been an increased interest in examining the neurodevelopment of speech processing in bilingual children in the past 15 years (e.g., Rivera-Gaxiola et al. 2005; Shafer et al. 2010, 2012, 2015; Ferjan Ramirez et al. 2017; Garcia-Sierra et al. 2016; Rinker et al. 2017). Most studies have used an auditory oddball paradigm, in which a frontocentral neural response called mismatch negativity (MMN) is used to measure auditory processing (Näätänen et al. 1997). The mature MMN is a negative peak between 100 and 250 msec poststimulus onset. Children older than approximately 4 years of age show somewhat similar neural measures of sound discrimination to adults but sometimes at later latencies (Gomes et al. 1999; Shafer et al. 2000; Shafer et al. 2010). For example, Shafer et al. (2010) observed an MMN in 4- to 7-year-old children to a speech sound contrast; this MMN peaked between 200 and 400 msec in the children, rather than between 100 and 250 msec, as expected for adults. The MMN shifts earlier in latency with increasing age, but does not reach adult latencies for more difficult sound contrasts until past puberty for some auditory and speech contrasts (Ponton et al. 2000; Shafer et al. 2000; Shafer et al. 2010).
Infants typically show a slow positive mismatch response (pMMR) in a passive oddball paradigm (Dehaene-Lambertz & Dehaene 1994; Dehaene-Lambertz & Baillet 1998; Trainor et al. 2001; Kushnerenko et al. 2002; Morr et al. 2002; Leppänen et al. 2004; Friederici et al. 2007; Lee et al. 2012; Shafer et al. 2012; Partanen et al. 2013a; Partanen et al. 2013b). However, a few studies have observed a negative MMR (nMMR) rather than a pMMR (e.g., Cheour et al. 1998; Trainor et al. 2001; Kushnerenko et al. 2002; He et al. 2007, 2009). With increasing age, the pMMR declines in amplitude and the nMMR emerges, peaking between 100 and 400 msec. The pMMR and nMMR may co-occur and partially overlap in infants and young children (e.g., Leppänen et al. 1997; Morr et al. 2002). It is likely that the timeframe for the emergence of the nMMR is dependent on the degree of acoustic or phonetic difference for the stimulus contrasts. For example, the nMMR emerges after 3 years of age to the subtle vowel contrast /I/ versus /ε/ (Shafer et al. 2010), but to a large acoustic difference, such as 1000 versus 2000 Hz, the nMMR is observed in infants under 1 year of age (Morr et al. 2002; see Kushnerenko et al. 2013, for review). The pMMR continues to be observed (preceding the nMMR, if it co-occurs) to somewhat more difficult contrasts into the grade-school ages (Shafer et al. 2010; Lee et al. 2012; Liu et al. 2014).

The nMMR may be equivalent to the MMN observed in adults and older children. The MMN indicates a mismatch between a stimulus or stimulus pattern (deviant) and the predictions generated by the prior context (standards) (Kushnerenko et al. 2013). The nature of pMMR is less clear. Some studies have suggested that the pMMR is an analog of adult P3a, indexing involuntary attention switch to stimulus change (Alho et al. 1990; Trainor et al. 2001; Kushnerenko et al. 2002). Others suggest that it reflects “an automatic categorization of the stimulus” (Friedrich et al. 2004), or an index of a general process as a result of neural adaptation (He et al. 2009). Recent studies have found larger pMMR to more difficult than to easier stimulus contrasts (e.g., Lee et al. 2012); these findings are more consistent with one of the latter two explanations for the pMMR. There is also evidence that a much later positive component (PC) peaking later than 300 msec and observed to novel sounds is the precursor of the adult P3a response, rather than the earlier occurring pMMR (Kushnerenko et al. 2013).

Infants, children, and adults may also show a second negative discriminative peak at superior anterior scalp sites. In an early study with adults, this response was called the reorienting negativity (Schroger & Wolff 1998); in studies with children, a similar response has been called the late discriminative negativity (e.g., Cheour et al. 2001), the late MMN, the MMN2 (Korpilati et al. 2001), or the late negativity (LN) (Shafer et al. 2005; Yu et al. 2017; Datta et al. 2019; Yu et al. 2018). The Nc in infants may be comparable to this LN in children and adults. The Nc is found in a fairly late time interval (500 to 600 msec), and has been suggested to reflect enhanced attention to unexpected auditory or visual change (Kurtzberg et al. 1984; Courschesne 1990; Kushnerenko et al. 2013).

Differences in the LN timing and topography, as well as different findings across studies, currently limit our understanding of what the LN reflects. Some studies suggest that the LN is modulated by phonological or lexical aspects of processing. For example, Korpilati and colleagues (2001) reported that a late MMN was significantly larger for deviant words than deviant tones, and suggested that it was related to automatic detection of lexical difference. Yu et al. (2017) found that the LN amplitude to lexical tone contrast was larger in native Mandarin listeners than monolingual English listeners when presented at a relatively long interstimulus interval (ISI, at the rate of approximately 3 sec). In other studies, the LN was related to language status (language-impaired versus typically-developing) (Bishop et al. 2010; Datta et al. 2010).
Alternatively, the LN may reflect enhanced auditory attention, considering that it is also observed to nonspeech auditory stimuli and is modulated by novelty (Kushnerenko et al. 2013). We observed positive responses in infants and younger children in the LN time window, and this positive response gradually became more negative. We therefore use the term “late MMR” to refer to the responses in the later time window for the remainder of the article. We refer to the MMR in the time window of 160 to 360 msec as early MMR (eMMR). Depending on age and language background, this eMMR may be positive or negative. Other studies have used the term eMMN to refer to early MMN in this time window (e.g., Choudhury et al. 2015).

The present study examines ERP correlates of processing the American English vowel contrast [ɛ] versus [i] in monolingually-exposed and bilingually-exposed infants and children between 3 and 47 months of age. The goals of this study were as follows: (1) to examine how level of maturity modulates the neurodevelopment of speech processing throughout the period in which the child is mastering language (infancy to 4 years of age); and (2) how age, sex, and language experience modulate ERPs associated with speech-sound discrimination. We hypothesized that even a minimal amount of experience with an L2 would influence speech processing in a young child. We also expected to find age-related modulation of the ERPs. In particular, we expected to see increased negativity in both early and late MMR amplitude with increasing age (Shafer et al. 2010). Finally, we predicted that sex would influence the early and late MMRs but more so during the first year of life than at later ages.

Materials and methods

The experiment reported here was approved by The Graduate Center, CUNY Institutional Review Board, and was conducted in compliance with the Declaration of Helsinki. Parental consent was obtained from the parent/legal guardian(s) for each participant.

Participants

Fifty-nine bilingual Spanish–English children and 59 monolingual English children were recruited. Ten children from the bilingual group were not included in this study for the following reasons: five refused to be tested (no data), two had too few trials, two had corrupt data, and one had missing data files. Of the remaining 49 bilingually exposed children, 30 were male and 19 were female. Seventy-eight data sessions were collected with 16 (nine males, seven females) of these bilingual-exposed children being tested two to five times at different ages in a longitudinal design. Ten children from the monolingual English group were not included for the following reasons: three were tested in a different paradigm, three refused to be tested, two showed indications of language impairment, and two had missing or corrupt data files. Of the remaining 49 monolingual children, 26 were male and 23 were female. Ninety data sessions were collected with 20 (10 males, 10 females) of the monolingual-exposed children being tested two to six times at different ages in a longitudinal design.

A total of 168 data sets were collected from the 98 children (Bilingual: 47 male and 31 female sessions; Monolingual: 48 male and 42 female sessions; Age: 3 to 5, 6 to 12, 13 to 24, and 25 to 47 months old). Table 1 presents a summary of the number of ERP sessions in terms of participant age, sex, and language for the cross-sectional and longitudinal data combined. There was no session number difference in terms of the language or sex factor ($\chi^2 = 0.558$, $p = 0.46$). An additional 11 bilingual sessions and 17 monolingual sessions are included here compared with our previous report (Shafer et al. 2011). In addition, the previous report did not include the
standardized language measures, or the subtypes of bilingual exposure (more English versus less English), and it examined maturation in a cross-sectional design. Shafer et al. (2012) focused exclusively on a comparison of the bilingual and monolingual 6- to 8-month-old infants.

Table 1. A total of 168 sessions from cross-sectional and longitudinal sessions combined by age (mos) and sex (M, male; F, female).

<table>
<thead>
<tr>
<th>Age (mos)</th>
<th>3–5</th>
<th>6–12</th>
<th>13–24</th>
<th>25–47</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monolingual</td>
<td>3M, 7F</td>
<td>16M, 17F</td>
<td>16M, 10F</td>
<td>13M, 8F</td>
<td>48M, 42F</td>
</tr>
<tr>
<td>Longitudinal</td>
<td>(3M, 4F)</td>
<td>(10M, 12F)</td>
<td>(12M, 9F)</td>
<td>(7M, 5F)</td>
<td>(32M, 30F)</td>
</tr>
<tr>
<td>Cross-sectional</td>
<td>(0M, 3F)</td>
<td>(6M, 5F)</td>
<td>(4M, 1F)</td>
<td>(6M, 3F)</td>
<td>(18M, 12F)</td>
</tr>
<tr>
<td>Bilingual</td>
<td>9M, 1F</td>
<td>15M, 16F</td>
<td>12M, 7F</td>
<td>11M, 7F</td>
<td>47M, 31F</td>
</tr>
<tr>
<td>Longitudinal</td>
<td>(4M, 0F)</td>
<td>(7M, 9F)</td>
<td>(9M, 4F)</td>
<td>(7M, 4F)</td>
<td>(27M, 17F)</td>
</tr>
<tr>
<td>Cross-sectional</td>
<td>(5M, 1F)</td>
<td>(8M, 7F)</td>
<td>(3M, 3F)</td>
<td>(4M, 3F)</td>
<td>(20M, 14F)</td>
</tr>
</tbody>
</table>

Note. Among a total of 98 children, 20 monolingual and 16 bilingual children were tested two to six times between 3 and 47 months of age. The sessions from the longitudinal and cross-sectional children are in parentheses.

Language input was estimated from a language background questionnaire (LBQ) filled out by the principal caretaker, using a seven-point scale for rating input across multiple persons (e.g., father, mother, babysitter, siblings) and situations (e.g., home, playground), with 3 indicating all Spanish (labeled on the questionnaire as “Spanish all the time”), −3 indicating all English (labeled on the questionnaire as “English all the time”), and 0 indicating equal Spanish and English input (labeled on the questionnaire as “equal English and Spanish”). Note that on the questionnaire the values 1 to 7 were used and were then transformed to a scale of −3 to 3 for ease of interpretation. None of the bilingual children had a language background score of 3 (indicating all Spanish exposure) or −3 (indicating all English exposure). Twenty-eight out of 49 children had relatively balanced English–Spanish exposure with an LBQ score between the values of −1 and 1. In one descriptive analysis and one mixed model regression analysis, we divided the bilingual children into two subcategories based upon LBQ scores (see Table 2). A total of 20 (eight females) children were categorized as Spanish-dominant with an LBQ score of 0 to 2.9 and a mean of 0.62 (SD = 0.80), and the remainder (N = 29, 11 females) were categorized as English-dominant with an LBQ score between −0.1 and −2.9, mean −1.47 (SD = 0.76). The mean LBQ score for all the bilingual children was −0.62 (SD = 1.3).

Language measures were obtained using the Preschool Language Scale-4 (PLS-4) in English and, for bilinguals, in both English and Spanish. Three-year-old children also received the English Peabody Picture Vocabulary Test-3 (Dunn & Dunn 1997) and, for bilinguals, the Spanish Test de Vocabulario en Imagenes Peabody (Dunn & Dunn 1986). Table 3 shows standard language scores on the PLS-4 (Zimmerman et al. 2002) by language group (monolingual and bilingual, Spanish-dominant versus English-dominant bilinguals). All children had language measures on the English tests; Spanish measures were available for 16 of the 20 Spanish-dominant children and for 19 of the 29 English-dominant bilingual children. Note that for children under 12 months of age, the test items on the PLS-4 reflected general communication development (until first words are reported), and thus we did not collect measures for both the English and Spanish versions of the PLS at this age. The only difference in these two tests on
early development is whether the question to the caretaker is asked in English or Spanish. Of the bilingual children lacking Spanish scores, only four of these children were older than 9 months of age. In these cases, absence of Spanish scores was due to failure to complete the Spanish testing due to fatigue.

Table 2. Bilingual sessions were further divided into more English (BME) and less English (BLE) subgroups in some analyses.

<table>
<thead>
<tr>
<th></th>
<th>3–5 mos</th>
<th>6–12 mos</th>
<th>13–24 mos</th>
<th>25–47 mos</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>BME</td>
<td>3M, 0F</td>
<td>5M, 5F</td>
<td>4M, 4F</td>
<td>6M, 3F</td>
<td>30</td>
</tr>
<tr>
<td>BLE</td>
<td>6M, 1F</td>
<td>10M, 11F</td>
<td>8M, 3F</td>
<td>5M, 4F</td>
<td>48</td>
</tr>
</tbody>
</table>

Table 3. Mean (SD) language scores for monolinguals and bilinguals.

<table>
<thead>
<tr>
<th>Group</th>
<th>English</th>
<th>English</th>
<th>English</th>
<th>Spanish</th>
<th>Spanish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monolingual</td>
<td>Lang-all</td>
<td>PLS_A</td>
<td>PLS_E</td>
<td>PLS_A</td>
<td>PLS_E</td>
</tr>
<tr>
<td>Bilingual</td>
<td>100 (12)</td>
<td>98 (16)</td>
<td>100 (12)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Bilingual total</td>
<td>100 (14)</td>
<td>100 (16)</td>
<td>102 (14)</td>
<td>101 (20)</td>
<td>96 (15)</td>
</tr>
<tr>
<td>Bil-Eng</td>
<td>100 (16)</td>
<td>100 (19)</td>
<td>103 (16)</td>
<td>94 (22)</td>
<td>94 (16)</td>
</tr>
<tr>
<td>Bil-Span</td>
<td>100 (11)</td>
<td>98 (15)</td>
<td>101 (12)</td>
<td>105 (12)</td>
<td>103 (12)</td>
</tr>
</tbody>
</table>

Note. Subgroups of bilinguals were included. PLS-A is the Auditory Comprehension and PLS-E is the Expressive skills. English Lang-all is the mean of the PLS-4 scores and the PPVT-3 (for older children). Bilingual-English (Bil-Eng) children had mean exposure scores on the LBQ between −0.6 and −2.9, whereas Bilingual-Spanish (Bil-Span) children had scores between −0.59 and 2.9.

In the main statistical analyses, we used language as a categorical measure (monolingual versus bilingual; monolingual, balanced bilingual versus Spanish-dominant bilingual) because of the following reasons: (1) the amount of English exposure does not follow a linear distribution since all monolingual children have 100% of English; (2) the monolingual and two bilingual groups of children did not significantly differ on the English language tests for any pairwise comparison between monolinguals and bilinguals in the longitudinal and cross-sectional groups, or between English- versus Spanish-dominant groups (t-tests, p > 0.20). The only differences were between English-dominant versus Spanish-dominant groups on the Spanish language tests. Unsurprisingly, bilingual-exposed children with relatively greater exposure to Spanish (0 to 2.9 on the LBQ) showed significantly higher receptive Spanish language standard scores than those with less Spanish exposure (−0.1 to 2.9; p = 0.05 for receptive language, and p = 0.11 for expressive language).

We used age as a continuous variable for the main statistical analyses. We also used age as a categorical variable in the t-tests, figures, and some tables for easy summary of participant characteristics. In some analyses, we divided children into the following age groups: 3 to 5 months, 6 to 12 months, 13 to 24 months, and 25 to 47 months. Our rationale for these four groups was derived from literature on early brain development and behavioral speech discrimination development. Moore and Linthicum (2007) divide early brain development into five distinct stages: transitioning to perinatal (27 to 29 fetal weeks), perinatal (third trimester to
sixth postnatal month), transition to childhood (6 to 12 months), early childhood (2 to 5 years of age), and later childhood (6 to 12 years of age). Up to 6 months of age, the auditory cortex is still quite immature. Behavioral studies indicate that by 6 months of age, monolingual infants already show language-specific modulation of vowel discrimination (Eilers et al. 1979; Kuhl et al. 1992). Between 6 months and 1 year of age, there is a marked reduction of axons in the marginal layer, and filament-filled axons become visible in the auditory radiations in the core of the temporal lobe and the deeper layers of the cortex. Behavioral studies indicate language-specific discrimination effects of more fine-grained speech information (Werker & Tees 1984; Polka & Werker 1994) and differences between monolinguals and bilinguals in vowel perception emerge during this period (Bosch & Sebastián-Gallés 2003). From 13 to 24 months is a time of rapid lexical growth, whereas from 24 to 48 months children are expected to acquire the grammatical patterns of language (Werker et al. 2005).

The Bayley Scales of Infant and Toddler Development (Bayley 1993) was used as a measure of general cognitive development. A t-test showed no group differences (monolingual mean = 103, SD = 15; bilingual mean = 98, SD = 17, p = 0.14). All infants were full-term, and had a normal birth history with no history of cognitive, neurological, speech, language, or hearing deficits in immediate family members. All had passed a newborn hearing screening according to parental report, and most passed a transient-evoked otoacoustic emissions (TOAE) hearing screening in the laboratory. We did not have records of TOAE for several children due to incomplete behavioral testing or participant lack of compliance. All infants in the analyses showed a clearly defined P1 obligatory auditory-evoked potential at frontocentral sites, indicating encoding of the auditory sound (Sharma et al. 2009). Socio-economic status (SES) information was collected for all children using the Hollingshead Four-Factor Index of Socioeconomic Status (Hollingshead 1975). Most of the children came from families with a middle-class or higher SES, but the monolingual group had a higher average SES than the bilingual group (Monolingual: mean = 50.7, SD = 12.1; Bilingual: mean = 39.5, SD = 16.9, p < 0.001).

Stimuli

Two, 250-msec resynthesized vowels [ɛ] as in “bet” and [I] as in “bit” that differed in F1 and F2 formant frequencies (F1: 650 Hz and 500 Hz; F2: 1980 Hz and 2160 Hz, respectively) were used (See Figure 1). These stimuli were from a continuum of nine stimuli (deviant [I]: step 3, standard [ɛ]: step 9 of the continuum) used in previous studies with 4- to 10-year-old children with and without language impairment (Datta et al. 2010; Shafer et al. 2010; Rinker et al. 2017). Previous studies in our laboratory using all nine stimuli (step 1 to 9) showed that the categorical perception crossover point was at step 6 for both adults and children; thus, step 3 and step 9 were equal step distances from the crossover point. Moreover, step 3 was consistently identified as [I], and step 9 as [ɛ] by both adults and children as reported in our previous studies (Shafer et al. 2007; Datta et al. 2010). In addition, other studies indicated poor categorization and neural responses MMN to these vowels by late adult learners of English with Spanish as a L1 (Hisagi et al. 2015), and poor categorization of these vowels in children with specific language impairment (e.g., Datta et al. 2010). The stimuli were presented at 86.5 dB SPL in sound field over two speakers, located 1 m in front of and 1 m slightly above and behind the child’s head.
Electrophysiological design and procedures

The electroencephalogram (EEG) was recorded from a 63-site Geodesic net using Net Station amplifiers and Net Station 4.1 software at a 250 Hz sampling rate and 0.1 to 30 Hz bandpass filter. The EEG was timelocked to the onset of the vowel stimuli. The stimuli were delivered via Eprime software (Psychology Software Tools, Pittsburgh, PA, USA) in sequences of 10 stimuli with an interstimulus interval of 400 msec between stimuli in the sequence, and 1500 msec between each sequence of 10. Three train types were randomly presented (see Table 4). The deviant stimulus occurred in the fourth and eighth position for 100 trains, in the fifth and tenth position for 50 trains, and in the sixth and tenth position for 50 trains, for a total of 1600 standards [ε] (80%) and 400 deviants [I] (20%). This is the same paradigm that was used in Shafer et al. (2010, 2011, 2012). We did not reverse the standard and deviant because the study time would be too long and our goal was to compare group effects.

*Table 4*. Experimental paradigm: 10 stimulus per train, and three types of stimulus trains were randomly presented with a stimulus onset asynchrony of 650 msec, and intertrain-interval of 1500 msec.

<table>
<thead>
<tr>
<th>Train Type</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[ε]</td>
<td>[ε]</td>
<td>[ε]</td>
<td>[I]</td>
<td>[ε]</td>
<td>[ε]</td>
<td>[ε]</td>
<td>[ε]</td>
<td>[I]</td>
<td>[ε]</td>
</tr>
<tr>
<td>2</td>
<td>[ε]</td>
<td>[ε]</td>
<td>[ε]</td>
<td>[ε]</td>
<td>[I]</td>
<td>[ε]</td>
<td>[ε]</td>
<td>[ε]</td>
<td>[ε]</td>
<td>[I]</td>
</tr>
<tr>
<td>3</td>
<td>[ε]</td>
<td>[ε]</td>
<td>[ε]</td>
<td>[ε]</td>
<td>[ε]</td>
<td>[I]</td>
<td>[ε]</td>
<td>[ε]</td>
<td>[ε]</td>
<td>[I]</td>
</tr>
</tbody>
</table>

The procedures were explained to the caretakers, who then read and signed informed consent. The electrode net was first soaked in a saline solution and then the excess solution was removed. The nets were placed on the participant’s scalp while the participant was entertained by a laboratory assistant. The impedances of electrodes were maintained at or below 50 kΩ, which is sufficient for high input impedance amplifiers (200 MΩ). During the experiment, older
children sat in a chair, whereas infants sat on the caretaker’s lap while watching a video (e.g., Elmo, Baby Einstein) with the sound muted. Care-takers wore headphones playing music so that they would not orient to the vowel changes. Up to 200 sequences of 10 stimuli were delivered (1600 standards, 400 deviants) in an approximately 40-min session (28 min of stimulus delivery plus preparation time). Sessions were ended if a child became fussy and could not be calmed by changing the distracting task (e.g., by a researcher blowing bubbles or showing silent toys).

The continuous EEG was refiltered off-line using a low pass filter of 20 Hz, and was segmented with a prestimulus duration of 200 msec and 800 msec poststimulus onset and baseline corrected using the prestimulus 100 msec amplitude. Any epochs with electrical activity exceeding ±140 µV at any electrode site were rejected, and bad channels (on 20% of segmented trials) were replaced by spline interpolation algorithm from surrounding channels. Over 90% of the children had more than 100 trials for the deviant stimulus after artifact rejection. The mean number of trials for the standard and deviant ERPs after artifact reject did not differ significantly between the monolingual and bilingual groups, although there was a trend toward fewer trials obtained for the bilingual groups ($p > 0.07$, standards: monolingual mean = 601, SD = 217, bilingual mean = 540, SD = 224; deviants: monolingual mean = 241, SD = 88, bilingual mean = 216, SD = 91). Segments (−200 to 800 msec) were averaged for type (standard, deviant).

Analysis

ERP data reduction. The first step in the analysis was to reduce the ERP data using a combination of an objective metric and the existing literature. F3, Fz, and F4 are the most often reported sites to represent the left, central, and right frontal areas of MMN responses in the current literature. We selected two other neighboring sites that were highly correlated (Pearson’s $r > 0.6$, df = 248) with each of these three sites building the three models, respectively. That is, we used the average of responses from Geodesic sites 8, 9, and 13, which are located near F3 for the left frontal model, sites 5, 55, and 4, which are located between Cz and Fz for the frontocentral model, and sites 3, 58, and 62, which are located near F4 for the right frontal hemisphere model (Figure 2) (see Shafer et al. 2012 for a similar approach). This method of data reduction reduces independent noise measured at each electrode for participants, and thus increases signal/noise ratio. Five intervals of 40 msec from 160 to 360 msec were selected to reflect the time interval where MMN emerges. Five later time intervals of 40 msec between 400 msec and 600 msec were selected to examine the late response where late MMR was expected. The subtraction waveforms (waveforms of deviant minus standard) were used for statistical analysis.

$t$-test to determine the presence of early and late MMRs. In the first step, Welch’s Paired Sample (two-tailed) $t$-tests were calculated for monolingual and bilingual language groups within each time interval of 40 msec (between 160 msec and 360 msec for eMMR, and 400 and 600 msec for late MMR) to determine the presence of significant early and late MMRs in four age groups (3 to 5 months, 6 to 12 months, 13 to 24 months, and 25 to 47 months of age). For simplicity, we pooled the left, central, and right hemisphere responses. Note that our earlier study showed a significant positive MMR from approximately 200 to 360 msec in monolingual children (Shafer et al. 2011), and significant negativity in 6-month-old children from 560 to 600 msec (Shafer et al. 2012). Thus, these $t$-tests were expected to replicate the previous findings.

Linear mixed-effects modeling analyses in the earlier and later time window. We hypothesized that MMR amplitude was modulated by both intrinsic factors (such as age and sex)
and extrinsic factors (such as language exposure during early childhood). We also expected hemispheric differences based on our earlier findings (e.g., Shafer et al. 2010). To examine whether such factors were predictive of the early and late MMR amplitude measures, linear mixed-effects models were developed by using R (R Core Team 2014) and the lme4 package (version 1.1-15) (Bates et al., 2015). We chose a linear mixed-effects model approach because it allowed us to combine the longitudinal data and cross-sectional data together; linear mixed-effect modeling does not require the assumption of data-point independence. It also allowed us to use age as a continuous variable. The amplitude of the subtraction waveform (deviant minus standard) was modeled as a function of age (in months), language background (first: monolingual versus bilingual, then: monolingual, bilingual-more English and bilingual-more Spanish), hemisphere (left, central, right), and time bins (five intervals between 160 and 360 msec for the earlier time interval, and five intervals between 400 and 600 msec for the later time interval, respectively). To account for within-subject variability, a random intercept and a random slope of time and hemisphere for each participant were included in the model. Two sets of mixed-effects modeling were performed separately, one between 160 and 360 msec and the other between 400 and 600 msec. The effects of age (in months), hemisphere, sex, and language exposure were then tested following a bottom-up theory-guided approach, starting with level one predictors, then progressively adding subject level independent variables and interaction terms. The Bayesian information criterion was used to compare the fit of the competing models (Burnham & Anderson 2004). Partial effect sizes for specific model parameters were reported using semipartial R-squared ($R^2$) (Edwards et al. 2008). To facilitate the interpretation of the results, we reported Type III Analysis of Variance Table; p values were obtained using the Satterthwaite approximation for degrees of freedom (lmerTest version 2.0-36; Kuznetsova et al. 2017).

Figure 2. The sensor net and sites used for analysis. Left front model: the average waveforms from site 8, 9, and F3; the fronto-central model: site 5, 55, and Fz; the right front model: site 3, 58, and F4.
All models included maximal random effects structure (Barr et al. 2013), that is, random intercept for subject and random slopes for the within-subject independent variables were included. Random slopes were excluded only in those cases in which maximal models failed to converge. Within- and between-subject outliers were trimmed following a two-stage procedure: first, we excluded values more than 3 SD below or above the mean; second, level one and level two standardized residuals were examined and we refitted the models without observations with residual values more than 3 SD below or above the mean. Not more than 0.3% of the data were excluded for any of the models. Outliers were also excluded from group averages (Table 5).

Table 5. The average amplitudes (SD) of the subtraction waveforms (deviant minus standard) for monolinguals and bilinguals between 3 and 47 months of age for mismatch responses (160–360 msec) and late negativity (400–600 msec).

<table>
<thead>
<tr>
<th>Age (mo):</th>
<th>3–5</th>
<th>6–12</th>
<th>13–24</th>
<th>25–47</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language:</td>
<td>Monolingual</td>
<td>Bilingual</td>
<td>Monolingual</td>
<td>Bilingual</td>
</tr>
<tr>
<td>df</td>
<td>29</td>
<td>26</td>
<td>95</td>
<td>92</td>
</tr>
<tr>
<td>160–200</td>
<td>0.43(2.55)</td>
<td>0.17(2.03)</td>
<td>0.38(1.73)</td>
<td>−0.06(1.79)</td>
</tr>
<tr>
<td>200–240</td>
<td>0.95(2.33)</td>
<td>0.43(2.53)</td>
<td>0.65(2)</td>
<td>0.28(1.98)</td>
</tr>
<tr>
<td>240–280</td>
<td>0.65(2.81)</td>
<td>0.69(3.13)</td>
<td>1.01(2.15)</td>
<td>0.79(2.08)</td>
</tr>
<tr>
<td>280–320</td>
<td>0.93(2.43)</td>
<td>1(2.95)</td>
<td>1.05(2.07)</td>
<td>0.72(2.51)</td>
</tr>
<tr>
<td>320–360</td>
<td>0.82(2.39)</td>
<td>0.83(2.72)</td>
<td>0.87(2.1)</td>
<td>0.46(2.5)</td>
</tr>
<tr>
<td>400–440</td>
<td>0.91(2.45)</td>
<td>0.72(2.54)</td>
<td>0.21(2.3)</td>
<td>0.31(2.48)</td>
</tr>
<tr>
<td>440–480</td>
<td>0.85(2.48)</td>
<td>0.72(2.51)</td>
<td>0.34(2.46)</td>
<td>0.35(2.7)</td>
</tr>
<tr>
<td>480–520</td>
<td>0.87(2.48)</td>
<td>1.04(2.77)</td>
<td>0.54(2.62)</td>
<td>0.47(2.62)</td>
</tr>
<tr>
<td>520–560</td>
<td>0.43(2.12)</td>
<td>0.76(3.01)</td>
<td>0.69(2.65)</td>
<td>0.73(2.66)</td>
</tr>
<tr>
<td>560–600</td>
<td>0.34(1.78)</td>
<td>0.32(3.03)</td>
<td>0.61(2.67)</td>
<td>0.53(2.63)</td>
</tr>
</tbody>
</table>

Note. Responses from hemispheres (left, central and right) were pooled. Significant p-values are in bold. *p < 0.05; **p < 0.01; ***p < 0.001.

Results

Presence of MMRs in the early time windows and MMRs in the later time windows

Figure 3 displays the grand average responses at the superior frontocentral site (Fz) for the standard and deviant conditions for monolingual and bilingual children, respectively. Each language group was divided into four subgroups according to age. In general, in the early time window (160 to 360 msec), positive MMR appeared in all four age groups for monolinguals; in contrast, for bilingual children, a positive MMR was significant in all but those between 13 and 24 months of age. In the later time window (400 to 600 msec), the amplitudes of the difference waves were significantly positive for both monolinguals and bilinguals under 12 months of age, and significantly negative for all age/language groups older than 12 months of age (Table 5).

Figure 4 presents the eMMR and late MMR (deviant minus standard) for each of the language/age subgroups at the left (near F3), midline (near FCz), and right (near F4) regions. Table 5 reports the results from the t-tests regarding the presence/absence of significant MMRs in the early and later time windows. Table 6 reports the Proportions of children with negative eMMR (160 to 360 msec) and negative late MMR (400 to 600 msec).
Table 6. Proportions of children with negative eMMR (160 to 360 msec) and negative late MMR (400 to 600 msec).

<table>
<thead>
<tr>
<th></th>
<th>Monolingual</th>
<th></th>
<th>Bilingual</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>eMMR</td>
<td>late MMR</td>
<td>eMMR</td>
<td>late MMR</td>
</tr>
<tr>
<td>3–5 mos</td>
<td>0.23</td>
<td>0.34</td>
<td>0.29</td>
<td>0.25</td>
</tr>
<tr>
<td>6–12 mos</td>
<td>0.26</td>
<td>0.39</td>
<td>0.28</td>
<td>0.34</td>
</tr>
<tr>
<td>12–24 mos</td>
<td>0.30</td>
<td>0.49</td>
<td>0.49</td>
<td>0.51</td>
</tr>
<tr>
<td>25–47 mos</td>
<td>0.27</td>
<td>0.44</td>
<td>0.33</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Figure 3. The waveforms for the standard [ε] and deviant [I] at the fronto-central sites for monolinguals and bilinguals between 3 and 47 months of age.

Mixed-effects modeling in the early time window (160 to 360 msec)

In the early time window (160 to 360 msec), linear mixed-effects regression analysis revealed that sex (F(1,168) = 5, \( p = 0.03, R^2 = 0.03 \)), age (F(1,168) = 5.586, \( p = 0.02, R^2 = 0.03 \)), language (F(1,168) = 5.699, \( p = 0.02, R^2 = 0.03 \)), time (F(4,212) = 16.2, \( p < 0.001, R^2 = 0.23 \)), and sex-by-age interaction (F(1,168) = 5.01, \( p = 0.03, R^2 = 0.03 \)) were significant predictors of MMR amplitude, while hemisphere was not a significant predictor of MMR amplitude. It appears that more nMMRs were observed in female than in male infants, especially at a younger age (Figure 5).

We further divided bilingual children into more English versus less English experience, as quantified on the LBQ, and repeated the analysis procedure using language as a ternary variable (English monolingual, bilingual-more-English, and bilingual-less-English exposure). The same main effects and interaction were observed as in the model above in which language was used as a binary variable (language: F(2, 168.2) = 3.185, \( p = 0.04 \); sex: F(1, 168.2) = 4.904,
Posthoc tests revealed that the only difference between the two models was that when language was used as a ternary variable, there was a difference between monolingual children and bilingual children with less English exposure (BLE) \( p = 0.01 \), but the difference between monolingual and bilingual children with more English exposure was not significant \( p = 0.22 \).

**Figure 4.** The subtraction waveforms (deviant minus standard) for the left, central, and right frontal regions for monolinguals and bilinguals. The dashed boxes indicate time windows of analysis. Early MMR: 160 to 360 msec, late MMR: 400 to 600 msec.

**MMRs in the later time window (400 to 600 msec)**

The final mixed-effects model revealed that only age and time had significant predictive effects (age: \( F(1, 168) = 7.557, p = 0.007, R^2 = 0.04 \); time: \( F(4, 214) = 2.643, p = 0.03, R^2 = 0.05 \)). The effects of sex, hemisphere, and language exposure were not statistically significant, and more nMMR amplitude was associated with older age.

In summary, \( t \)-tests confirmed significant differences between the standard and deviant responses in the early and the late time windows, which indicated stimulus discrimination. The results from the mixed-effects modeling revealed language background, sex, age, and sex-by-age...
interaction are significant predictors of MMR amplitude between 160 and 360 msec. In the late time window of 400 to 600 msec, only age is a significant predictor of MMR amplitude.

Figure 5. Topomaps of the subtraction waveforms for early mismatch responses.

**Discussion**

This study examined the modulation of MMRs to English vowels in 3- to 47-months-old children by the factors of age, sex, hemisphere, and language experience. The current analysis showed that age, sex, and language experience influenced the amplitude of MMR in the early time window of 160 to 360 msec. In contrast, age was the only factor that influenced the amplitude of the late MMR in the time window of 400 to 600 msec. The new findings that we contribute to the current literature include the following: (1) Relative amount of bilingual exposure in English versus Spanish modulated the amplitude of the eMMR, but not the late MMR. Specifically, bilingual children with relatively balanced English and Spanish exposure or more English than
Spanish had similar MMR amplitude to their monolingual peers in the early time window (i.e., eMMR). However, bilingual children with Spanish-dominant exposure showed a difference in MMR amplitude from English monolinguals in the early interval. (2) Both the early and late MMRs to the vowel contrast shifted gradually from robust positivity to negativity from infancy as age increased. (3) No language exposure or sex effects were observed for the amplitude of the MMR in the late time window between 400 and 600 msec (i.e., late MMR). Later, we discuss our findings in greater detail in relation to the general literature.

**Maturational factors for the eMMR**

Age was a significant predictor of the amplitudes of the eMMR. The amplitudes of eMMR for this vowel contrast became less positive (or more negative) between 3 and 47 months of age. This pattern was independent of language exposure and sex. This result is consistent with the majority of previous developmental studies (e.g., Morr et al. 2002; He et al. 2009). He and colleagues (2009) found that in response to musical tone differences (specifically, a large fundamental frequency pattern contrast), 4-month-old infants showed a robust nMMR, but 2-month-old infants showed a positive MMR. However, the pattern of findings is quite complicated in young infants and appears to be dependent on stimulus factors, as well as age. For example, Leppänen and colleagues (2004) found that in neonates, a large positive MMR amplitude is associated with longer gestational age and more mature heart period at F3 and C3. The studies by Shafer and colleagues (Shafer et al. 2000; Morr et al. 2002) and Kushnerenko and colleagues (2002) showed a positive MMR to tone contrasts (20% and 50% frequency [F0] change) in infants that was largest between 3 and 8 months of age and decreased in positivity and was largely absent at 2 years of age, with emerging negativity between 3 and 4 years of age. The findings of the present study using these vowel contrasts are highly consistent with this pattern. Future studies will be needed to explore how degree of stimulus difference as well as stimulus quality (speech versus nonspeech) influence this pattern of development. The study by Kusherenko et al. (2007) has begun this task, revealing differences in neural discrimination of broadband white noise and narrower pitch differences.

Studies of brain maturation indicate that cortical auditory regions are highly immature in the newborn infant (Kostović & Jovanov-Milosević 2006; Moore & Linthicum 2007). Moore and Linthicum (2007) pointed out that intense sounds with large spectral differences are necessary to drive neural firing for encoding and discrimination at this young age. By 6 months of age, neural assemblies that receive auditory input have matured sufficiently to allow for encoding and discrimination of finer stimulus differences (Shafer et al. 2015). Kushnerenko et al. (2013) suggested that resolution of pitch information improves between 2 and 4 months of age. Fine-grained frequency resolution is essential for discriminating spectral differences between the vowels used in this study. The presence of a positive MMR indicates that the auditory cortex is sufficiently mature to resolve the spectral difference between vowels. However, we suggest that this pattern does not indicate the predictive detection that is signaled by MMN. Rather, the emerging nMMR reflects this process. Future studies will need to examine how behavioral correlates of perception relate to the positive versus negative eMMR in this early time interval.

As we predicted, we did observe a sex difference that appears to suggest more rapid maturation of female than male infants. Female children had significantly less positive eMMR (i.e., more negative eMMR responses). In particular, none of the male infants under approximately 4 months of age showed negative eMMRs to the vowel difference. In addition, 15 of the 41 female infants under 12 months of age did showed negative eMMRs. Evidence that
negative eMMR is an index of more mature processing comes from a study in which infants who demonstrated neural discrimination of complex higher order patterns (e.g., rapid learning of a linguistic rule) were the same as those who showed a negative eMMR to a pitch change (Mueller et al. 2012). Thus, we hypothesize that the more negative eMMR in female than male infants in the present study reflects faster maturation of the speech perception system, or more broadly, faster maturation of the auditory cortex, as reported in earlier studies (Shucard et al. 1981; Shucard & Shucard 1990; Shafer et al. 1999; Friederici et al. 2008).

In the present study, hemisphere was not a significant predictor of eMMR amplitude. We evaluated a larger age range (3- to 47-months) than some studies that have reported sex–hemisphere interaction in infancy (e.g., Shucard et al. 1977). For example, Friederici et al. (2008) reported that at 4 weeks of age among infants with low testosterone levels, female infants showed more bilateral discriminative responses while male infants showed more left lateralized responses. More rapid maturation of the left hemisphere has been observed in 3-months- old females and of the right hemisphere for 3-month-old males (Shucard et al. 1981; Shafer et al. 1999). Lack of a hemispheric effect or sex–hemisphere interaction in a larger age range suggests that the left and right hemispheric difference is less robust or disappears in males versus females, as infants grow older.

**Development of automaticity as indexed by MMR**

We had previously suggested that automaticity of speech perception is not present to the fine, phonetic contrast examined in the present study until around 4 years of age (Shafer et al. 2010; Shafer et al. 2011). Specifically, children are not initially automatic with discriminating this contrast because it takes time and experience to establish robust, selective perceptual routines for native-language phonological categories (Strange 2011). It is possible that larger acoustic-phonetic differences, for example between [i] and [ɑ], would allow for automaticity of discrimination because these are sufficiently salient. Crick and Koch (1990) argued that highly salient distinctions can be processed (discriminated or identified) with fewer attentional resources than less salient information. However, they also pointed out that less salient information can be made more salient through over-learning. Lack of robust nMMR (or the presence of positive MMR) in the early time window suggests that “over-learning” of these subtle differences is not yet achieved in young children. Some studies have reported the pMMR to nMMR polarity shift for auditory contrasts at younger ages (e.g., at 4 months for piano tone in He et al. 2007; at 6 months for lexical tone in Cheng et al. 2013). Cheng and colleagues (2015) found the pMMR in neonates shifted into nMMR for a large deviant /du/–/da/ contrast, but remained positive for smaller deviant /di/–/da/ contrast in 6-month-olds. Some of our 6-months-olds also showed nMMR rather than pMMR. However, at the group level, we did not observe nMMR in children under 12 months of age.

An alternative explanation for the absence of an nMMR is that filtering of the data at 0.3-Hz high-pass does not attenuate a large-slow positivity, and this positivity overlaps with the faster-rate nMMR. Some studies reporting group level MMN (i.e., negativity) used a high-pass filter of 1 Hz, which would attenuate slow activity with a time-constant > 1 sec (Cheng et al. 2013, Cheng et al. 2015, Garcia-Sierra et al. 2016). However, it is important to recognize that this filtering choice of 1-Hz high-pass would also attenuate a slow negativity, such as the late MMR/Nc. Note also that a faster eMMR (of 5 to 10 Hz) would still be discernable when riding on a slow positivity but this pattern is not apparent in Figure 4.
Monolingual versus bilingual experience on MMR

We observed a sex-by-language interaction in our earlier, cross-sectional study on the 6-month-old monolingual versus bilingual infants (e.g., Shafer et al. 2012). In the present study, we examined a wider age range and bilingual children, rather than female infants who showed a more negative eMMR. Part of the reason for the difference between studies may be related to the increased n, but also the wider age range. In the present study, a larger proportion of bilingual (41% in children > 12 months) than monolingual children (29% in children > 12 months) showed nMMR. Thus, this result indicated that bilingual experience was associated with more nMMR.

This result challenges the interpretation that nMMR is more adult-like, or is associated with more developed processing of speech contrasts, as reported in some studies (e.g., Cheng et al. 2015). Specifically, bilingual infants/children had less experience with this English vowel contrast than did monolingual infants/children. In Shafer et al. (2012), we reported evidence that regardless of language experience, infants showed more nMMR to a deviant that occurred in the final position compared with the mid position of sequences of 10 stimuli. A longer ISI of 1500 msec separated the final position from the onset of the next sequence of 10. We suggested that the final position drew more attention, and took as evidence an increase in positivity to a standard stimulus in this 10th position. Based on this finding, we speculate that more nMMR in bilinguals in the present study is due to more attention allocation by the bilingual children than the monolingual children. Behavioral studies have reported attention differences between monolingual and bilingual infants (e.g., Lewkowicz & Hansen-Tift 2012; Pons et al. 2015). Interestingly, in our recent study using short versions of these vowels comparing monolingual and bilingual adults and 8- to 11-year-old children, we found differences in attention-related neural measures, but not in the MMN (Datta et al. 2019). A study of infants is needed that directly manipulates attention to and away from the speech stimuli to test this hypothesis that the difference in MMR amplitude between the language groups is related to attention.

Amount of English versus Spanish input

We had predicted that the amount of language input would modulate development of neural discrimination of this English vowel contrast. Our LBQ allowed us to divide the bilingual children into two subgroups. The bilinguals with more English exposure (BME) had similar MMR amplitude to the monolingual group, whereas the BLE differed from BME and their monolingual peers, showing less positive MMR amplitudes in the early time range. Interestingly, Garcia-Sierra and colleagues (2016) reported that monolingual children with low language input had similar amplitude MMR to bilingual children, specifically showing more positive MMR than that observed for monolinguals with high language input. For bilingual children, they found that more language input was positively correlated with the amplitude of pMMR to a contrast from that language. These findings suggest a complex relationship between MMR amplitude and language input. Specifically, less positivity in infants can indicate poorer discrimination or less positivity can indicate the emergence of an overlapping nMMR (namely, the MMN). Specifically, with increasing age, the MMN moves earlier in latency as shown in our study with 4- to 7-year-old children using the same stimuli (Shafer et al. 2010). As a result, the MMN can overlap with and reduce the amplitude of the positive MMR (Morr et al. 2002; Shafer et al. 2010).

An important question is how much English language exposure is necessary for a bilingual toddler to achieve a similar language profile to his/her monolingual peers. Cattani and
colleagues (2014) reported that the proportion of English exposure is the main predictor of bilingual Spanish–English toddlers’ language performance, and bilingual toddlers who hear 60% or more English have matching language skills to their monolingual peers. Our study provided neural evidence that is consistent with the findings of Cattani and colleagues (2014). Our BME subgroup consisted of children who hear more than 50% of English in the ambient environment. These results have important clinical implications in that the amount of input can help decide whether an apparent language delay in a bilingual child is possibly due to insufficient input rather than to a true language disorder. However, it is important to recognize that even these higher levels of English input in children who are bilingual may still result in language differences from monolingual children, including lower-level English skills (Hoff 2013).

One caveat in the present study was that SES was not matched across groups. Monolingual children from different SES backgrounds can have a language input difference between low and high SES as large as 30 million words during their first 3 years of life (Hart & Risely 2003), and SES affects language production growth measures (Hoff 2013). In the present study, the BLE (less English) subgroup had similar SES scores (mean = 46.6, SD = 14.8) to the monolingual group (mean = 50.7, SD = 12.1), whereas the BME subgroup had lower SES scores (mean = 30.8, SD = 15.5) than the other two groups. The subgroup with BME, rather than the subgroup with similar SES (BLE), however, showed similar-amplitude MMR responses to the monolingual children. Thus, differences in SES appeared to be less important than amount of English exposure in the present study. It will be necessary to have a larger number of bilingual children, varying SES, proportion of English/Spanish input, and a more fine-grained measure of amount of input to allow a fuller exploration of how SES and amount of input interact to impact MMR amplitude.

The second limitation to generalizing our result broadly across all bilinguals is that Spanish and English are prosodically quite different. These prosodic differences may have aided the children in keeping the input from the two languages separate. In contrast, prosodically similar languages such as Spanish and Catalan might result in a significantly different pattern of neural responses, as suggested by the different behavioral trajectory for Catalan monolingual and Spanish–Catalan bilingual infants (Bosch & Sebastián-Gallés 2003). Thus, research examining other language pairs will be essential to fully understand the development of speech processing in bilingual children.

The late MMR

We compared the amplitude of the MMR in the late time window in monolingual and bilingual children, and found that age is the only factor that influences this response. The emerging negativity in older children may be equivalent to the LN found in grade school children or, alternatively, it may be the emerging MMN. Recall that researchers have used various names to refer to this late MMR, including LN (Kushnerenko et al. 2013). In infants, the presence of positivity in this window could be equivalent to the P3a. However, it is difficult to interpret these responses in isolation. The studies of Kushnerenko and colleagues (2007) examined MMRs to different stimulus types in different contexts and observed a LN to novelty sounds and not to an infrequent deviant (that was always the same stimulus). These novelty sounds, which changed on each trial, are more likely to draw attention. Our previous study also suggested that attention is necessary to elicit a late MMR, at least in 6-month-old infants (Shafer et al. 2012). In addition, this late MMR is consistent with the Nc observed to novel events in infants (e.g., Courchesne 1990). In the present study, after 12 months of age, both monolinguals and bilinguals showed
significant late nMMR. The lack of language exposure effects for the late MMR was similar to the findings of Garcia-Sierra and colleagues (2016). Specifically, they found increased negativity to the deviant between 350 and 550 msec, and that amount of language input did not affect MMR amplitude in the late time window in either monolingual or bilingual toddlers. These findings, however, contrasted with other studies that showed an effect of language on the late MMR, specifically in older children with language impairment (e.g., Bishop et al. 2010; Datta et al. 2010) or typically developing children who were learning a L2 (Shestakova et al. 2003).

Our results cannot fully address the mechanisms indexed by late MMR because we only used one contrast. Putkinen and colleagues (2012) examined multiple deviant types (e.g., duration, intensity, frequency, silent gap) with large, medium, or small magnitudes of acoustic differences. They obtained a robust LN for all deviant types in toddlers between 2 and 3 years of age. The authors argued that the fact that the smallest frequency contrast can elicit robust LN is inconsistent with the hypothesis that LN reflects reorienting of attention after a distracting stimulus (Shestakova et al. 2003; Horvath et al. 2009). If this was the case, then LN should be larger to greater stimulus difference. In our recent study with 8- to 11-year-olds and adult bilinguals, LN was larger for bilingual than monolingual adults but did not differ between monolingual and bilingual children (Datta et al. 2019). The stimuli in Datta et al.’s study were 50-msec versions of the [ε] versus [I] vowels used in the present study, and, thus, more difficult to discriminate (cf. Shafer et al. 2005; Datta et al. 2010). However, Datta et al. also included conditions focusing attention to the auditory modality. Thus, whether the mechanism indexed by the LN reflects reorienting has not been definitively decided. While the nature and functional role of the late MMR and LN need to be further examined, the developmental changes of the late MMR found in the present study together with other studies suggest that this late response has the potential of being used to evaluate the development of speech processing in toddlers. This is important because the MMR in the earlier time frame (150 to 400 msec) is often not significant to the subtle speech contrasts that are of particular interest in studies of language development.

Conclusion

Our findings demonstrate that neural indices of speech processing during the first few years of life undergo considerable development. Both maturational factors (age and sex) and experiential factors (language experience) modulate the infant/child MMRs. These developmental changes cannot be inferred from behavior. The considerable changes in neural discrimination (e.g., from a pMMR to a nMMR dominating the response) provide clues indicating that the mechanisms for encoding and discriminating auditory information differ greatly from those used by older children and adults. Specifically, we found that immaturity in neural discrimination was indexed by a positive MMR prior to 350 msec to the English [ε] versus [I] speech contrast. Increased maturity, in terms of age, was indexed by increased negativity of the MMR in this early time frame. In addition, female infants showed increased negativity of the MMR, consistent with more rapid maturation than male infants. The finding that language experience modulated the MMR in the early time frame provides further evidence for the importance of early speech experience in auditory-speech development. We hypothesize that this increasing negativity reflects increasing automaticity of discriminating the target contrast. We also suggested that the different pattern found for bilinguals could be the result of increased attention to the speech signal. Future studies need to further explore language-specific experience using other language pairs to fully understand the development of neural measures of speech processing.
Acknowledgments

The authors thank A. Barias and M. Wroblewski for helping with data collection, B. Tagliaferri for technical support, and W. Strange and R. G. Schwartz for advice on the design. This research was supported by NIH HD46193 to V. L. Shafer. V. L. S. oversaw the project, designed the experiments, and was involved in writing the article; Y. H. Y. helped with data collection, performed data analyses, and wrote the initial draft in conjunction with V. L. S., and led the manuscript revision process; C. T. helped with data collection and interpreting the language measures; H. H. and L. C. performed the early stages of the Mixed-Effect Modeling analysis in conjunction with Y. H. Y.; N. V. helped design the language background questionnaire and collect the data; J. G. helped collect the data; K. G. and H. D. helped design and pilot the electrophysiological paradigm and helped collect the data. All authors were involved in revising the article. The authors have no conflicts of interest to disclose.

Correspondence

Address for correspondence: Yan H. Yu, Department of Communication Sciences and Disorders, St. John’s University, 8000 Utopia Parkway, Queens, NY 11437, USA. E-mail: yuy1@stjohns.edu and Valerie L. Shafer, Ph.D. Program in Speech-Language-Hearing Sciences, The Graduate Center, City University of New York, 365 Fifth Avenue, New York, NY 10016, USA. E-mail: vshafer@gc.cuny.edu

References


Rinker, T., Alku, P., Brosch, S., et al. (2010). Discrimination of native and non-native vowel contrasts in bilingual Turkish-German and monolingual German children: Insight from the mismatch negativity ERP component. *Brain Lang, 113*, 90–95.


