

Research Article

Cite this article: Datta H, Hestvik A, Vidal N, Tessel C, Hisagi M, Wróbleski M, Shafer VL (2019). Automaticity of speech processing in early bilingual adults and children. *Bilingualism: Language and Cognition* 1–17. <https://doi.org/10.1017/S1366728919000099>

Received: 30 March 2018
Revised: 11 December 2018
Accepted: 22 January 2019

Address for correspondence: Hia Datta,
E-mail: hdatta@molloy.edu

Automaticity of speech processing in early bilingual adults and children

Hia Datta, Arild Hestvik, Nancy Vidal, Carol Tessel, Miwako Hisagi,
Marcin Wróbleski and Valerie L. Shafer

Molloy College; University of Delaware; Iona College; Florida Atlantic University; University of Connecticut; Pacific University Oregon and CUNY Graduate Center

We examine whether early acquisition of a second language (L2) leads to native-like neural processing of phonemic contrasts that are absent in the L1. Four groups (adult and child monolingual speakers of English; adult and child early bilingual speakers of English and Spanish, exposed to both languages before 5 years of age) participated in a study comparing the English /ɪ/ - /ɛ/ contrast. Neural measures of automatic change detection (Mismatch Negativity, MMN) and attention (Processing Negativity, PN and Late Negativity, LN) were measured by varying whether participants tracked the stimulus stream or not. We observed no effect of bilingualism on the MMN, but adult bilinguals differed significantly from adult monolinguals on neural indices of attention. The child bilinguals were indistinguishable from their monolingual peers. This suggests that learning a L2 before five years of age leads to native-like phoneme discrimination, but bilinguals develop increased attentional sensitivity to speech sounds.

1. Introduction

Learning a second language (L2) during childhood, compared to learning the L2 in adulthood, typically leads to superior L2 speech perception (Hisagi, Garrido-Nag, Datta & Shafer, 2015) and production skills (Baker, Trofimovich, Flege, Mack & Halter, 2008; Piske, Flege, MacKay & Meador, 2002; Yeni-Komshian, Flege & Liu, 2000). Perception is the listener's experience of the stimulus and is measured using behavior, such as phoneme-category identification or discrimination tasks. It is uncertain whether the neural processes that support speech perception differ between early bilinguals and monolingual listeners. The current study addresses this question.

Three alternative models have been proposed for how bilinguals perceive speech in their two languages (Hisagi et al., 2015). The first model suggests that bilinguals favor one phonology over the other (Cutler, Norris & Williams, 1987; Snijders, Kooijman, Cutler & Hagoort, 2007). The second model proposes that bilinguals compromise between the phonological systems of the two languages (Williams, 1977). The third model is that bilingual listeners adjust their phonological categories based on linguistic context (Elman, Diehl & Buchwald, 1977; Gonzales & Lotto, 2013).

Studies of adult bilinguals generally rely on self-reports of early language experience, which is likely to be imperfect, and, thus, could account for disparate findings in L2 speech perception across studies. However, language history of children reported by parents/guardians is more immediate and likely to be more accurate. Previous studies indicate that adult and child bilinguals do not necessarily show the same pattern of processing compared to monolingual age-matched participants (Baker et al., 2008; Brice, Gorman & Leung, 2013; Rinker, Shafer, Kiefer, Vidal & Yu, 2017; Tong, Lee, Lee & Burnham, 2015). For example, children who begin learning the L2 before five years of age may still demonstrate differences from monolinguals and these differences may be related to insufficient input. In addition, few studies have examined neural measures of speech processing in typically-developing children and only a few have focused on bilingual children (Kuipers & Thierry, 2015; Rinker, Alku, Brosch & Kiefer, 2010; Rinker et al., 2017). Thus, there is a clear need for investigations of speech processing in both monolingual and bilingual children.

1.1 Development of speech perception in monolingual and bilingual children

Monolingual children in the grade-school years generally show good phonological skills, but these skills are not yet fully developed (Nittrouer, 2006). Specifically, in speech perception, grade-school children rely more heavily on global cues (e.g., spectral formant transitions) than more fine-grained spectral cues (Nittrouer, 2002). Child learners who begin learning English as an L2 after three years of age generally show good English-language skills within four and half to six and half years of exposure in school (Paradis & Jia, 2017). However, a

language gap may persist into middle school, even for children learning English before six years of age (Farnia & Geva, 2011), if they speak a different language at home consistently. Socio-economic status (SES) and language background factors (e.g., language use in the home) may account for some of the differences in English language performance (Jia & Fuse, 2007). L2 performance can also vary across different aspects of the L2 (Paradis & Jia, 2017). For example, lexical knowledge may fall within the typical range, whereas phonology will continue to lag behind.

To date, only a few studies have closely examined L2 speech processing in bilingual, grade-school children. Perception studies indicate that there is some influence of the L1 on the L2, at least at younger ages. For example, a study of Korean–English children showed poorer perception of English vowels than monolingual English-speaking 2- to 5-year-old children, but better perception than adult late-learners of English (Tsukada, Birdsong, Bialystok, Mack, Sung & Flege, 2005). Immersion in the L2 at school leads to native-like perception of L2 vowels (McCarthy, Mahon, Rosen & Evans, 2014), but there can still be lingering differences at older ages (Darcy & Krüger, 2012). Speech and language skills in the L2 of early bilinguals, however, can be comparable to those of monolinguals by about 10 years of age, for children who attend schools in which the L2 is the dominant language (Paradis & Jia, 2017).

1.2 Neurophysiological measures of speech discrimination

Several neurophysiological studies of L2 speech processing have found differences between monolinguals and early bilinguals that are not apparent at the behavioral level (Hisagi et al., 2015; Sebastian-Gallés, Rodríguez-Fornells, de Diego-Balaguer & Díaz, 2006). Event Related Potentials (ERPs) reflect information processing that precedes the behavioral response. Specifically, the Mismatch Negativity (MMN) component indexes speech sound discrimination under conditions where attention is directed away from the stimulus of interest, thereby revealing more automatic processes. MMN is elicited in an oddball paradigm where one stimulus is repeated frequently (the standard) and a second stimulus is presented infrequently (the deviant) and computed as the difference between the response to these two conditions. The MMN is seen as increased negativity at fronto-central sites to the deviant compared to the standard, generally peaking between 100 and 300 ms following onset of the stimulus (Näätänen, Paavilainen, Rinne & Alho, 2007; Näätänen, Sussman, Salisbury & Shafer, 2014).

Many MMN studies have found that early bilingual experience results in differences from monolinguals (Molnar, Polka, Baum & Steinhauer, 2014; Peltola, Tuomainen, Koskinen & Aaltonen, 2007; Sebastian-Gallés et al., 2006; Shafer, Yu & Datta, 2011; Tamminen, Peltola, Toivonen, Kujala & Näätänen, 2013). It is unclear which of the three models presented above better fits the results from various studies. Some studies of bilingual adults suggest that their sensitivity to L2 phonological contrasts is influenced by linguistic context (Garcia-Sierra, Ramirez-Esparza, Silva-Pereyra, Siard & Champlin, 2012; Masapollo & Polka, 2014). Other studies support the claim that bilinguals have compromised between two phonological systems (Peltola et al., 2005, 2007; Tamminen et al., 2013), or support the claim that bilinguals favor one system over another (Hisagi et al., 2015).

Hisagi et al. (2015) found that adults with early bilingual experience showed accurate behavioral categorization and

discrimination of an L2 vowel contrast, similar to the monolingual listeners; this finding is inconsistent with the model suggesting that bilinguals ‘compromise’. In contrast, for the neural measure, MMN was smaller in early bilinguals (as well as in late bilinguals) than monolinguals; this pattern initially appears to be inconsistent with the model suggesting that bilinguals have flexibility to adjust to linguistic context. Taken together, these findings may indicate that at an attention-independent level, early bilinguals favor the L1, but that, with attention, they show flexibility, which allowed native-like behavioral judgments. Thus, it is of particular interest to understand how attention influences the MMN index of speech discrimination.

Speech perception studies using an oddball design may also elicit a late negativity (LN) response (Cheour, Korpilahti, Martynova & Lang, 2001; Wetzel & Schröger, 2014). Several studies with children have observed LN following the MMN (Datta, Shafer, Morr, Kurtzberg & Schwartz, 2010; Hestvik & Durvasula, 2016; Moreno & Lee, 2015; Putkinen, Niinikuru, Lipsanen, Tervaniemi & Huottilainen, 2012; Shafer, Morr, Datta, Kurtzberg & Schwartz, 2005; Shestakova, Huottilainen, Ceponiene & Cheour, 2003). The LN may be an index of re-orienting attention (Ceponiene, Lepistö, Soininen, Aronen, Alku & Näätänen, 2004). This response is of interest because it might serve as an index of attentional resource allocation during speech discrimination and allow testing of differences between monolinguals and bilinguals.

1.3 ERP measures of speech processing in children

ERP studies suggest that speech processing may not be adult-like until well past puberty (Shafer, Yu & Datta, 2010; Shafer, Yu & Wagner, 2015). This immature system may allow increased plasticity in learning an L2 (Kuhl, Conboy, Coffey-Corina, Padden, Rivera-Gaxiola & Nelson, 2008). In addition, some studies suggest that early learning of speech depends on implicit processes, which allows for native-like speech perception (Archila-Suerte, Zevin, Bunta & Hernandez, 2012). Most ERP studies of child speech processing focus on disorders (Kujala & Leminen, 2017). The few that focus on speech processing in child L2 acquisition have largely examined children between three and seven years of age.

The first ERP studies of child L2 learners suggested differences from age-matched monolingual controls. In several studies, experience with an L2 in a daycare or school setting led to a larger MMN to an L2 speech contrast compared to non-native child listeners (Cheour, Shestakova, Alku, Ceponiene & Näätänen, 2002; Peltola et al., 2005; Shestakova et al., 2003). For example, three- to six-year-old Finnish children exposed to French in pre-school showed an increased MMN to the French vowel contrast /e/ and /ɛ/ versus standard /i/ after six months of experience (Cheour et al., 2002). The P3a and LN also increased from pre- to post-exposure. Cheour et al. (2002) suggested that the LN indicated involuntary attention shifts to the deviant sounds and that this was more apparent as time of L2 exposure increased. The P3a is an index of involuntary attention orienting to non-target deviant stimuli (Polich, 2012).

Other studies of child L2 learning have failed to observe an increase in MMN amplitude to an L2 speech contrast as L2 experience increases. For example, Peltola et al. (2007) did not observe increased MMN for eight-year-old Finnish children immersed in learning English compared to Finnish controls. Interestingly, the Finnish children learning English did not show robust MMNs to native Finnish contrasts either. The

authors suggested that neural circuitry was not committed to native sounds at this age. In another study, five- to six-year old children from Turkish–German-speaking homes with two to three years of German exposure exhibited smaller MMN to German vowel contrasts compared to German monolingual children (Rinker et al., 2010). The authors suggest that inadequate L2 input beginning after three years of age may account for the small MMN to the L2 contrast.

ERP studies of L2 speech in children have not reported whether an LN or P3a is modulated by language experience. In our previous studies of children, the LN was not modulated by attention (Shafer et al., 2005; Datta et al., 2010), but it was present in both children with specific language impairment (SLI) and their typically developing peers. In addition, for a more salient vowel contrast (longer 250-ms /ε/ versus /ɪ/), children with typical development showed a P3a, indicating that the stimulus difference was sufficiently great to lead to an orienting response (Datta et al., 2010). Thus, it is of interest to examine whether bilingual experience modulates neural indices of attention orienting in children.

1.4 Effects of attention on speech perception in adults and children

Attention plays a role in the development of speech perception. Differences in maturation of attentional skills and in how attention is employed during speech processing tasks could influence the pattern of results observed in studies of bilingual children and adults. Several developmental models suggest that infants initially direct attention to relevant cues in the ambient language to acquire a weighting scheme, or selective perception routines (SPRs) for the native language phonology (Jusczyk, Cutler & Redanz, 1993; Kuhl et al., 2008; Strange, 2011; Werker & Curtin, 2005). Over the first four years of life, SPRs are hypothesized to become automatized to allow for efficient recovery of the phonological form from the acoustic-phonetic information (Shafer et al., 2010, 2011). Late L2 learners often do not exhibit automaticity of L2 speech perception but use their L1 SPRs instead. This leads to poorer perception under conditions of high cognitive load, such as perception in background noise (Strange, 2011).

The time course for developing automaticity of speech perception in an L1 or L2 is unknown. We have suggested elsewhere that monolingual American-English children do not show automaticity in processing a contrast between the vowel /ɪ/ in “bid” and /ε/ in “bed” until after four years of age (Shafer et al., 2010, 2011), based on the finding that MMN was not observed in the majority of children until after four years of age (also see Lee et al., 2012). The latency of the MMN was also later for children compared to adults, suggesting that children’s processing at this stage of development is less automatic.

In two previous studies which examined speech perception in 8–10 year-old children (Shafer et al., 2005; Datta et al., 2010) and one of which used the same stimuli as in the current study (Shafer et al., 2005), little effect of attention was found on the MMN to the /ɪ/ vs. /ε/ contrast. In these studies, the children were asked to attend to a tone occurring infrequently among the vowel stimuli in one condition and ignore the auditory stimuli and watch a video in the other condition. The only difference in the MMN was that the response began earlier when attending, but only for the long-vowel contrast (Datta et al., 2010). These studies suggest that by 8–10 years of age, children are sufficiently automatic in

L1 speech perception and, like adults, attention has little effect on neural discrimination.

In the studies with the tone target (Shafer et al., 2005; Datta et al., 2010), there was evidence of greater attention to the auditory modality, seen as a P3a response following the MMN to the long vowels in a condition requiring attention to the stimuli (compared to ignoring the stimuli). In addition, the “processing negativity” (PN), which indexes attention orienting and is seen as a negative shift in the ERP (Näätänen, 1982), was observed in the attention task (Shafer et al., 2007). An interesting question is whether children and adults who are bilingual from an early age would show a similar pattern of neural responses in processing speech under different attention conditions as monolinguals. This is the central question that we address in the current study.

1.5 The present study

Our first aim was to investigate whether early Spanish–English bilingual speakers exhibit different neural processing of the contrast between /ɪ/ and /ε/ in American English – a contrast that is not phonemic in the Spanish L1 of these speakers – in comparison to monolingual native speakers. Our earlier study revealed a smaller MMN in a passive task to this contrast in bilinguals compared to monolinguals, even when the bilinguals learned English before five years of age (Hisagi et al., 2015). This smaller MMN was attributed to reliance on SPRs (Strange, 2011) tailored to Spanish vowels rather than to American English vowels. This previous study did not fully examine the amount of language input received in English vs. Spanish by the participants. Thus, the current study includes more language background information, and aims to replicate the MMN difference observed in our previous study, and at the same time establish the relationship between MMN and language use measures.

Our second aim focused on whether 8–10 year-old bilingual children, who began learning English no later than five years of age, would show maturational differences and/or show similar language group differences observed between adult early bilinguals and monolinguals. The L2 is becoming well-established between four and a half to six and a half years of experience in the school system (Paradis & Jia, 2017), but there may still be subtle differences in phonological processing that are not apparent in behavior. In addition, auditory maturation is incomplete (Shafer et al., 2000). We hypothesized that bilingual children would show smaller MMN than their monolingual peers in a passive task (watching a muted movie) than when paying attention to the auditory stimuli. Specifically, the bilinguals might rely on their L1 speech perception routines. In addition, we predicted that both groups would show a late negativity (LN) discriminative response, because this response is robustly present even in children with weak language skills (Shafer et al., 2005).

Our third aim was to examine whether early bilingual experience affected automaticity of processing of this vowel contrast, as indexed by the MMN and LN discriminative responses. We hypothesized that monolingual English-speaking adults would show no modulation of MMN to this contrast, as a function of attention directed to versus away from the stimulus stream; Spanish–English bilinguals, however, might show enhanced MMN amplitude when attending to the speech stream, because they are less automatic. This prediction follows from the claim that L1 speech perception is highly automatic, and thus, not influenced by attention, whereas L2 speech perception is more effortful (Hisagi et al., 2015). With regards to the LN, we predicted that

bilingual listeners would show a larger LN (Ortiz-Mantilla, Choudhury, Alvarez & Benasich, 2010). In addition, we hypothesized that both child groups would show increased MMN and LN when attending to the speech signal as compared to ignoring the speech sounds because their speech processing skills are still developing, and thus are less automatic.

To manipulate attention, we instructed participants to carry out two different attention-related tasks. One condition drew attention to the auditory stream via a speech target (participants had to identify an infrequent /ba/ target interspersed among the vowels and an infrequent /da/). A second condition drew attention to the auditory stream via a non-speech target (identify a high tone target among the vowels and an infrequent low tone). The third condition drew attention away from the auditory events via a muted video. We predicted that the PN to the vowels would be larger in the conditions where attention was drawn to the auditory modality for all participants (Hansen & Hillyard, 1980; Näätänen, 1990). Some studies suggest that bilingual experience enhances executive functions, including attentional control (Bialystok, Craik & Luk, 2012). In this case, bilinguals and monolinguals would differ in the PN effect; however, it is unclear whether better attentional control would result in a larger or a smaller effect (but see Astheimer, Berkes & Bialystok, 2016).

2. Experiment I: monolingual vs. bilingual adults

2.1 Methods

Participants

All participants were recruited from the New York metropolitan area through public postings on the internet or via letters sent to the homes (addresses obtained using Experian). After a telephone screening to determine eligibility based on language background, participants were scheduled to visit the lab, where they signed consent forms. Participants were screened for any history of speech-language, attention or neurological problems through interview and questionnaire. Hearing was screened at 25 dB hearing level from 500 to 4000 Hz; one monolingual adult was excluded because of a failed hearing screening.

Twenty-five adults were monolingual speakers of American English and 15 adults were bilingual Spanish-English speakers. The bilingual adults met the inclusion criteria of either being born in the US, or having arrived before five years of age, and having acquired both English and Spanish at this age or earlier.

The monolingual group consisted of 14 women and 11 men (mean age = 29.9, range = 19 to 40; SD = 7), and the bilingual group consisted of 11 women and 4 men (mean age = 28.6, range = 19 to 40; SD = 6.3). Participants completed a language background questionnaire (LBQ), the results of which are summarized in Table 1. Adult bilinguals' mean reported age of first words in English was 36.5 months (SD = 20.2, range = 12-60), and first words in Spanish was 25.1 months (SD = 16.5, range = 12-60). Most participants indicated that English acquisition began later than Spanish acquisition.

Participants rated amount of input in various contexts (e.g., home, community, school, media) on a seven-point scale (1 = all Spanish to 7 = all English, with 4 = balanced input) and proficiency on a five point scale. All scales were rescaled to a 7-point scale. The mean of the home input, school input and proficiency scores were used to compute a composite Spanish proficiency/use score, ranging from 1 to 7. The mean Spanish proficiency score for the adult bilinguals was 4.45 (SD = 1.5, range = 1.8-6.3).

Stimuli

The stimuli were two vowels; /i/ as in American English 'bit', and /e/ as in American English 'bet.' /i/ and /e/ constitute different phonemes in English, but not in Spanish: in Spanish [e] is an allophone of Spanish /e/. In contrast, American English /i/ is not perceived as a good exemplar of Spanish /i/ or Spanish /e/ and Spanish adult late learners of English perform at chance levels in a forced choice categorization task of American English /i/ when the alternative is /e/ (Hisagi et al., 2015). The two tokens used in the experiment were taken from a continuum of nine vowels that were created by editing the first and second formant values of a re-synthesized token produced by an American-English female (Hisagi et al., 2015). The bandwidth for each formant was maintained from the original recordings, which gave the stimuli a natural quality (timbre). The final speech stimuli were 50 ms in duration with a rise and fall time of 5 ms. F₀ was maintained at 190 Hz. The third (F3) and the fourth (F4) formants were constant at 2174 Hz and 3175 Hz, respectively. The nine exemplars were made by increasing F1 and decreasing F2 in equal steps from /i/ to /e/. The two tokens selected for the experiment had mean center frequencies of F1 at 500 and 650 Hz and F2 at 2160 and 1980 Hz respectively, and were the same as those used in previous studies (Hisagi et al., 2015; Shafer et al., 2005). In addition to the vowels, the experiment also included auditory stimuli that served as targets in an "attend-to-auditory-stream" condition (and were included but ignored along with all stimuli in an "ignore-auditory-stream" condition, see below). The attention task target stimuli were two 100-ms pure tone stimuli of 500 Hz and 2000 Hz, and two naturally recorded syllables /ba/ and /da/ that were 250 ms in duration. All stimuli were presented at 72 dB SPL sound field over two speakers.

Experimental design

The within-subject design of the experiment consisted of the factor CONDITION (standard vs. deviant) crossed with the ATTENTION conditions: *Attend* to the stimulus stream vs. *Ignore* the stimulus stream. In addition, the factor TARGET (speech vs. tone) was fully crossed even though there was no task associated with targets in the Ignore condition, resulting in a CONDITION (Standard vs. Deviant) x ATTENTION (Attend vs. Ignore) x TARGET (speech vs. tone) x LANGUAGE (monolingual vs. bilingual) design. The vowel and target stimuli were identical across the three tasks, but the task instructions differed. The vowel /e/ (standard) was delivered for 79% of the trials, and /i/ (deviant) was presented for 17% of the trials. The interspersed targets (speech and tones) for the Attend level comprised 4% of the total trials. The Attend-speech condition was designed to focus attention on spectral information in speech (higher resonances of the first, second and third formants); the participants were asked to respond to the /ba/ stimulus. To do this, they needed to reject [da], as well as the vowel stimuli. In the Attend-tone condition, the target was the 2000 Hz pure tone. The 500 Hz pure tone was included to give participants a choice between two tones. The stimuli were followed by a 600 ms inter-stimulus interval (ISI). The Attend-speech vs. Attend-tone was introduced to determine whether focus on speech versus non-speech auditory targets would modulate the automatic vowel discrimination.

Procedure

Each participant was asked to fill out a case history form designed to screen for any prior speech, language, hearing, psychological or

Table 1. Language background for adults calculated from 7-point scale. The number of questions used to calculate the response is shown in the column labeled “No. Questions”. The number of participants who responded in the range is provided (Mostly Spanish = 1 or 2; Equal = 3, 4, or 5; Mostly English = 6 or 7). Participants’ origin country: Puerto Rico (N = 4); Dominican Republic (N = 2), Colombia, Peru, Venezuela, Cuba, Spain, Guatemala (N = 1 each). Three did not report country of origin. Only one participant reported hearing more English than Spanish in the home. The mean self-rating for which language was spoken at home was 3.4, and the mean self-rating for which language was spoken in the community was 4.7. Literacy skills, however, favored English (5.8), which is consistent with participants growing up in New York City and attending public schools.

Category	No. Questions	Mostly Spanish	Equal	Mostly English	Not Reported/NA
Home	11	3	10	1	1
Community	6	0	14	1	0
Media	8	0	7	5	3
Literacy	7	0	1	11	3
Parent	2	7	6	0	2
Grandparents	2	9	1	0	4

neurological issues. All participants passed a hearing screening to ensure hearing was within normal limits. The participant was then seated in a comfortable chair in a sound-shielded audiometric booth in front of a PC monitor (17” screen) placed approximately 1 meter distant in the center. The stimuli were presented at a comfortable hearing level via two loudspeakers speakers that were suspended above the participants. One speaker was placed at a distance of 1.5 meters from the participant at a vertical angle of 45 degrees in the front while the other was 0.5 meters at 25 degrees behind the participant. Participants were monitored from outside the sound booth via a video camera. E-Prime software version 1.1 controlled stimulus delivery (Schneider et al., 2002).

The stimuli were delivered in 12 blocks of trials with a break after each block. A total of 200 deviants vs. 952 standards were presented in each of the three conditions. Forty-six target sounds (tones or syllables) were included in each Attend condition. Participants were tested on two different days. They received the Attend-tone and Ignore-tone condition (with non-attended tone targets) during one of the visits and the Attend-speech and Ignore-speech condition (with non-attended speech targets) during the other visit. The order of the tone and speech target conditions were counter-balanced across participants. All participants received the Ignore conditions of the experiment first on each day to prevent participants from having heightened awareness for the stimuli in this condition.

EEG acquisition

The electroencephalogram (EEG) was recorded with a 65-channel Geodesic Sensor Net with silver/silver-chloride (Ag/AgCL) plated electrodes using Net Station Software version 4.1. The electrodes were sheathed in sponge-encasings, which were dampened using a potassium chloride solution. The impedances of the electrodes were maintained below 40K Ω . Vertical and horizontal electrode montages near the eyes were used to monitor eye movement and eye blinks. The vertex served as the reference during data acquisition. The EEG was sampled at 250 Hz, and then amplified with a band pass filter of 0.1-30Hz using a 64-channel Net Amps 200.

EEG post-processing

The continuous EEG was segmented into single trial epochs of 850 ms duration, including a 200 ms pre-stimulus onset baseline period. After baseline subtraction, the segments were submitted to Netstation artifact detection procedures, for detecting eye blinks/movements (using a 70 μ V threshold) and bad channels. Trials

with eye blinks or movements were marked for exclusion, and channels marked as bad were then replaced with the spherical interpolation. An average was computed for each stimulus type and condition (standard and deviant in Attend-speech, Attend-tone, Ignore-speech, Ignore-tone). The data were then re-referenced to the average of all sites.

2.2 Results

2.2.1 PCA pre-processing.

We first identified temporal and spatial “regions-of-interest” by conducting sequential temporo-spatial Principal Component Analysis (Dien, Spencer & Donchin, 2003; Spencer, Dien & Donchin, 1999, 2001) using the PCA ERP toolbox (Dien, 2010). The decomposition was conducted on the difference waves (deviant minus standard), so that the PCA focused on the temporal and spatial distribution of the experimental effects, rather than the obligatory components of the auditory evoked potential (Luck, 2014). We then based selection of time windows and electrode regions in the voltage data on the latency of temporal and spatial factors uncovered by the PCA. This strategy avoids researcher bias in selecting electrodes and time samples for analysis, and mitigates against increased Type I error rate (Luck & Gaspelin, 2017).

The scree plot test and the parallel test was used to determine the number of components to retain (Horn, 1965). This retained 11 initial temporal factors, which accounted for 91% of the variance; the PCA was then rerun limited to 11 factors, which were rotated using the covariance matrix (without Kaiser normalization) to simple structure, using PROMAX ($k = 3$) (Hendrickson & White, 1964; Richman, 1986; Tataryn, Wood & Gorsuch, 1999). We then limited analysis to temporal factors that accounted for at least 5% of the total variance. Only the first four temporal factors met these criteria. Visual inspection of the topographical distribution of these temporal components revealed that only two of the four factors corresponded to the typical time course and spatial distribution of the MMN (208 ms, TF2) and the LN (644 ms, TF1).

Each temporal factor was decomposed into its spatially independent components using ICA (Bell & Sejnowski, 1995), following the same procedure for factor reduction as for temporal PCA. After inspecting the scree plot, the ICA was limited to 6 spatial factors. The combined or sequential temporo-spatial PCA results in one factor score per subject and condition, which represents

the weighted average of all electrodes and time samples for each underlying temporo-spatial factors (Dien, 2012; Dien & Frishkoff, 2005) and constrains selection of time windows and electrode regions for the observed ERPs in the data.

2.2.2 Mismatch Negativity

The time course of the MMN factor was identified as the second temporal PCA factor TF2, peaking at 208 ms. The electrode region for MMN was identified as the third subfactor of the spatial decomposition of this temporal factor (TF2SF3), which had a central distribution typical of the MMN. A subset of electrodes with factor loadings exceeding 0.6 was selected as the most highly weighted electrodes and used to represent the MMN (a “virtual channel”) in the voltage data (the blue region of the topoplot in Figure 1). Figure 1, top left panel shows the main difference wave of TF2SF3 expressed as microvolt-scaled factor scores; the top right panel shows the topographical distribution of the factor score difference wave; and the bottom panel shows the mean voltage for the MMN electrode region with deviants, standards and the difference waveforms for monolingual and bilingual adults.

We analyzed the raw voltage data constrained by the temporal and spatial properties of TF2SF3 as follows: first, the time window 152-244 ms was constructed by selecting the time samples in the TF2 factor with loadings greater than 0.6. The electrode region of EGI sites E4, E5, E17, E18, E22, E30, E43, E47, E54, E55, E58, E65 was then selected using electrodes with TF2SF3 factor loadings greater than 0.6 (see above); corresponding to the dark blue region in Figure 2. A mean voltage difference score (deviant minus standard) for this time/space region was computed for each subject and cell, and submitted to the same mixed factorial ANOVA as for the factor score analysis. The statistical results matched the factor score analysis, with a significant intercept ($F(1,38) = 44.9$, $p < .00001$) (i.e., a significant mismatch effect, as the dependent measure was a difference wave), and a significant ATTENTION x TARGET TYPE interaction ($F(1,38) = 4.21$, $p < .05$). The interaction was that speech targets resulted in greater MMN in the Attend condition than in the Ignore condition (cf. Figure 2).

There was no other main effect or interactions; specifically, no interactions involving group.

The only effect of attention on the MMN was that when subjects tracked the stimulus stream for target stimuli, the MMN was enhanced by tracking speech-like targets, but not by non-linguistic targets, and this effect on the MMN was the same for monolinguals and bilinguals. Thus, attention to speech properties in the signal led to greater MMN for both groups. There was no difference between adult monolinguals and adult bilinguals in the MMN or in the speech-attention related enhancement of the MMN response. There was also no significant correlation between the mean MMN across conditions and the participants' Spanish proficiency/use scores (Spearman rank order correlation, $N = 14$, $r = 0.14$, $t(N-2) = 0.51$, $p = 0.6$).

2.2.3 Late Negativity

The Late Negativity (LN) was captured by the first temporal factor TF1, peaking at 644 ms, and which also accounted for the largest amount of variance in the data. Examining the 6 spatial subfactors within temporal factor 1, the first sub-factor TF1SF1 best matched the observed anterior negativity in the undecomposed grand average data. The remaining 5 spatial factors did not have topographical distributions indicative of cognitive ERPs and were discarded from further analysis. Figure 3, top panel, shows the temporal and spatial distribution of the main

effect difference wave for the LN temporo-spatial factor, and the lower panel shows the raw voltage data averaged for the electrode region defined by the spatial factor (as above by selecting electrodes with factor loadings greater than 0.6), by language group.

Bilinguals clearly show greater negativity than monolinguals in this ERP component.

We next used the temporal and spatial distribution of TF1SF1 to constrain the selection of a region of interest from the undecomposed voltage data, and constructed an electrode region defined by electrodes that had factor loadings greater than 0.6 (the dark blue area in Fig 3; specifically electrodes E6, E7, E10, E11, E12, E14), and the 376-648 ms time window, defined by these time samples in TF1 that exceeded 0.6. The average difference-wave voltage for this time window over this electrode region was used as the dependent measure in an ANOVA with LANGUAGE, ATTENTION and TARGET as factors. The statistical analysis mirrored the findings of the factor score analysis, resulting in a main effect of intercept, that is, the LN ($F(1,38) = 16.3$, $p < 0.001$); and a main effect of LANGUAGE ($F(1,38) = 5.6$, $p < 0.05$), in which the effect was significantly greater for the bilinguals. In the voltage analysis, the ATTENTION x TARGET interaction reached significance in the voltage data analysis ($F(1,38) = 4.17$, $p < 0.05$), driven by a greater LN in the condition where participants were tracking non-speech target tones; that is, in this condition, the difference in brain response to deviants and standards were enhanced, compared to when participants were tracking speech targets (i.e., the opposite of the ATTENTION x TARGET effect observed in the MMN).

2.2.4 Processing Negativity

Finally, we examined whether the two adult groups differed in the Processing Negativity (PN) component. In order to isolate the temporal and spatial region for statistical analysis of the N1, we first conducted a temporo-spatial PCA limited to the standards in the Attend conditions and the standards in the Ignore conditions. We retained 8 temporal factors, the three first of which accounted for at least 5% of the variance. The fourth temporal factor TF4 (accounting for 5% of the variance) matched the temporal and spatial distribution of the N1 peak (120 ms) observed in the grand average voltage data. In the next spatial step, we retained 5 spatial factors. The second spatial factor TF4SF2 had a topography that matched the N1 part of the Auditory Evoked Potential (AEP; see Figure 4, top panel).

We next analyzed the voltage data constrained by the factor analysis. The time window 104-148 ms was defined by the samples with TF4 and the electrode region (E22, E29, E30, E34, E42, E43, E47; a region slightly posterior from Cz) was selected where the factor loadings exceeded 0.6. The same ANOVA was run with these time/space voltage means as dependent measures. This resulted in a main effect of ATTENTION ($F(1,38) = 19.7$, $p < 0.0001$); and an interaction ATTENTION x LANGUAGE ($F(1,38) = 5.7$, $p < 0.05$), such that bilinguals had a greater difference between Attend and Ignore: that is, a significantly larger amplitude PN than the monolinguals. Finally, the voltage analysis also showed an interaction between Attention and Target type, ($F(1,38) = 9.7$, $p < 0.01$), such that PN to the standard vowels was greater when subjects attended to tones.

2.3 Discussion

The adult monolingual and bilingual groups did not differ in MMN amplitude, and the MMN was not modulated by attention

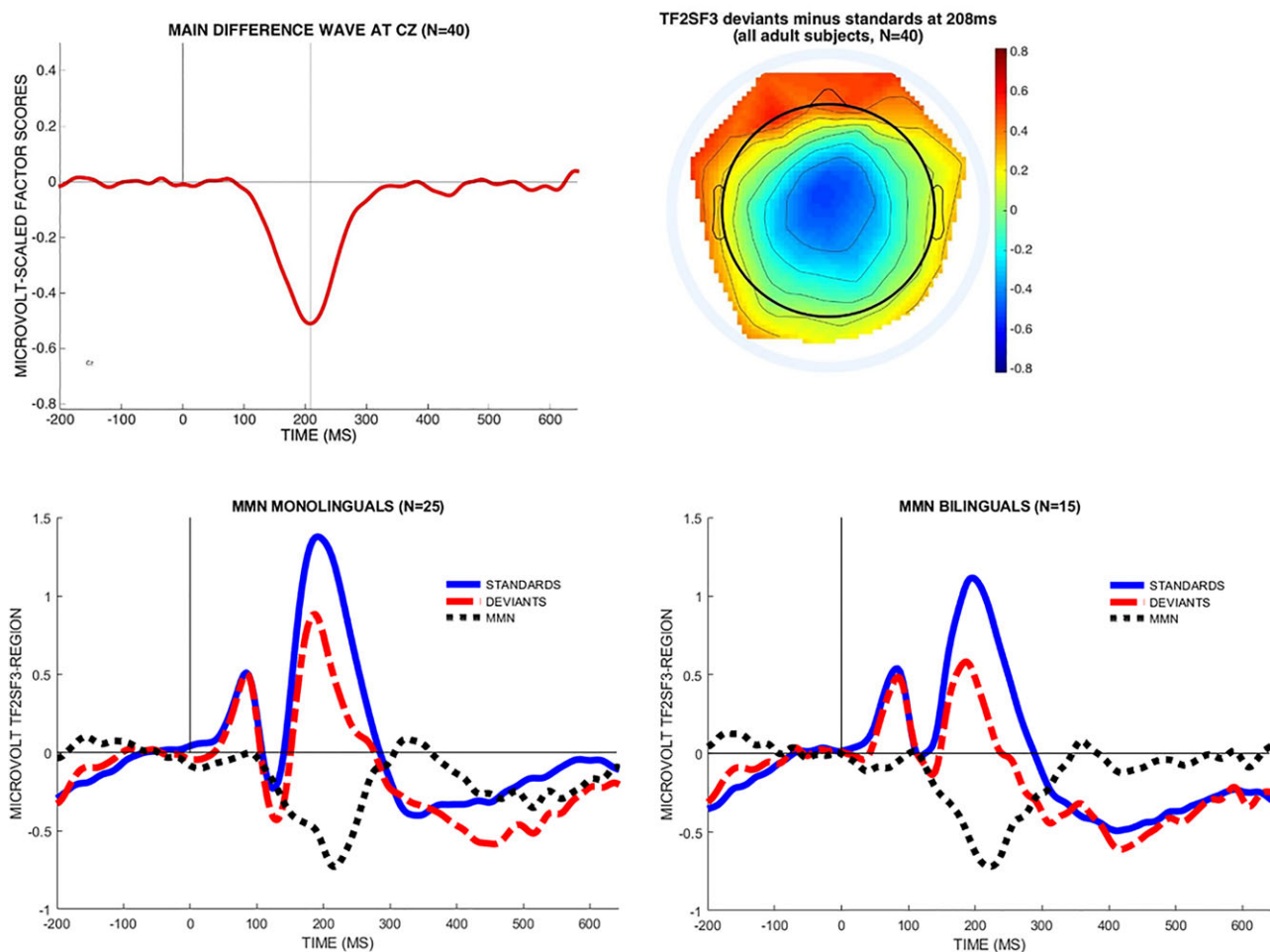


Fig. 1. Top row, left panel: Factor score main effect difference wave. Right panel: Spatial distribution of TF2SF3 factor score difference wave. Lower row, left panel: main effect MMN in monolinguals, average of TF2SF3 electrode region; right panel; MMN main effect for bilinguals.

to or away from the speech. Both groups showed a slightly larger MMN in the Attend condition when the target was a speech sound. This suggests that processing of the vowel contrast /i/ versus /ε/ is native-like for these early bilinguals. On the other hand, whereas both adult groups exhibited a LN to the deviants in both the Attend and Ignore conditions, the LN amplitude was significantly larger for the bilingual group. We also observed a bilingual effect on the PN. Both groups showed increased negativity to the (standard) stimuli in the Attend compared to the Ignore condition, but the PN amplitude was greater for the bilinguals. This suggests that bilinguals may be allocating more resources to processing the stimuli in the Attend condition than the monolingual group.

Another surprising finding was that the LN had greater amplitude in the Attend-tone condition compared to Attend-speech condition for both monolinguals and bilinguals. LN is an index of reorienting attention (Ceponiene et al., 2004). It may indicate greater effort in discriminating speech (vowel stimuli) when auditory attention is directed towards non-speech targets (tones).

3. Experiment II: monolingual vs. bilingual children

Experiment II was identical to Experiment I in all respects, except the participants were monolingual and bilingual children. For this

experiment, parents or guardians provided written consent for participants and completed the case histories. The goal of this experiment was to examine the maturation of neural responses to the English vowel contrast in relation to the attention manipulation.

3.1 Methods

3.1.1 Participants

Fifteen monolingual children (6 males and 9 females, mean age = 8.9, range 9-11; SD = 0.9) and 15 bilingual children (10 males and 5 females, mean age = 9.3, range 9-11; SD = 0.85) were tested. Except for one child born in the UK, the bilingual children were born in the US.

Parents completed a Language Background Questionnaire regarding early language exposure (simultaneous, sequential), when the child's first words in English and Spanish were observed, schooling (preschool, kindergarten, etc.) and amount of Spanish versus English use in the child's environment. (This questionnaire was incomplete for one child.) Table 2 summarizes the results of the Language Background Questionnaire.

English language abilities were tested for all children using *Clinical Evaluation of Language Fundamentals*, Fourth Edition (CELF-4) (Semel, Wiig & Secord, 2004) and *Peabody Picture*

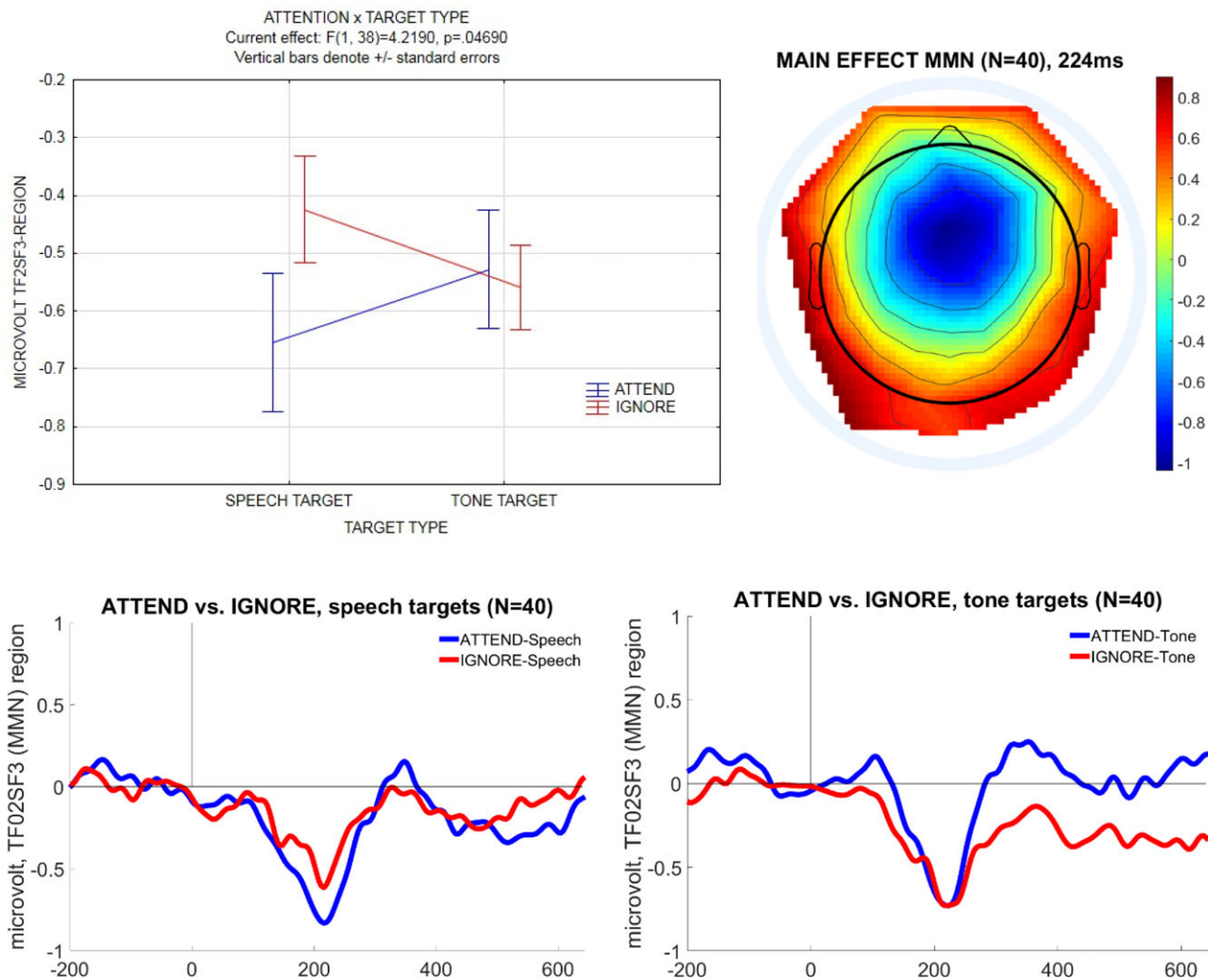


Fig. 2 - Colour online, B/W in print

Fig. 2. MMN difference waveforms showing the interaction between attention and target type on the mismatch effect. The interaction plot shows that the MMN to the unattended stimulus stream is enhanced when subjects are tracking speech targets, but not when subjects are tracking non-speech targets. Left panel shows distribution of MMN across subjects and conditions; lower panel shows the difference waveforms under the different attention and target conditions. Topoplots shows the general distribution of the difference wave at its peak (224ms). The waveforms represent the average voltage in the raw data for the PCA-determined electrode region for MMN.

Vocabulary Test, 3rd edition (PPVT-3) (Dunn & Dunn, 1997). Spanish language abilities were tested for bilingual children with *Clinical Evaluation of Language Fundamentals-Spanish*, Fourth Edition (CELF-4) (Wiig, Semel & Secord, 2006) and *Test de Vocabulario en Imágenes Peabody* (TVIP) (Dunn & Dunn, 1986). The mean first words in English use was reported as 15.6 months (SD = 10.6, range = 7–36), and the mean first words in Spanish use was 21.7 (SD = 17.6, range = 8–60). The mean Spanish use score was calculated similarly to the calculation for adults, using home environment (1–7 point scale), order of acquisition (Spanish first = 1, simultaneous = 4, English first = 7) and preschool (Spanish = 1, both = 4, English = 7) showing a mean of 4.39 (SD = 1.31, range = 2.26–6.4). The mean and standard deviations are shown in Tables 3a and 3b. All children scored in the normal range (within 1 SD, 15 points) in at least one language. Three of the four children receiving more English than Spanish in the home also showed weak Spanish scores (< 85).

The Spearman rank order correlation between Spanish use score and Spanish CELF scores was significant ($N = 13$, $r = -0.72$, $t(N-2) = -3.42$, $p < .01$). The correlations between Spanish use score

and the TVIP (Spanish version of the PPVT) ($N = 13$, $r = -0.37$, $t(N-2) = -1.35$, $p = 0.2$), and the correlation between the Spanish use score and the English CELF ($N = 14$, $r = 0.1$, $t(N-2) = 0.35$, $p = 0.73$) were not significant.

3.1.2. Procedure, equipment and stimuli

The procedure and stimuli were identical to that of the adults.

3.2 Results

The children showed the typical pattern of ERP responses (AEPs) to auditory stimuli, with a large fronto-central positivity (P100), followed by a negativity (N250) (Shafer et al., 2000; Shafer et al., 2010). As with the adult data, we conducted a temporal PCA on the difference waves (deviants – standard), with LANGUAGE (monolingual, bilingual), ATTENTION, and TARGET as conditions. The temporal PCA (PROMAX rotation, covariance matrix, $k = 3$) retained 14 temporal factors, accounting for 93% of the total variance. Of these, only the first three accounted for more than 5% of variance. TF1 accounted for

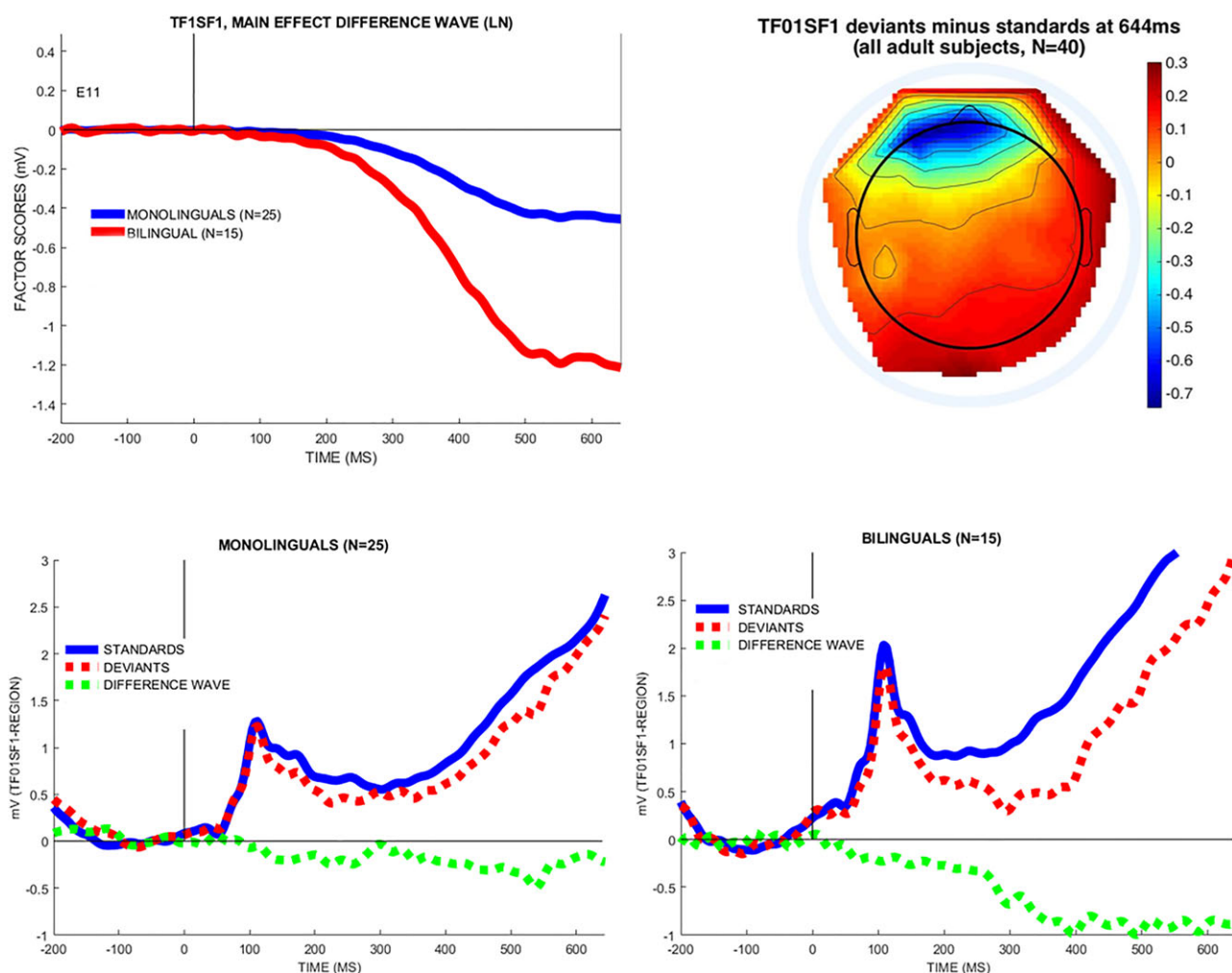


Fig. 3. The top, left panel shows the difference waves for the factor scores for monolinguals vs. bilinguals in the Late Negativity ERP; the right top panel shows the spatial distribution of this main effect difference wave. The bottom row shows the raw voltage average of TF01SF1 electrodes exceeding 0.6 factor loadings (channels 6, 7, 10, 11, 14) for monolinguals (left panel) vs. bilinguals (right panel).

47% of the variance, with the highest loading at 556 ms. The topography and time course of TF1 was very similar to the first temporal factor corresponding to the LN in the adult data. TF2 accounted for 11% of the variance, and corresponded temporally with the negative peak of the difference wave in the grand average, at 280 ms. TF3 peaking at 164 ms and accounting for 7% did not have a clear, interpretable spatial distribution. Statistical analysis of its factor scores resulted in no significant effects, and it was therefore not analyzed further. Spatial ICA was conducted on TF1 and TF2, resulting in, six retained spatial factors. These time windows and electrode regions were used to compute single subject averages for each condition in the undecomposed data.

3.2.1 Mismatch negativity

TF2, peaking at 280 ms, accounted for the earliest latency mismatch effect for children (see Figure 5). The first spatial sub-factor of TF2 closely matched the difference score topographical distribution in the raw data (not shown). Figure 5, top panel shows the time course of the difference wave in the factor analysis and the topographical scalp distribution of this factor in voltage. An electrode region was constructed using electrodes E2, E3, E4, E6, E7, E8, E9, E12, E13, E58, E62 with factor loadings exceeding 0.6 in

TF2SF1; the lower panel in Figure 5 displays no apparent difference between the two groups of children.

The mean voltage of the undecomposed data in the time window 244-312 ms was calculated from EGI electrode sites E2, E3, E4, E6, E7, E8, E9, E12, E13, E58, E62 (with factor loadings exceeding 0.6 for TF2SF1) for each subject and condition and were analyzed in a repeated measures ANOVA with LANGUAGE x ATTENTION x TARGET. Results revealed a main effect of intercept ($F(1,28) = 7.9, p < .01$); as the dependent measures were difference scores (deviant minus standard), this translates into a main effect of mismatch. No other main effects or interactions were observed. The same analysis conducted after excluding four bilingual child participants who were predominantly English in their usage, based on parent report on the LBQ, did not change these statistics.

We also examined whether the MMN-effect (mean across all four conditions) correlated with a Spanish use/proficiency score (standard language scores were transformed and added to the composite use/proficiency measure described in the methods on the following scale: Spanish > English standard scores by more than 1 SD (equivalent to 15) = 1; English > Spanish standard scores by more than 1 SD = 7; Spanish = English scores, within

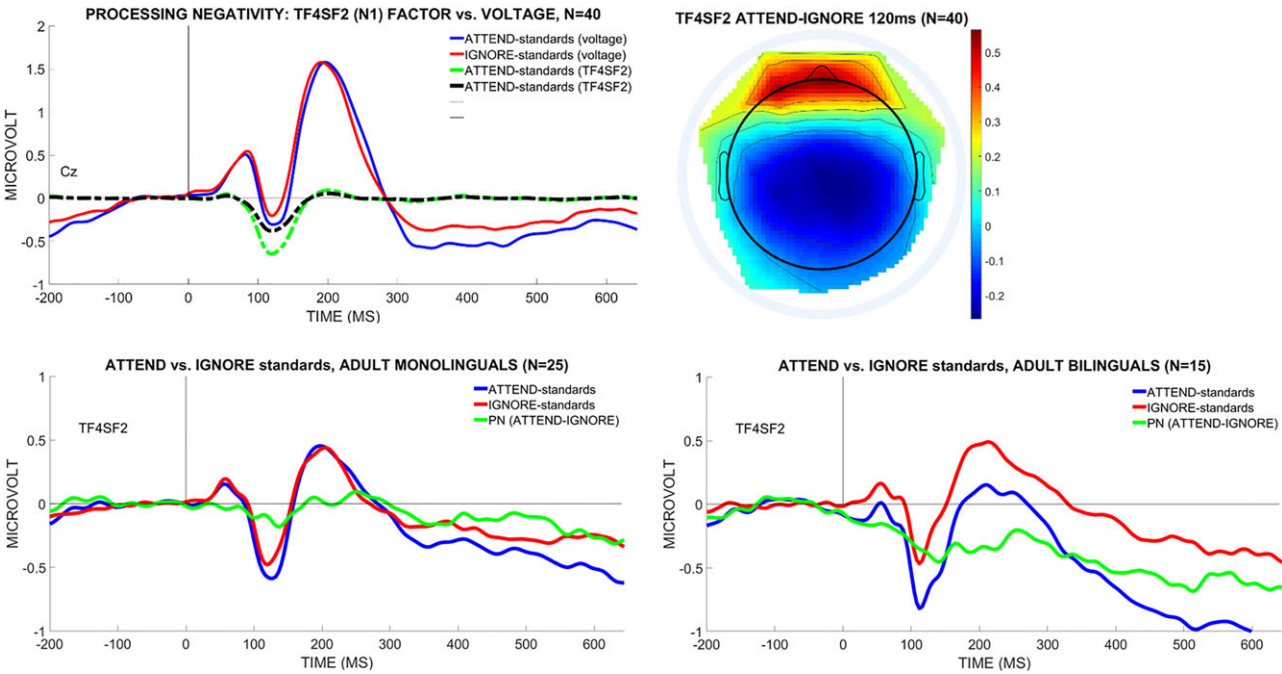


Fig. 4. Top panel: main effect waveforms for standards under Attend and Ignore conditions; superimposed on the N1 temporo-spatial factor; right top panel shows the distribution of the main effect difference scores (Attend minus Ignore). Lower panel shows the mean voltage for the electrode region isolated by the PCA (E22, E29, E30, E34, E42, E43, E47) for monolinguals vs. bilinguals with the PN difference wave in green.

Table 2. Language background for children calculated from 7-point scale. The number of questions used to calculate the response is shown in column No. Questions. The number of participants who responded in the range is provided. (Mostly Spanish=1 or 2; Equal =3,4, or 5; Mostly English =6 or 7). Three of the children received more input in Spanish than English (values < 3.5), nine showed balanced input (values between 3.5 and 5) and two showed more English than Spanish input (values greater than 5). Twelve out of the 15 children reported that his/her communication was best in English. The majority of children showed more Spanish than English or balanced input at home, whereas in other contexts, they received equal or more English than Spanish.

Category	# of Questions	Mostly Spanish	Equal	Mostly English	Not Reported/NA
Home	11	5	5	4	0
Community	12	0	4	9	1
Media	8	1	3	9	1
Literacy	8	0	2	11	1
First spoken Lang	1	6	5	3	0

1 SD = 4). A Spearman rank order correlation showed no significant relationship between these measures ($r = -0.32$, $t(N-2) = -1.22$, $p = 0.24$).

3.2.2 Late negativity

The temporal PCA identified the later part of the waveform as a separate event. Figure 6 below shows the main effect of the Attend variable in the first spatial sub-factor of this temporal factor, along with its spatial scalp distribution. As is apparent, there is a clear effect of attention in this later interval but no difference between monolingual and bilingual children. The lower panel in Figure 6 shows the electrode region calculated from the electrodes with factor loadings greater than 0.6 in the spatial factor, and also shows no apparent difference between the groups.

For the undecomposed voltage data, we calculated the mean of electrodes E6, E7, E8, E10, E11, E12 in the time-range 400-648 ms (those times and sites with factor loadings exceeding .6). Again, only a main effect of intercept was observed ($F(1,28) = 9.03$, $p < .01$).

Table 3. Standard test scores.

Test	Mean	SD	Range
A. Standard test scores for bilingual children.			
Mean English CELF	111	16	82-133
Mean Spanish CELF	100	25	57-136
Mean PPVT	106	18	77-123
Mean TVIP	91	22	55-120
B. Standard test scores for monolingual children. (Scores are missing for two children, but the parental report indicated that they were performing at or above grade level.)			
Mean English CELF	106	12	90-122
Mean PPVT	100	13	80-118

3.2.3 Processing negativity

Figure 7 shows the responses to standards at Fz for the two groups under the two ATTENTION conditions, along with the scalp topography of the difference.

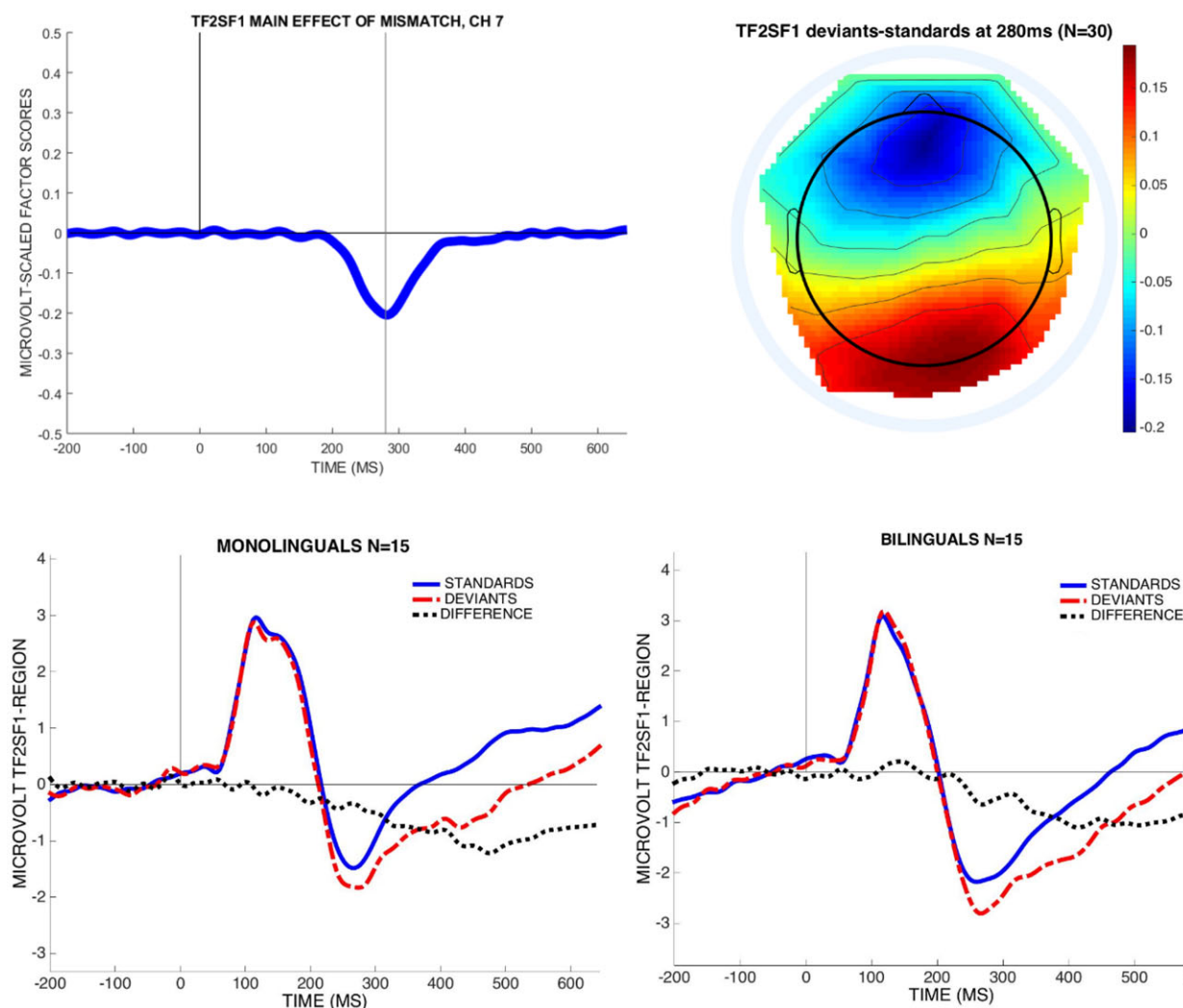


Fig. 5. Upper panel: Main effect factor waveform for the first negative peak and spatial distribution of this effect in the factor. Lower panel: mean waveforms for each child group for standard and deviants computed for the region in blue in the topographical plot.

In order to capture the PN directly, we subtracted the response to the standards in the Attend condition from the response to the standards in the Ignore condition as input to a temporo-spatial PCA followed by voltage analysis constrained by the factor solution. The initial temporal PCA retained 8 factors, but only the first three accounted for more than 5% variance (TF1, 644 ms: 60%; TF2, 264ms: 10%; TF3, 168 ms, 9%). The third temporal factor TF3 at 168 ms matched the temporal and spatial distribution of the difference between attended and ignored standard stimuli observed in the voltage data. Four spatial factors were retained in the spatial step; the second spatial factor TF3SF2 had a distribution similar to that of adults and was selected for analysis, see Figure 8.

For analysis, we chose the time window during which temporal factor loadings exceeded 0.6 (100-216 ms), and electrodes that exceeded 0.5 from the spatial decomposition (electrodes E5, E17, E18, E1, E22, E25, E30, E43, E47, E54, E55). A repeated measures ANOVA with LANGUAGE as between-subject resulted in a significant intercept ($F(1,28) = 25.06$, $p < 0.0001$), but no main effects or interactions. In other words, the children exhibited a clear PN, but there was no difference between monolingual and bilingual children. Excluding the four bilingual subjects who

were predominantly English in their usage did not change these statistics.

3.3 Discussion

Children exhibited the same pattern of two ERP responses to mismatch as adults: an MMN and a LN. The onset latency of the MMN effect, however, was later (280 ms) than in adults (208 ms). A later MMN latency for children than for adults is consistent with previous studies (Shafer, Morr, Kreuzer & Kurtzberg, 2000) showing maturational effect in the MMN. There was no difference between monolingual and bilingual children in the MMN, and no effect on MMN by attention condition or target stimulus. The LN in children had a similar distribution and time course to that of adults. However, the LN was not modulated by attention in the children. Similarly, both groups of children exhibited significant Processing Negativity (PN) when attending to the auditory modality compared to ignoring the auditory stimuli and watching the muted movie in the Ignore condition. However, unlike the adult groups, the two groups of children did not differ in the amplitude of the PN.

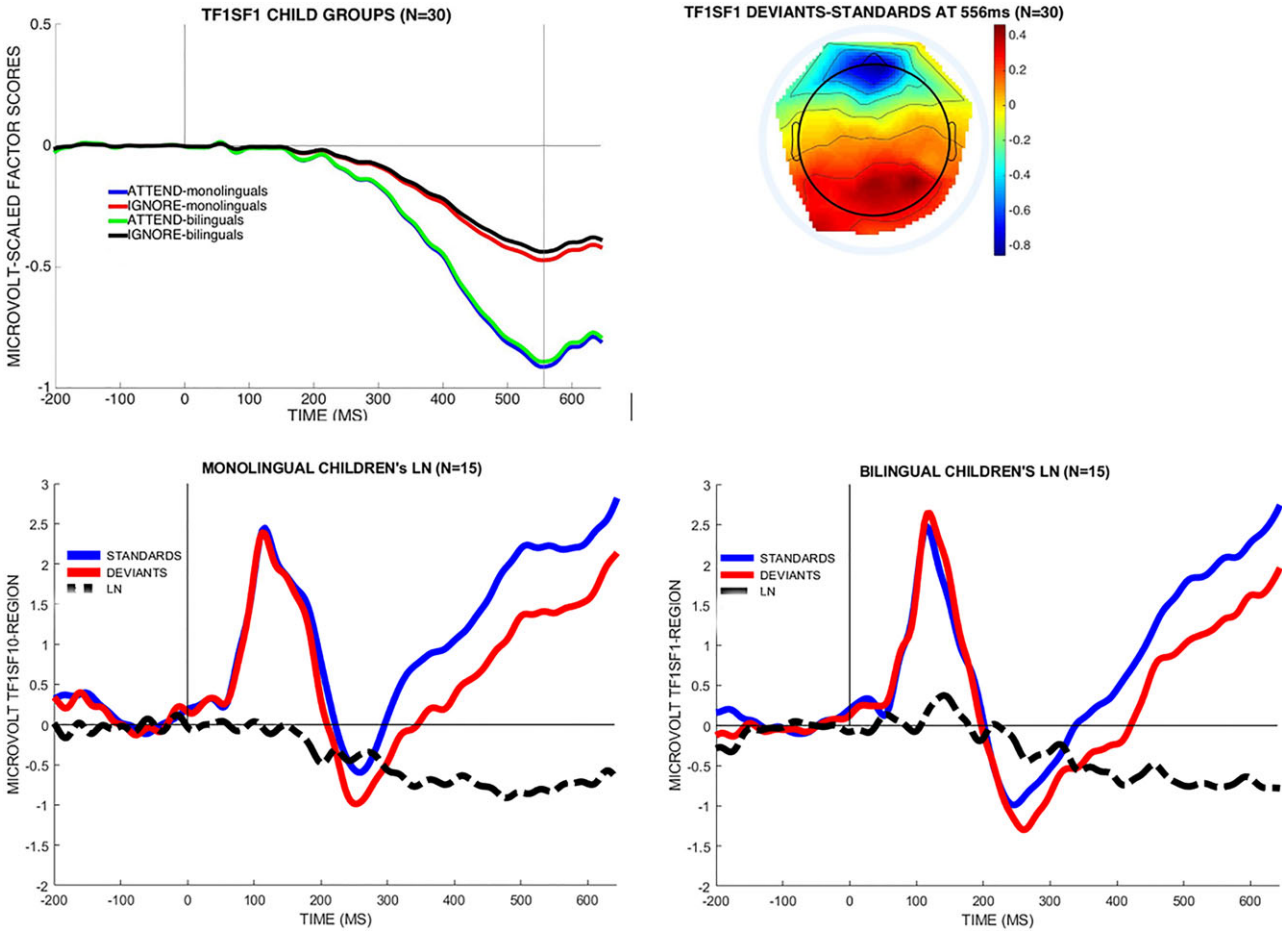


Fig. 6. Late Negativity in child groups. The left top panel shows the factor waveform by ATTENTION and LANGUAGE; right top panel shows the scalp distribution of the factor. Lower panels show the mean waveforms for the spatial region over time by the standard and deviant stimulus conditions along with the difference waveform.

4. General discussion

Our first aim was to replicate and extend findings from our previous study of bilingual speech processing. In Hisagi et al. (2015), we had observed a smaller MMN to the English /i/ versus /ε/ contrast in Spanish learners of English, whether they learned English at or before five years (early Spanish–English bilinguals) of age or after 14 years (late Spanish–English bilinguals) of age. The current study did not replicate this finding. Specifically, early Spanish–English bilinguals showed an equivalent amplitude MMN to monolingual English speakers. It is possible that this discrepancy reflects a difference in our samples. As in the current study, all participants in Hisagi et al. (2015) reported learning both languages before 5 years of age, but we did not obtain language use ratings in various settings. In the current sample, we used a much more detailed questionnaire. We know that all but two of the adult and two child participants had been exposed to English by pre-school. The adult bilinguals generally showed greater use of Spanish at home and in the residential community, but all reported more use of English in school and literacy contexts. The child bilingual participants, however, were evenly split with 6 favoring Spanish in the home and 8 favoring English. A substantial number of our participants initially were exposed only to Spanish, but most report English as the dominant language. It is possible that the early bilinguals in Hisagi et al. (2015) favored Spanish as adults to a greater extent than the early bilinguals in the current study.

Age of acquisition is unlikely to be a factor, since the adults in the two studies report similar experience.

Our second aim was to examine whether bilingual children differed from their monolingual peers on neural measures of automatic speech sound processing and attentional resource allocation. We found that they were identical to monolinguals in all respects. A number of studies of bilingual children suggest that it can take four and a half to six and a half years of immersion in school to fully acquire a second language (Paradis & Jia, 2017). Previous studies using neural measures have shown that one to two years of second language experience in four- to six-year old children is insufficient to lead to native-like responses (Rinker et al., 2010, 2017; but see Cheour et al., 2002 and Peltola et al., 2005, 2007), but the children in those studies only had one year of public school (beginning at five years of age in the US and 6 years of age in Germany). The current study is consistent with these findings, in that the children in our current sample had four to six years of experience with English in the school system, and also reported to be using mostly English in the residential community, whereas Spanish was used mostly at home. Thus, our current sample of children might have had more extensive experience with English than the participants reported on in Hisagi et al. (2015).

Our third goal was to examine the degree to which attention modulated the MMN index of speech sound discrimination, and whether any attentional effects interacted with the

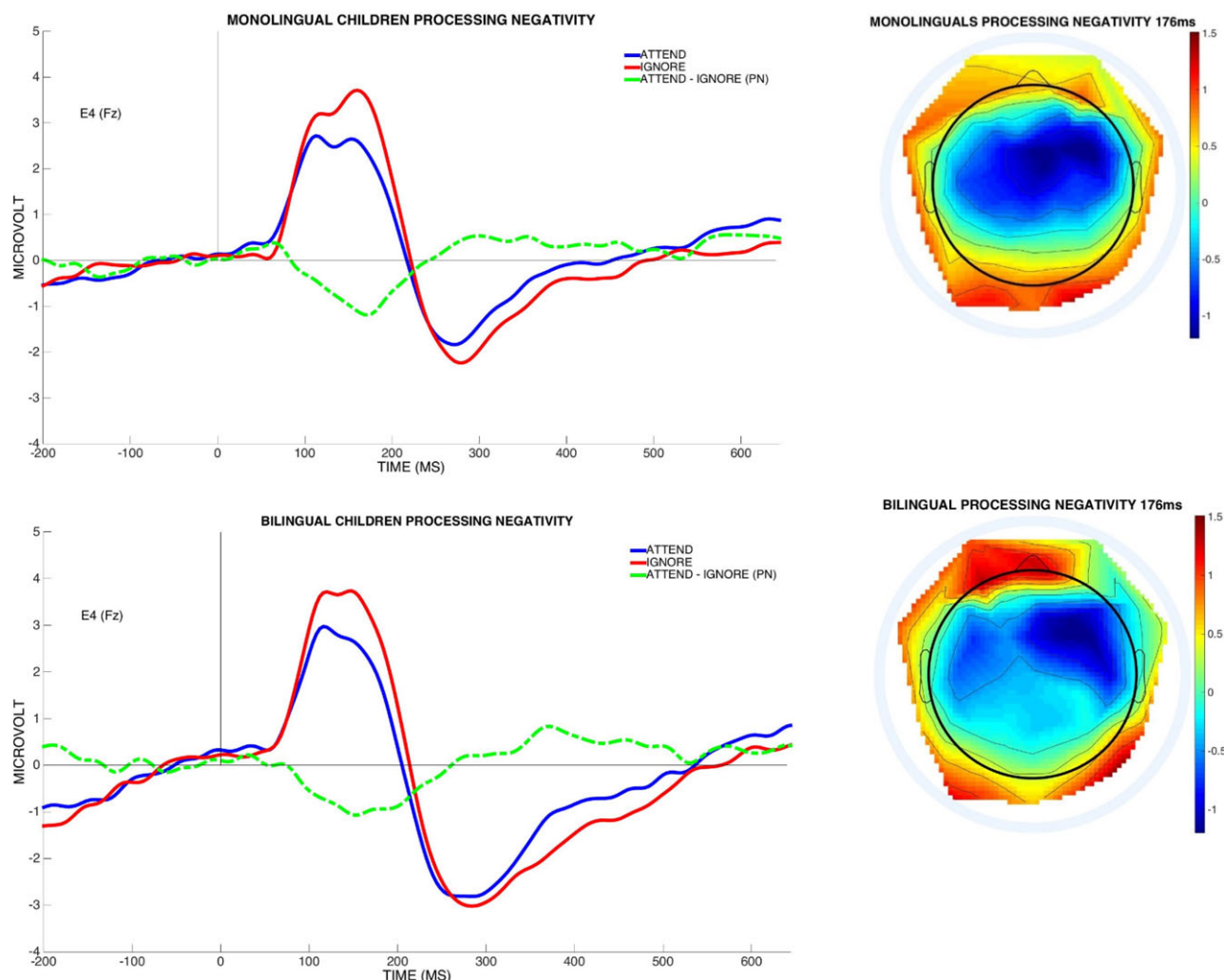


Fig. 7. Waveforms for standards at Fz for Attend vs. Ignore conditions and the spatial distribution of the difference for monolingual vs. bilingual children.

monolingual/bilingual difference. In Hisagi et al. (2015), the smaller MMN to the English vowel contrast in bilinguals suggested that they were not fully automatic in discrimination of this vowel contrast. A goal of the current study was to find out whether directing the L2 participants' attention to the stimuli would enhance their MMN response. We observed an increase in MMN amplitude to the vowel contrast when the adult participants attended to the speech target, which suggests that the vowel contrast used in this study was sufficiently difficult to benefit from attention. However, this effect was observed for both language groups: that is, being bilingual afforded no advantage. We also found an increase in LN when both groups of adults attended to tone targets, which suggests that MMN and LN are affected differently by attentional modulation.

Another goal was to examine the interaction between MMN, attention, and being bilingual vs. monolingual. We observed no effect of attention on the MMN. However, adult bilinguals exhibited a significantly larger PN amplitude than monolinguals. This suggests that adult bilinguals are more attentive to the auditory environment (Garcia-Sierra et al., 2012; Peltola et al., 2005, 2007), and allocate more resources in order to achieve the same linguistic efficiency; this finding may be related to the greater cognitive flexibility claimed for bilinguals than their monolingual

counterparts (Molnar et al., 2014). This inference is also supported by the fact that adult bilinguals had a significantly larger LN response than adult monolinguals, suggesting a greater involuntary attention shift toward vowel discrimination. This pattern is similar to the finding of Ortiz-Mantilla et al. (2010). Thus, the enhanced "bilingual LN" might indicate that bilingual listeners more often need to make additional decisions about the speech: namely, what the target language is. It also converges with other recent neurophysiological findings of increased attentional control mechanisms developing in bilingual children, seen as a 'spill-over' effect into non-verbal tasks (Arredondo et al., 2017), and more top-down auditory attentional control for bilingual adults relative to monolinguals (Krizman, Skoe, Marian & Kraus, 2014). These studies suggest that the need for increased attentional control in the process of selecting the target language may underlie these neural patterns.

Turning to Experiment II, we observed no effect of attention on children's MMN amplitude, but, unlike adults, we also did not observe an effect of attention manipulation on the children's LN. The failure to see this pattern in the children may indicate immaturity, or that the children are more English-dominant than the adults. This finding matches our previous study of 8-10 year-old children using the same stimuli (but delivered

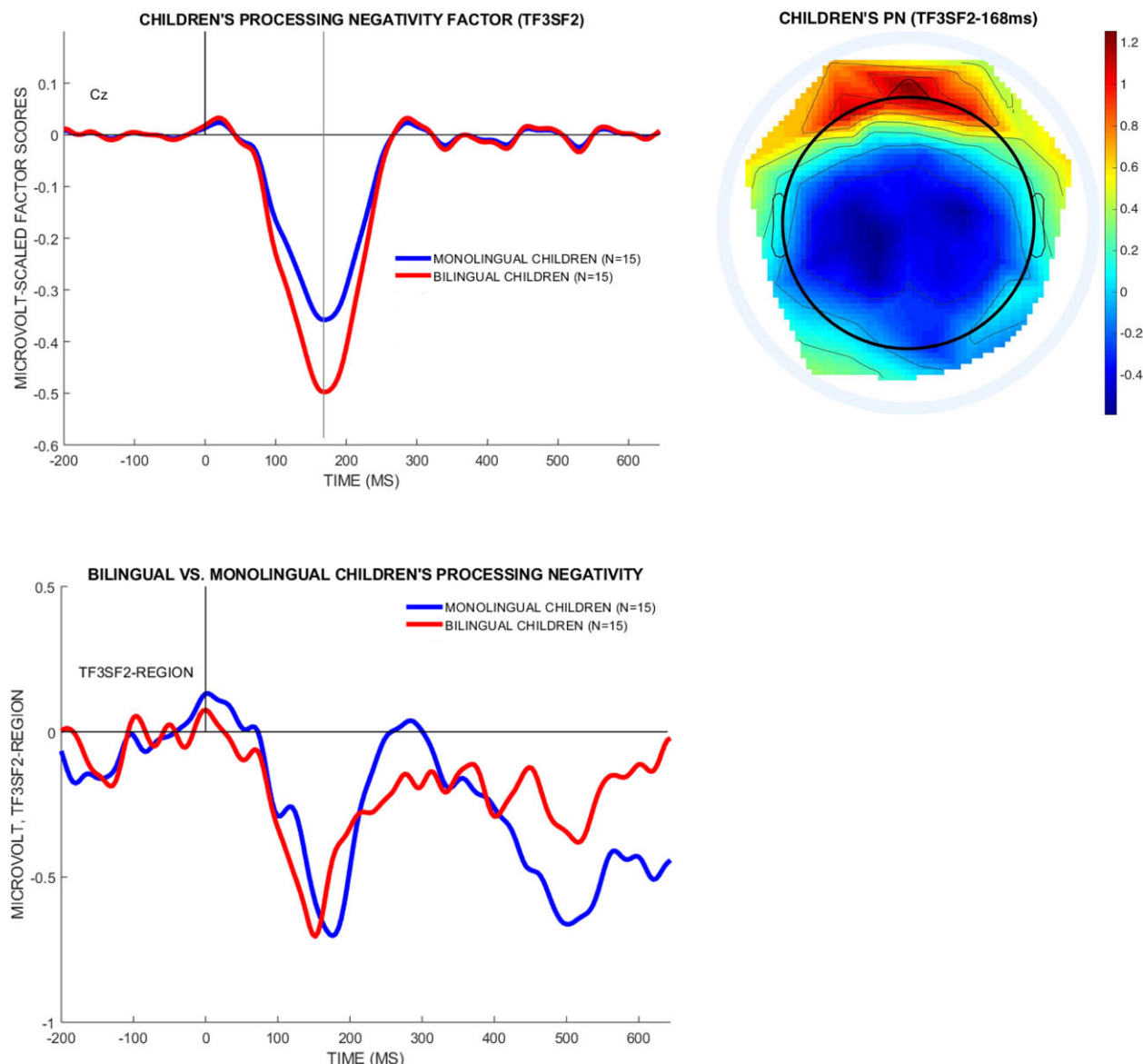


Fig. 8. Processing Negativity factor waveforms and scalp distribution (upper panel) and voltage waveforms for the ICA-defined region by LANGUAGE group.

over ear-insert phones) (Shafer et al., 2005). Specifically, attention to the speech stimuli did not affect MMN amplitude. Even so, the reason for enhanced MMN in adults when attending to speech targets, but not in children, remains unaccounted for. We do know that grade school children weigh speech cues differently from adults (Nitttrouer & Miller, 1998), but we would expect attention to have a greater effect for children than adults.

The children also showed the PN effect to the attention conditions; but again, no difference was found between the monolingual and bilingual children. These findings suggest that three-to five years of English input in the NYC public schools for the bilingual Spanish–English children was sufficient to allow for native-like speech processing skills in English. The later timing of the mismatch responses, however, indicates that speech processing is not fully mature in this age group.


We introduced three possible models of bilinguals' speech perception. Our findings are clearly inconsistent with the model that claims that bilinguals compromise between the two phonological systems (Cutler et al., 1987; Snijders et al., 2007), considering that

we observed native-like L2 processing. Our results are consistent with the other two models (Elman et al., 1977; Gonzales & Lotto, 2013; Williams, 1977), but cannot fully address which of the two models is better. We only had data for an L2 contrast and, thus, it is possible that our cohort of bilinguals favored English over Spanish. It is also possible that the robust L2 speech processing indicated that the early bilinguals we tested were able to adjust their phonological contrasts based on linguistic contexts (see Casillas & Simonet, 2018). We did not attempt to manipulate context, and the experimental setting decidedly favored English, with the consent and instructions carried out in English. Future studies that test processing in both Spanish and English and that manipulate context will be necessary to select between these models. Finally, the lack of a relationship between the neural discriminative measure, MMN and our language use/proficiency measure may be the result of fairly high English proficiency for all our participants. It will be important in a future study to increase the number of participants and examine a wider range of proficiency levels and amount of use in English and Spanish

to further explore how these factors influence speech processing in early bilinguals.

The current study did not address whether these bilinguals had maintained L1 phonological categories. This limitation could be addressed by comparing mismatch responses between Spanish monolinguals and bilinguals to a native Spanish contrast that is not found in English, although a Spanish consonant contrast might serve as a better test, given that the five Spanish vowels are assimilated into five non-overlapping English phoneme categories. Alternatively, we could test a Mandarin–English bilingual population on the Mandarin /i/-/y/ versus our English /i/ to /ε/ (which is difficult for Mandarin listeners), since we know monolingual English speakers show poorer discrimination of the Mandarin contrast (Yu, Shafer & Sussman, 2017).

In conclusion, our results showed that adult and child bilinguals who began acquiring English by five years of age show native-like neural discrimination of a spectrally-difficult English vowel contrast. However, they also showed differences from monolinguals in neural measures that are likely to be related to attentional processes. These findings support the claim that bilingual experience leads to differences in executive functions, such as attentional control.

Author ORCIDs.  Arild Hestvik, 0000-0003-4561-7584; Valerie Shafer, 0000-0001-8551-1878.

Acknowledgements. This research was funded by the National Institutes of Health, National Institute of Child Health and Human Development grant HD46193 (Valerie Shafer, P.I.)

References

- Archila-Suerte P, Zevin J, Bunta F and Hernandez AE (2012) Age of acquisition and proficiency in a second language independently influence the perception of non-native speech. *Bilingualism* 15(1), 190–201. <https://doi.org/10.1017/S1366728911000125>
- Arredondo MM, Hu XS, Satterfield T and Kovelman I (2017) Bilingualism alters children's frontal lobe functioning for attentional control. *Developmental Science* 20(3). <https://doi.org/10.1111/desc.12377>
- Astheimer LB, Berkes M and Bialystok E (2016) Differential Allocation of Attention During Speech Perception in Monolingual and Bilingual Listeners. *Language, Cognition and Neuroscience* 31(2), 196–205. <https://doi.org/10.1080/23273798.2015.1083114>
- Baker W, Trofimovich P, Flege JE, Mack M and Halter R (2008) Child-adult differences in second-language phonological learning: the role of cross-language similarity. *Language and Speech* 51(Pt 4), 317–342.
- Bell AJ and Sejnowski TJ (1995) An Information-Maximization Approach to Blind Separation and Blind Deconvolution. *Neural Computation* 7(6), 1129–1159. <https://doi.org/10.1162/neco.1995.7.6.1129>
- Bialystok E, Craik FIM and Luk G (2012) Bilingualism: consequences for mind and brain. *Trends in Cognitive Sciences* 16(4), 240–250. <https://doi.org/10.1016/j.tics.2012.03.001>
- Brice AE, Gorman BK and Leung CB (2013) Spanish-English speech perception in children and adults: developmental trends. *Clinical Linguistics & Phonetics* 27(3), 220–234. <https://doi.org/10.3109/02699206.2012.757805>
- Casillas JV and Simonet M (2018) Perceptual categorization and bilingual language modes: Assessing the double phonemic boundary in early and late bilinguals. *Journal of Phonetics* 71, 51–64. <https://doi.org/10.1016/j.wocn.2018.07.002>
- Ceponiene R, Lepistö T, Soininen M, Aronen E, Alku P and Näätänen R (2004) Event-related potentials associated with sound discrimination versus novelty detection in children. *Psychophysiology* 41(1), 130–141. Retrieved from <http://dx.doi.org/10.1111/j.1469-8986.2003.00138.x>
- Cheour M, Korpilahti P, Martynova O and Lang AH (2001) Mismatch Negativity and Late Discriminative Negativity in Investigating Speech Perception and Learning in Children and Infants. *Audiology and Neurotology* 6(1), 2–11. Retrieved from <http://www.karger.com/DOI/10.1159/000046804>
- Cheour M, Shestakova A, Alku P, Ceponiene R and Näätänen R (2002) Mismatch negativity shows that 3–6-year-old children can learn to discriminate non-native speech sounds within two months. *Neuroscience Letters* 325(3), 187–190. [https://doi.org/10.1016/S0304-3940\(02\)00269-0](https://doi.org/10.1016/S0304-3940(02)00269-0)
- Cutler A, Norris D and Williams J (1987) A note on the role of phonological expectations in speech segmentation. *Journal of Memory and Language* 26(4), 480–487. [https://doi.org/10.1016/0749-596X\(87\)90103-3](https://doi.org/10.1016/0749-596X(87)90103-3)
- Darcy I and Krüger F (2012) Vowel perception and production in Turkish children acquiring L2 German. *Journal of Phonetics* 40(4), 568–581. <https://doi.org/10.1016/j.wocn.2012.05.001>
- Datta H, Shafer VL, Morr ML, Kurtzberg D and Schwartz RG (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. [https://doi.org/10.1044/1092-4388\(2009/08-0123\)](https://doi.org/10.1044/1092-4388(2009/08-0123))
- Dien J (2010) The ERP PCA Toolkit: An open source program for advanced statistical analysis of event-related potential data. *Journal of Neuroscience Methods*. <https://doi.org/10.1016/j.jneumeth.2009.12.009>
- Dien J (2012) Applying Principal Components Analysis to Event-Related Potentials: A Tutorial. *Developmental Neuropsychology* 37(6), 497–517. <https://doi.org/10.1080/87565641.2012.697503>
- Dien J and Frishkoff GA (2005) Principal components analysis of ERP data. In T Handy (Ed.), *Event-Related Potentials: A Methods Handbook*. Cambridge: MIT Press.
- Dien J, Spencer KM and Donchin E (2003) Localization of the event-related potential novelty response as defined by principal components analysis. *Cognitive Brain Research* 17(3), 637–650. [https://doi.org/10.1016/S0926-6410\(03\)00188-5](https://doi.org/10.1016/S0926-6410(03)00188-5)
- Dunn LM and Dunn LM (1986) *Test de Vocabulario en Imágenes Peabody (TVIP)*. Circle Pines, MN: American Guidance Service.
- Dunn LM and Dunn LM (1997) *Peabody Picture Vocabulary Test* (3rd ed.). Circle Pines, MN: American Guidance Service.
- Elman JL, Diehl RL and Buchwald SE (1977) Perceptual switching in bilinguals. *The Journal of the Acoustical Society of America* 61(S1), S65–S66. <https://doi.org/10.1121/1.2015834>
- Farnia F and Geva E (2011) Cognitive correlates of vocabulary growth in English language learners. *Applied Psycholinguistics* 32(04), 711–738. <https://doi.org/10.1017/S0142716411000038>
- García-Sierra A, Ramírez-Esparza N, Silva-Pereyra J, Siard J and Champlin CA (2012) Assessing the double phonemic representation in bilingual speakers of Spanish and English: An electrophysiological study. *Brain and Language* 121(3), 194–205. <https://doi.org/10.1016/j.bandl.2012.03.008>
- Gonzales K and Lotto AJ (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>
- Hansen JC and Hillyard SA (1980) Endogenous brain potentials associated with selective auditory attention. *Electroencephalography and Clinical Neurophysiology* 49(3–4), 277–290. [https://doi.org/10.1016/0013-4694\(80\)90222-9](https://doi.org/10.1016/0013-4694(80)90222-9)
- Hendrickson AE and White PO (1964) Promax: a Quick Method for Rotation To Oblique Simple Structure. *British Journal of Statistical Psychology* 17(1), 65–70. <https://doi.org/10.1111/j.2044-8317.1964.tb00244.x>
- Hestvik A and Durvasula K (2016) Neurobiological evidence for voicing underspecification in English. *Brain and Language*. <https://doi.org/10.1016/j.bandl.2015.10.007>
- Hisagi M, Garrido-Nag K, Datta H and Shafer VL (2015) ERP indices of vowel processing in Spanish–English bilinguals. *Bilingualism: Language and Cognition* 15(02), 1–19. <https://doi.org/10.1017/S1366728914000170>
- Horn JL (1965) A rationale and test for the number of factors in factor analysis. *Psychometrika* 30(2), 179–185.
- Jia G and Fuse A (2007) Acquisition of English Grammatical Morphology by Native Mandarin-Speaking Children and Adolescents: Age-Related Differences. *Journal of Speech Language and Hearing Research* 50(5), 1280. [https://doi.org/10.1044/1092-4388\(2007/090\)](https://doi.org/10.1044/1092-4388(2007/090))

- Jusczyk PW, Cutler A and Redanz NJ (1993) Infants' Preference for the Predominant Stress Patterns of English Words. *Child Development* 64(3), 675–687. <https://doi.org/10.1111/j.1467-8624.1993.tb02935.x>
- Krizman J, Skoe E, Marian V and Kraus N (2014) Bilingualism increases neural response consistency and attentional control: Evidence for sensory and cognitive coupling. *Brain and Language* 128(1), 34–40. <https://doi.org/10.1016/j.bandl.2013.11.006>
- Kuhl PK, Conboy BT, Coffey-Corina S, Padden D, Rivera-Gaxiola M and Nelson T (2008) Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 363(1493), 979–1000. <https://doi.org/10.1098/rstb.2007.2154>
- Kuipers JR and Thierry G (2015) Bilingualism and increased attention to speech: Evidence from event-related potentials. *Brain and Language* 149, 27–32. <https://doi.org/10.1016/j.bandl.2015.07.004>
- Kujala T and Leminen M (2017) Low-level neural auditory discrimination dysfunctions in specific language impairment – A review on mismatch negativity findings. *Developmental Cognitive Neuroscience* 28, 65–75. <https://doi.org/10.1016/j.DCN.2017.10.005>
- Lee C-Y, Yen H, Yeh P, Lin W-H, Cheng Y-Y, Tzeng Y-L and Wu H-C (2012) Mismatch responses to lexical tone, initial consonant, and vowel in Mandarin-speaking preschoolers. *Neuropsychologia* 50(14), 3228–3239. <https://doi.org/10.1016/j.NEUROPSYCHOLOGIA.2012.08.025>
- Luck SJ (2014) *An introduction to the event-related potential technique* (2nd ed.). Cambridge, Mass: MIT Press. Retrieved from <https://mitpress.mit.edu/books/introduction-event-related-potential-technique-0>
- Luck SJ and Gaspelin N (2017) How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology* 54(1), 146–157. <https://doi.org/10.1111/psyp.12639>
- Masapollo M and Polka L (2014) Asymmetries in vowel perception: Do they arise from focalization, perceptual magnets, or both? *The Journal of the Acoustical Society of America* 135(4), 2224–2224. <https://doi.org/10.1121/1.4877267>
- McCarthy KM, Mahon M, Rosen S and Evans BG (2014) Speech perception and production by sequential bilingual children: A longitudinal study of voice onset time acquisition. *Child Development* 85(5), 1965–1980. <https://doi.org/10.1111/cdev.12275>
- Molnar M, Polka L, Baum S and Steinhauer K (2014) Learning two languages from birth shapes pre-attentive processing of vowel categories: Electrophysiological correlates of vowel discrimination in monolinguals and simultaneous bilinguals. *Bilingualism: Language and Cognition* 17 (03), 526–541. <https://doi.org/10.1017/S136672891300062X>
- Moreno S and Lee Y (2015) Short-term Second Language and Music Training Induces Lasting Functional Brain Changes in Early Childhood. *Child Development* 82(2), 389–400. <https://doi.org/10.1530/ERC-14-0411.Persistent>
- Näätänen R (1982) Processing negativity: An evoked-potential reflection. *Psychological Bulletin* 92(3), 605–640. <https://doi.org/10.1037/0033-2909.92.3.605>
- Näätänen R (1990) The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences* 13, 201–288.
- Näätänen R, Paavilainen P, Rinne T and Alho K (2007) The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology* 118(12), 2544–90. <https://doi.org/10.1016/j.clinph.2007.04.026>
- Näätänen R, Sussman ES, Salisbury D and Shafer VL (2014) Mismatch negativity (MMN) as an index of cognitive dysfunction. *Brain Topography* 27(4), 451–466. <https://doi.org/10.1007/s10548-014-0374-6>
- Nittrouer S (2002) Learning to perceive speech: How fricative perception changes, and how it stays the same. *The Journal of the Acoustical Society of America* 112(2), 711–719. <https://doi.org/10.1121/1.1496082>
- Nittrouer S (2006) Children hear the forest. *The Journal of the Acoustical Society of America* 120(4), 1799–1802. <https://doi.org/10.1121/1.2335273>
- Nittrouer S and Miller ME (1998) Developmental weighting shifts for noise components of fricative-vowel syllables. <http://Dx.Doi.Org.Udel.Idm.Ocl.Org/10.1121/1.419730>. <https://doi.org/10.1121/1.419730>
- Ortiz-Mantilla S, Choudhury N, Alvarez B and Benasich AA (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>
- Paradis J and Jia R (2017) Bilingual children's long-term outcomes in English as a second language: language environment factors shape individual differences in catching up with monolinguals. *Developmental Science* 20(1). <https://doi.org/10.1111/desc.12433>
- Peltola MS, Kuntola M, Tamminen H, Hämäläinen H and Aaltonen O (2005) Early exposure to non-native language alters preattentive vowel discrimination. *Neuroscience Letters* 388(3), 121–125. <https://doi.org/10.1016/J.NEULET.2005.06.037>
- Peltola MS, Tuomainen O, Koskinen M and Aaltonen O (2007) The effect of language immersion education on the preattentive perception of native and non-native vowel contrasts. *Journal of Psycholinguistic Research* 36(1), 15–23. <https://doi.org/10.1007/s10936-006-9030-y>
- Piske T, Flege JE, MacKay IRA and Meador D (2002) The production of English vowels by fluent early and late Italian-English bilinguals. *Phonetica* 59(1), 49–71. <https://doi.org/10.1159/000056205>
- Polich J (2012) Neuropsychology of P300. In ES Kappenman and SJ Luck (Eds.), *The Oxford Handbook of Event-Related Potential Components*. Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780195374148.013.0089>
- Putkinen V, Niinikuru R, Lipsanen J, Tervaniemi M and Huotilainen M (2012) Fast Measurement of Auditory Event-Related Potential Profiles in 2–3-Year-Olds. *Developmental Neuropsychology* 37(1), 51–75. <https://doi.org/10.1080/87565641.2011.615873>
- Richman MB (1986) Rotation of Principal Components. *Journal of Climatology* 6, 293–335.
- Rinker T, Alku P, Brosch S and Kiefer M (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 and Language* 113(2), 90–95. <https://doi.org/10.1016/J.BANDL.2010.01.007>
- Rinker T, Shafer VL, Kiefer M, Vidal N and Yu YH (2017) T-complex measures in bilingual Spanish-English and Turkish-German children and monolingual peers. *PloS One* 12(3), e0171992. <https://doi.org/10.1371/journal.pone.0171992>
- Sebastian-Gallés N, Rodríguez-Fornells A, de Diego-Balaguer R and Díaz B (2006) First- and second-language phonological representations in the mental lexicon. *Journal of Cognitive Neuroscience* 18(8), 1277–1291. <https://doi.org/10.1162/jocn.2006.18.8.1277>
- Semel E, Wiig EH and Secord WA (2004) *Clinical evaluation of language fundamentals* (4th ed.). Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company.
- Shafer VL, Morr ML, Datta H, Kurtzberg D and Schwartz RG (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 VL, Morr ML, Kreuzer JA and Kurtzberg D (2000) Maturation of mismatch negativity in school-age children. *Ear Hear* 21(3), 242–251. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&opt=Citation&list_uids=10890733
- Shafer VL, Ponton C, Datta H, Morr ML and Schwartz RG (2007) Neurophysiological indices of attention to speech in children with specific language impairment. *Clinical Neurophysiology* 118(6), 1230–1243. <https://doi.org/10.1016/J.CLINPH.2007.02.023>
- Shafer VL, Yu YH and Datta H (2010) Maturation of Speech Discrimination in 4- to 7-yr-old Children as Indexed by Event-related Potential Mismatch Responses. *Ear and Hearing* 31(6), 735–745. <https://doi.org/10.1097/aud.0b013e3181e5d1a7>
- Shafer VL, Yu YH and Datta H (2011) The development of English vowel perception in monolingual and bilingual infants: Neurophysiological correlates. *Journal of Phonetics* 39(4), 527–545. <https://doi.org/10.1016/J.WOCN.2010.11.010>
- Shafer VL, Yu YH and Wagner M (2015) Maturation of cortical auditory evoked potentials (CAEPs) to speech recorded from frontocentral and temporal sites:

- Three months to eight years of age. *International Journal of Psychophysiology* 95(2), 77–93. <https://doi.org/10.1016/j.ijpsycho.2014.08.1390>
- Shestakova A, Huottilainen M, Ceponiene R and Cheour M** (2003) Event-related potentials associated with second language learning in children. *Clinical Neurophysiology* 114(8), 1507–1512. Retrieved from <http://www.sciencedirect.com/science/article/B6VNP-48XJSR7-2/2/aff2d70f08ebf29097566dc7697c5aa>
- Snijders TM, Kooijman V, Cutler A and Hagoort P** (2007) Neurophysiological evidence of delayed segmentation in a foreign language. *Brain Research* 1178, 106–113. Retrieved from <http://www.sciencedirect.com/science/article/B6SYR-4PFDD57-1/1/3b423de9d9b25d8e7c8b48db8b86b9b7>
- Spencer KM, Dien J and Donchin E** (1999) A componential analysis of the ERP elicited by novel events using a dense electrode array. *Psychophysiology* 36(3), 409–414. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10352565
- Spencer KM, Dien J and Donchin E** (2001) Spatiotemporal analysis of the late ERP responses to deviant stimuli. *Psychophysiology* 38(2), 343–358. <https://doi.org/10.1017/S0048577201000324>
- 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>
- Tamminen H, Peltola MS, Toivonen H, Kujala T and Näätänen R** (2013) Phonological processing differences in bilinguals and monolinguals. *International Journal of Psychophysiology* 87(1), 8–12. <https://doi.org/10.1016/j.ijpsycho.2012.10.003>
- Tataryn DJ, Wood JM and Gorsuch RL** (1999) Setting the Value of K in Promax: A Monte Carlo Study. *Educational and Psychological Measurement* 59(3), 384–391. Retrieved from <http://epm.sagepub.com/cgi/content/abstract/59/3/384>
- Tong X, Lee SMK, Lee MML and Burnham D** (2015) A Tale of Two Features: Perception of Cantonese Lexical Tone and English Lexical Stress in Cantonese-English Bilinguals. *PLoS ONE* 10(11). <https://doi.org/10.1371/journal.pone.0142896>
- Tsukada K, Birdsong D, Bialystok E, Mack M, Sung H and Flege J** (2005) A developmental study of English vowel production and perception by native Korean adults and children. *Journal of Phonetics* 33(3), 263–290. <https://doi.org/10.1016/j.wocn.2004.10.002>
- Werker JF and Curtin S** (2005) PRIMIR: A Developmental Framework of Infant Speech Processing. *Language Learning and Development* 1(2), 197–234. <https://doi.org/10.1080/15475441.2005.9684216>
- Wetzel N and Schröger E** (2014) On the development of auditory distraction: A review. *PsyCh Journal* 3(1), 72–91. <https://doi.org/10.1002/pchj.49>
- Wiig EH, Semel E and Secord WA** (2006) *Clinical Evaluation of Language Fundamentals-Spanish*. The Psychological Corporation/A Harcourt Assessment Company.
- Williams L** (1977) The perception of stop consonant voicing by Spanish-English bilinguals. *Perception & Psychophysics* 21(4), 289–297. <https://doi.org/10.3758/BF03199477>
- Yeni-Komshian GH, Flege JE and Liu S** (2000) Pronunciation proficiency in the first and second languages of Korean-English bilinguals. *Bilingualism: Language and Cognition* 3(2), S1366728900000225. <https://doi.org/10.1017/S1366728900000225>
- Yu YH, Shafer VL and Sussman ES** (2017) Neurophysiological and Behavioral Responses of Mandarin Lexical Tone Processing. *Frontiers in Neuroscience* 11(March), 1–19. <https://doi.org/10.3389/fnins.2017.00095>