Brain & Language 212 (2021) 104890

Contents lists available at ScienceDirect

Brain and Language

journal homepage: www.elsevier.com/locate/b&l

Language learning as a function of infant directed speech (IDS) in Spanish: Testing neural commitment using the positive-MMR☆

Adrián García-Sierra[a,b,*], Nairán Ramírez-Esparza[c], Noelle Wig[a,b], Dylan Robertson[a]

[a] Speech, Language, and Hearing Sciences, University of Connecticut, 2 Alethia Dr. Unit 1085, Storrs, CT 06269, USA
[b] Connecticut Institute for the Brain and Cognitive Science, University of Connecticut, 337 Mansfield Rd Unit 1272, Storrs, CT 06269, USA
[c] Department of Psychological Sciences, University of Connecticut, 406 Babbidge Rd, Unit 1020, Storrs, CT 06269, USA

ARTICLE INFO

Keywords:
Infant-directed-speech
Spanish
Positive-MMR
Language development

ABSTRACT

Spanish-English bilingual families (N = 17) were recruited to assess the association between infant directed speech (IDS) in Spanish and their degree of neural commitment to the Spanish language. IDS was assessed by extracting the caregivers' Vowel Space Area (VSA) from recordings of a storybook reading task done at home. Infants' neural commitment was assessed by extracting the positive mismatch brain response (positive-MMR), an Event-Related Potential (ERP) thought to be indicative of higher attentional processes and early language commitment. A linear mixed model analysis demonstrated that caregivers' VSA predicted the amplitude of the positive-MMR in response to a native speech contrast (Spanish), but not to a non-native speech contrast (Chinese), even after holding other predictors constant (i.e., socioeconomic status, infants' age, and fundamental frequency). Our findings provide support to the view that quality of language exposure fosters language learning, and that this beneficial relationship expands to the bilingual population.

Infants' early language experiences shape neural commitment underlying language learning. For example, García-Sierra et al. (2016) showed that the quantity of language input infants receive in the household systematically varies with their ability to discriminate native speech sounds as measured using brain responses. Quality of language exposure also has been associated with infants' later language learning measured using brain responses (Bosseler, Teinonen, Tervaniemi, & Huotilainen, 2016; Peter, Kalashnikova, Santos, & Burnham, 2016; Uther, Giannakopoulou, & Iverson, 2012; Zangl & Mills, 2007; Zhang et al., 2011). Specifically, research has shown that infants benefit from the exaggerated, hyperarticulated, singsong-like qualities known as infant-directed speech (IDS) (e.g., 'Hiiii babyyy'; Fernald & Kuhl, 1987; Werker et al., 2007). The current investigation uses brain responses to assess the extent to which IDS spoken by Spanish dominant caregivers is associated with infants' Spanish language commitment.

1. Infant directed speech and brain measures

Most studies that have explored infants' speech perception using brain responses have shown that the exaggerated pitch contours and/or exaggerated formants, which differentiate IDS from ADS, facilitate infants' neural processing of linguistic information as early as newborns (Bosseler et al., 2016; Saito et al., 2007; Zhang et al., 2011). For example, Zhang et al. (2011) explored the specific brain signatures associated with the perception of the English vowel /i/ in formant exaggerated form (i.e., IDS) vs. non-exaggerated form (ADS) in a sample of 6–12-month-old infants. Brain measures in the form of Event Related Potentials (ERPs) were recorded using an alternating block design, in which each stimulus (IDS vowel/ADS vowel) was equally probable to occur. The results showed that the IDS vowel produced a significantly larger N250 ERP response; a component that has been associated with phonetic processing (Rivera-Gaxiola, Klarman et al., 2005; Rivera-Gaxiola, Silva-Pereyra et al., 2005) and lexical processing (Mills,

☆ The data that support the findings of this study are available from the corresponding author upon reasonable request: Adrián García-Sierra adrian.garcia-sierra@uconn.edu.

* Corresponding author at: Department of Speech, Language & Hearing Science, Affiliate Faculty, Cognitive Science, University of Connecticut, Storrs, 2 Alethia Dr., Unit 1085, Storrs, CT 06269-1085, USA.

E-mail addresses: adrian.garcia-sierra@uconn.edu (A. García-Sierra), nairan.ramirez@uconn.edu (N. Ramírez-Esparza), noelle.wig@uconn.edu (N. Wig), dylan.robertson@uconn.edu (D. Robertson).

URL: http://speechlanguage.uconn.edu/ (A. García-Sierra).

https://doi.org/10.1016/j.bandl.2020.104890
Received 3 February 2020; Received in revised form 27 October 2020; Accepted 17 November 2020
Available online 8 December 2020
0093-934X/

Coffey-Corina, & Neville, 1997; Zangl & Mills, 2007). Even more, Zhang et al. reported neural synchronization between the N250 and IDS (delta and theta waves) that was heightened in frontal-central and frontal-central-parietal regions, which have been associated with linguistic processing of affective speech. Relevantly, the ERP data suggests that IDS was processed by a cortical neural network connecting Broca's area and perceptual-motor systems, which has been suggested to be associated with learning. Hence, the authors concluded that formant exaggeration, characteristic of IDS, alters infants' brain activation patterns to benefit phonetic processing.

ERPs have also shown that IDS also facilitates the neural processing of linguistic information at the lexical. For example, at 6 months of age, infants are becoming familiar with word sounds, but not word meaning. Consequently, Zangl and Mills (2007) revealed that familiar, but not unfamiliar, words spoken in IDS elicited brain responses typically associated with increased attentional processing (N200-400) in 6-month old infants. However, a separate group of 13-month old infants showed different patterns of brain activation. Specifically, familiar words not only elicited brain responses typically associated with increased attentional processing (N200-400), but also brain responses typically associated with the ability to map words to meaning (N600-900). In contrast, unfamiliar words elicited only an increase in attentional processing in the 13-month old infants. Altogether, this developmental trend of brain responses supports the idea that the brain is sensitive to the rich acoustic properties of IDS, such that increased attentional processes benefit early word learning.

2. The mismatch response (MMR) in infants as an assessment of neural commitment

The adult ERP component known as the Mismatch Response (MMR; Näätänen, 1992) reflects the brain's ability to detect an infrequent sound (deviant) among a sequence of repetitive sounds (standard). Thus, the MMR can be used to indicate early stages of auditory change detection in speech stimuli in both adults and infants. However, research has shown that the infant MMR can include two responses: a mismatch response (MMR) with positive polarity occurring at a shorter latency (positive-MMR at 150–350 ms after stimulus onset), and a MMR with negative polarity occurring at a later latency (negative-MMR at 350–550 ms after stimulus onset). The later, negative-MMR has been associated with the brain's ability to process the acoustic distinctions within one's native phonology (Cheour et al., 1998a,b), and thus can be used as a neural marker of language commitment (Kuhl et al., 2008; Rivera-Gaxiola, Klarman et al., 2005). Simply, as infants gain experience to the native language, the brain learns to process it more efficiently. Thus, stimuli that are representative of one's native-language will show substantially different brain activation compared to stimuli that are not representative of their native-language (Kuhl, 2000, 2004). In simple words, as infants gain experience to the native language, the brain shows neural specialization to native-speech sounds.

The amplitude of the negative-MMR evoked in the infant's brain has been used as index of neuronal specialization to native speech sounds. For example, 12-month old infants have shown a larger negative-MMR in response to a native phonetic contrast compared to a non-native phonetic contrast (Cheour et al., 1998a,b; Näätänen et al., 1997). This aligns with the wealth of evidence that illustrates how infants lose the ability to distinguish non-native speech sounds over time as a result of exposure to their native language (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Werker & Tees, 1983).

On the other hand, the earlier, positive-MMR is claimed to reflect high attentional demands in detecting an acoustic change. Friederici, Friedrich, and Christophe (2007) proposed that a positive-MMR represents the "extra effort" recruited as a consequence of processing acoustic deviances. In line with this hypothesis, the positive-MMR has been shown to decline with age as the brain matures and accumulates exposure to their native language(s) (García-Sierra, Ramírez-Esparza, & Kuhl, 2016; Morr, Shafer, Kreuzer, & Kurtzberg, 2002; Trainor et al., 2003), and accordingly has been shown to be predictive of infants' later language outcomes. For example, Friedrich et al. (2009) found that monolingual 4-month-old infants showed a positive-MMR in response to unfamiliar (foreign) stress patterns across familiar sounds (consonants and vowels). In addition, it was shown that atypical positive-MMR responses (larger and prolonged) to the non-native stress pattern predicted poorer language abilities at 2.5 years of age.

Hence, the literature reviewed so far suggests that neural specialization in the infant brain is observed as a reduction in attentional demands (i.e., positive-MMR decreases, while negative-MMR increases over time to native distinctions; positive-MMR remains for non-native distinctions). This is in accordance with Zhang, Kuhl, Imada, Kotani, and Tohkura (2005), which showed that the adult brain is more efficient (i.e., focal activity, shorter duration) when processing native compared to non-native speech contrasts.

Further research has confirmed that the positive-MMR is related to the attentional demands associated with the discrimination of a signal. For example, Cheng et al. (2015) created a clever experiment to tease apart the interaction between brain maturation, language experience, and the ease of discrimination of the signals. The researchers recorded MMRs in response to both easy, and hard to discriminate Chinese speech contrasts (long vs. short Mandarin vowels, respectively) in 3 groups of Chinese monolinguals: newborns, 6-month-old infants, and adults. The newborns showed positive-MMRs for the easy and hard speech contrasts. The 6-month-old infants showed positive-MMRs for the difficult contrast, but negative-MMRs for the easy contrast. Importantly, the adults showed negative-MMRs for both contrasts. The fact that positive- and negative-MMRs coexisted in the 6-month-old infants, and that the positive-MMR is absent in adults, suggests that ease of discriminability interacts with brain maturation and language experience. In other words, as infants gain language experience, infants become more familiar with the sounds of their native language such that processing demands require less attentional resources or occurs in a more automatic way (see Liu, Chen, & Tsao, 2014 for similar finding).

The idea that infants use enhanced attention to process native speech sounds, and that these demands are reduced with increasing language experience and brain maturation, is not new (Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993, Jusczyk et al., 1994). In line with Jusczyk and colleagues' studies, Strange (2011) proposed that infants develop attentional mechanisms, named selective perceptual routines (SPRs), to process speech sounds efficiently and automatically. Hence, the processing of non-native or infrequent speech sounds relies on weak SPRs and requires more attentional demands compared to native or frequent speech sounds with strong SPRs. In line with Kuhl (2004) neural specialization idea, perceptual routines for native speech sounds that have not been fully learned require increased attention and are manifested in the form of positive-MMRs, while fully learned SPRs to native speech sounds require low attentional demands and result in explicit cortical representation (i.e., negative-MMR). Further, research has corroborated that the quantity of language exposure infants receive in the household is associated with the brain response a given speech contrast elicits in infants (García-Sierra et al., 2016 see below for more details).

The Positive-MMR and Negative-MMR as a function of language experience. The interaction between the positive-MMR and negative-MMR is complex, but overall there is agreement that both can be used to assess language specialization. For example, in a longitudinal study with English monolingual infants, Rivera-Gaxiola, Silva-Pereyra et al. (2005) reported different patterns of brain activation at 7- and 11-months of age. The authors divided the sample in two groups based in the type of ERP response observed. Group 1 showed positive-MMRs to both native and non-native speech contrasts at 7-months-old, but showed negative-MMRs to the native contrast while maintaining a positive-MMR for the non-native speech contrast at 11-months-old. Group 2 showed negative-MMRs to both the native and non-native speech contrasts, at both ages. Rivera-Gaxiola and colleagues' proposed that the positive-MMR reflects a response associated to the acoustic properties of the signal, while the negative-MMR reflects a more mature response. Hence, the Group 1′s transition from positive- to negative-MMR observed only to the native contrast is thought to index infants' developmental trend towards neural language commitment. In contrast, Group 2′s consistent negative-MMRs across all ages and contrasts suggest that the infants were processing non-native phonetic distinctions like they would a native one. Hence, Group 2 index less language specialization. Importantly, further research with English monolingual infants has shown that the amplitude and duration of the positive-MMRs to a non-native contrast early in development can predict later language abilities (Rivera-Gaxiola, Klarman et al., 2005; Friedrich, Herold, & Friederici, 2009).

In a more resent study, García-Sierra et al. (2016) reported positive- and negative-MMRs in a group of 11–14 month-old bilingual (Spanish / English) and monolingual (English) infants. Differently from Rivera-Gaxiola, Silva-Pereyra et al.'s (2005) study, the infants were divided in two groups based on the amount of language input received in the household: high and low. Further, a native speech contrast for both groups (English), a non-native speech contrast for both groups (Chinese), and a speech contrast native to bilinguals, but not monolinguals (Spanish) were presented. The results showed that monolinguals with high amounts of language input displayed a negative-MMR for the English contrast, a positive-MMR for the Chinese contrast, and no positive- nor negative-MMR for the Spanish contrast. Even though, monolinguals with low amounts of language input similarly showed a positive-MMR for the Chinese contrast and no positive- nor negative-MMR for the Spanish contrast; they instead showed a positive-MMR to the English contrast. Consequently, and in accordance with Rivera-Gaxiola et al. only the brain responses to the native contrast showed a change in the polarity of the MMR, but the change was as a function of the amount of language input received in the household. Further, it can be inferred that the positive-MMR to native and non-native contrasts represent different brain mechanisms. Bilinguals, too, showed relationships between brain responses and amount of language input. Specifically, bilinguals with high amounts of English input showed positive-MMRs to the English contrast, and bilinguals with high amounts of Spanish input showed positive-MMRs to the Spanish contrast, while neither group showed no positive- nor negative-MMRs to the other two relevant contrasts.

García-Sierra's et al. (2016) results agree with Strange (2011) idea and expands on Rivera-Gaxiola, Silva-Pereyra et al.'s (2005) results in the sense that the positive-MMR can reflect more than one process. Namely, the positive-MMR to native speech sounds represents the development of attentional mechanisms associated with the formation of selective perceptual routines to facilitate speech processing (see also Gomes et al., 2000; Cheng et al., 2015). Specifically, it seems as though the quantity of language input is important to facilitate neural commitment in both monolingual and bilingual infants (García-Sierra et al., 2016). On the other hand, the positive-MMR to non-native speech perception represents attentional mechanisms associated with discriminating unfamiliar sounds and/or complex acoustic sounds aside from native language learning (Hisagi, Shafer, Strange, & Sussman, 2010; Lee et al., 2012; Marklund, Schwarz, & Lacerda, 2019).

Summary. Overall, the positive-MMR has been thought to represent a precursor to language commitment given its association with language experience, and thus can be used to understand beneficial factors in language learning across diverse populations. Support for this idea comes from studies showing that infants' speech discrimination of native contrasts transition from a positive- to negative-MMR as a function of maturation and/or language experience (García-Sierra et al., 2011; Kuhl et al., 2008; Rivera-Gaxiola, Klarman et al., 2005). Relevantly, the positive-MMR can reflect more than one process since it is not present for all native speech contrasts (hard- present; easy- absent; Cheng et al., 2015), but also only shows a relation with the amount of language input received in the household to native speech contrasts (García-Sierra et al., 2016).

Positive-MMR and IDS. To our knowledge there is one study that has explored the positive-MMR as a function of IDS. Peter et al. (2016) exposed adults and 9-month-old infants to an exemplar of the native vowel /i/ in two conditions: one used the vowel in IDS as the repetitive, standard sound but in ADS as the infrequent deviant sound, and the other used ADS as the standard but in IDS as the deviant. This design allowed the difference waveforms (deviant vs. standard) to best reflect phonetic and lexical processing of the vowel, independent from its general auditory properties. Adults' results showed that IDS and ADS produced similar MMRs. On the other hand, infants' results showed that IDS produced an adult-like MMR (i.e., negative), while ADS produced a positive-MMR. Thus, the researchers suggested that the infants had more difficulty discriminating speech contrasts in ADS. In other words, IDS facilitates infants' neural processing of speech at the phonetic level, similar to how high language exposure facilitates infants' language commitment (García-Sierra et al., 2016).

In the present investigation, we focus on the positive-MMR to further understand how the quality of speech impacts higher attentional processes related to language learning. Specifically, we will correlate the amount of IDS Spanish-English bilingual infants receive with their positive- MMRs observed to native (Spanish) and non-native (Chinese) speech contrasts. This will help better understand the degree to which IDS fuels language learning and commitment in diverse language learners.

2.1. *Infant directed speech and language learning*

Different methodologies have been used to understand how IDS promotes language learning, which has led to some questioning of how infants benefit from IDS. For example, Ramírez-Esparza, García-Sierra, & Kuhl (2017a) used a qualitative approach to quantifying Spanish-English bilingual infants' exposure to IDS, that then could be associated with infants' later language abilities. Specifically, infants wore a LENA digital recorder (LENA foundation, Boulder, Colorado) over four days. Coders later identified the presence of IDS during caregiver-infant social interactions. The researchers found a positive association between IDS exposure at 1 year of age and vocabulary size at 2 years of age assessed via parental report (Ramírez-Esparza et al., 2017a). These findings were comparable to their monolingual peers (Ramírez-Esparza , García-Sierra, & Kuhl, 2014).

Although Ramírez-Esparza et al. (2014); (2017a;) demonstrated that IDS promotes language learning, their studies qualitatively assessed IDS, rather than quantifying its main phonetic characteristic: "hyperarticulation." Hyperarticulation refers to the expansion and elongation of vowels, which is typically measured by extracting the Vowel Space Area (VSA; e.g., Kuhl et al., 1997; Liu & Kuhl, 2003; Song, Demuth, & Morgan, 2010). Explicitly, the VSA plots the distance between the average F1 and F2 formant frequencies of each point vowel: /i/, /a/, and /u/. As a result, hyperarticulation creates larger distances between the formant frequencies of each vowel category, and thus an expanded vowel space (Kuhl et al., 1997). Arguably, this hyperarticulation makes the exemplar clearer, and promotes language learning. For example, Taiwanese infants exposed to a larger vowel space better discriminate speech in head turn procedures (Liu & Kuhl, 2003). Similarly, Song et al. (2010) found that slow speaking rate and hyperarticulation were associated with an increase in infants' word recognition ability. In a more recent study, Hartman, Ratner, and Newman (2017) found that hyperarticulation used by caregivers during a free play activity in the lab with their 18-month old children was associated with expressive and receptive child language outcomes at 2 years of age.

While the above laboratory tasks are telling, there is ongoing debate about whether infants benefit from IDS as a result of its hyperarticulation in the vowels, by-products from the slower rate of speech, different prosodic structure (McMurray, Kovack-Lesh, Goodwin, & McEchron, 2013), or the breathiness in the vowels that are associated with positive affect (Miyazawa, Shinya, Martin, Kikuchi, & Mazuka, 2017). Similarly, it has been questioned whether IDS produces significantly higher quality vowel tokens as compared to ADS. Martin et al. (2015) demonstrated with a computational algorithm that ADS in Japanese (assessed from spontaneous speech between caregivers and experimenters in the lab) was discriminated better than the corresponding IDS contrasts (assessed from spontaneous speech between caregivers and child at home), suggesting that ADS produced clearer vowels. This result could be associated with the inherent characteristics of the Japanese language, or the methodological approach used in the study (i.e., analyses of spontaneous speech). For example, the Japanese language differs from the English language in their vowel inventory (i.e., 5 in Japanese vs. 11–12 in English), as well as on the fact that Japanese vowels are not as heavily diphtongized as English vowels (Miyazawa et al., 2017). Furthermore, the study analyzed spontaneous speech, which contains more variable hyperarticulation than controlled speech. On this note, Miyazawa et al. (2017) showed that although both Japanese mothers' spontaneous IDS samples and carefully read speech (CS) samples in the lab had expanded vowel spaces compared to their spontaneous ADS samples, their CS speech had a larger space than IDS. This suggested that IDS in spontaneous speech does not always produce clear tokens as the hyperarticulation hypothesis suggests.

In this study, we introduced several methodological approaches to address the mixed findings on the effects of IDS in language development. First, we extracted the VSA from bilingual caregivers while using their dominant language, Spanish. This will further expand the literature by including another language that is comparable to Japanese (Miyazawa et al., 2017) in the sense that it has 5 vowels and is mostly monophthongal (Swingley & Alarcón, 2018). Second, we used an approach that compromised between spontaneous IDS and carefully read speech samples (e.g., Miyazawa et al., 2017). Specifically, instead of analyzing IDS during a free play or reading activity (e.g., Hartman et al., 2017; Liu & Kuhl, 2003; Martin et al., 2015; Miyazawa et al., 2017), we used a storybook task that has been used in other studies for acoustic analyses (Burnham et al., 2015; Inoue, Nakagawa, Kondou, Koga, & Shinohara, 2011; McMurray et al., 2013). However, unlike those studies that asked caregivers to read storybooks in the laboratory settings, we asked caregivers to do it at home in a moment and time of their choice. Further, caregivers were not instructed to use IDS in order to capture the natural variability of hyperarticulation used by the caregivers when reading to their infants (Ramírez-Esparza et al., 2014; 2017a; 2017b). Essentially, our approach aimed to observe the variability of the hyperarticulation across caregivers (as we expect some to do it more than others), but yet control for the variability of tokens by asking all caregivers to do the same storybook reading activity. Importantly, in order to assess if hyperarticulation is used at home, we did acoustic analyses on caregivers' adult speech produced from tokens in the lab. Finally, to complement the literature, we included an outcome variable to assess early language development

3. Study overview

The general goal of this investigation was to observe if the quality of IDS spoken by caregivers is associated with early language commitment in infants. In order to accomplish this goal, Spanish-English bilingual families with infants of 11 and 14 months of age were recruited from a larger study. As part of that study, infants wore the LENA digital language processors (DLPs; LENA foundation, Boulder, Colorado) in order to capture everyday social interactions between caregivers and their infants (Ramírez-Esparza et al., 2017a), as well as the association between amount of language exposure and brain measures (García-Sierra et al., 2016). The LENA DLPs are ideal for recording natural data at home, continuously and unobtrusively (see Mehl & Holleran, 2007; Ramírez-Esparza et al., 2019). For example, parents read to the child in a time and moment of their choice, as opposed to during a task set by the experimenter while being audio-recorded. Furthermore, the LENA DLPs collect high quality sound files that can be later processed with specialized speech algorithms to detect adult word count among other speech categories (e.g., Oller et al., 2010). Thus, presumably, these sounds files can be considered optimal for VSA analyses. For the purpose of this study, the caregivers were instructed to read a storybook with a set of sentences in Spanish, once a day for four days, to their infants. Caregivers also read the same set of sentences in the lab, but now were explicitly instructed to do so as if they were reading to an adult. Caregivers also read sentences in English; however, we do not report VSA for the English language given that the bilingual sample was mostly Spanish dominant, and thus the English sentences were confounded with Spanish-accented speech (see Fish, García-Sierra, Ramírez-Esparza, & Kuhl, 2017).

Early language commitment was assessed by recording infants' ERPs in response to two sets of speech sounds (native or nonnative) which were contrastive in only one language of interest (Spanish or Chinese). Specifically, the native contrast was a Spanish voicing stop consonant contrast (/da/ vs. /ta/; see García-Sierra et al., 2016), while the non-native contrast was a Chinese-Mandarin consonant contrast (alveolo-palatal affricate /tɕhi/ vs. alveolo-palatal fricative /ɕi/; see Tsao, Liu, & Kuhl, 2006). Both speech contrasts were created with acoustic properties representative of ADS. Therefore, our first goal was to extract the positive-MMR for each of the speech contrasts independently. Based on previous research with bilinguals (García-Sierra et al., 2016), we assumed that infants would show positive-MMRs to both speech contrasts, suggesting high attentional demands during speech processing (Cheng et al., 2013; 2015; Lee et al., 2012; Marklund et al., 2019; Peter et al., 2016). Then, we compared infants' brain responses elicited across the native and non-native contrasts to assess if infants showed enhanced neural sensitivity to the native speech contrast.

The second goal was to test if the positive-MMR is also associated with language commitment. In order to accomplish this goal, we observed the relation between VSA and the positive-MMRs associated with the Spanish contrast and the Chinese contrast independently. Although no studies have tested this idea before, we speculated VSA to be associated with early language learning as exemplified by the positive-MMR observed for the native language (Spanish). We hypothesized this given that IDS has been associated with language learning as measured by head-turn procedures (Liu & Kuhl, 2003), word recognition (Song et al., 2010), and parental reports (e.g., Hartman et al., 2017; Ramírez-Esparza et al., 2014; 2017a). In line with the idea that language exposure influences subsequent language learning, we speculated that VSA would not be associated with the positive-MMR observed for the non-native contrast (Chinese contrast).

Finally, we carried out a linear mixed model by treating the brain measures as the bivariate response, and VSA, along with SES, infants' age, and fundamental frequency (F0) as predictors. VSA was the main predictor of interest; however, we wanted to control for other predictors associated with language to isolate its true role in our study. SES was considered as a predictor given the evidence that demonstrates how language development, as well as caregivers' quality of speech (i.e., IDS), are associated with SES (Hoff, 2006; Gilkerson et al., 2017; Ramírez-Esparza et al., 2014; 2017a; 2017b; Rowe, 2008; Rowe & Goldin-Meadow, 2009). Age was considered as a predictor in the model given that our sample included infants belonging to two age groups and language commitment has been associated with age (e.g., Morr et al., 2002; Trainor et al., 2003). Finally, F0 was considered as a predictor given that F0 facilitates infants' neural processing of linguistic information (Bosseler et al., 2016; Saito et al., 2007; Zhang et al., 2011; Zangl & Mills, 2007; Peter et al., 2016). To summarize, we tested the hypothesis that caregivers' VSA would predict the amplitude of the positive-MMR to the native speech contrast, but not to the non-native speech contrast, even after holding the other predictors constant. This finding would suggest that caregivers' hyperarticulated speech fosters (mediates) the attentional demands that facilitate native speech processing during the first years of life.

4. Methods

4.1. Participants

A total of seventeen infants and their caregivers were recruited as part of a larger study, which aimed to understand Spanish-English bilingual infants' language development in terms of brain measures (García-Sierra et al., 2016) and social measures (Ramírez-Esparza et al., 2017a). Therefore, inclusionary criteria dictated that both English and Spanish were spoken in the home. Infants were 11 months old (N = 11; 5 female) or 14 months old (N = 6; 3 female) at the time of data collection. Infants were full-term (37–43 weeks), of normal birth weight (6–10 lbs), and encountered no major birth or postnatal complications, recurrent ear infections, or any known hearing impairments. An additional two infants failed to complete testing due to noisy EEG and failure to follow instructions. Participants were paid 100 dollars for their participation. The University of Washington Institutional Review Board approved this project.

Bilingual caregivers' socioeconomic status. Socioeconomic status (SES) was assessed using four domains: marital status, retired/employed status, educational attainment, and occupational prestige (Hollingshead, 2011). The higher caregivers' Hollingshead score is, the higher their SES (range from 8 to 66). Caregivers' mean score was 45.41 (*SD* = 17.67), which suggests middle SES.

4.2. Bilingual caregivers' cultural and language background

Caregivers' mean age was 28.82 (*SD* = 8.8). On average, caregivers reported to have lived in the U.S. for 13.40 years (*SD* = 9.50). Six bilingual caregivers were born in Mexico, five were born in the U.S., three in Colombia, one in Venezuela, one in Peru, and one in Puerto Rico. Bilingual caregivers completed a background questionnaire to assess their exposure related to English and Spanish (García-Sierra et al., 2016). The overall ***confidence*** in understanding and speaking English and Spanish at the time of the experiment was assessed using a Likert scale which asked bilingual caregivers to rate themselves on a 1–5 scale (1 = "I cannot speak the language, I have a few words or phrases, and I cannot produce sentences"; 5 = "I have a native-like proficiency with few grammatical errors and I have good vocabulary"). The overall mean for bilingual caregivers' confidence in speaking English was 4.5 (*SD* = 0.64) and 5.0 (*SD* = 0.0) for Spanish. The overall mean for bilingual caregivers' confidence in understanding English was 4.78 (*SD* = 0.42) and 5.0 (*SD* = 0.0) for Spanish. Furthermore, Fig. 1 depicts that bilingual caregivers were dominant in Spanish in early life, but transitioned to being more balanced to both languages at the time of the experiment.

4.3. Procedure

Participants were recruited at the Institute for Learning and Brain Sciences in Seattle, Washington. The families participated in the study's two parts: (1) assessment of caregivers' speech at home (IDS) and in the lab (adult speech), and an (2) assessment of infants' brain measurements (ERPs) in the lab.

4.4. Part 1: Speech recordings

Infant Directed Speech (IDS). The LENA digital recorders (DLPs) were used to capture IDS during the reading sentence task. Caregivers received two LENA digital recorders (DLPs) and vests designed with a chest pocket to hold the DLP for infants to wear. The DLP can record up to 16 h. Accordingly, caregivers were instructed to record 8 h a day, for 4 days (2 weekdays and 2 weekend days; analyses to portions of this naturalistic data collection can be found in García-Sierra et al., 2016; Ramírez-Esparza et al., 2017a; 2017b).

Caregivers read the storybook to their infants once a day for four days while wearing the DLP, and indicated the time at which they did so in a provided time-log sheet. The storybook contained 12 pages where each page displayed a picture associated with a sentence (e.g., *La pala es mi juguete preferido* / The shovel is my favorite toy). Each sentence had a target word that was bisyllabic and constructed as CVCV or CVV (e.g., *pala*, *día*, *buzo* / shovel, day, diver; respectively). In all of the target words, the first vowel was a point vowel (/i/, /a/, and /u/) which allowed us to construct the vowel space area (VSA) within caregivers' speech. We used the time-log sheet to identify the target words containing these point vowels of interest within the recording. The storybook was read multiple times to increase the sample size of productions, and to become familiarized with the acoustic properties of each participants' productions. The total number of planned vowel productions in IDS per participant was /i/ = 8, /a/ = 32, and /u/ = 8. This familiarization helped to exclude as much coarticulation as possible when placing the vowel boundaries used to quantify the hyperarticulation of IDS (see acoustic analyses below).

Adult speech. Caregivers carried out the same sentence reading task done at home, now in the lab, to assess adult speech. Specifically, caregivers were seated in a sound-attenuated recording booth and instructed to read the experimental sentences at a normal speed and volume, as though they were reading to another adult. Each caregiver read the set of experimental sentences 3 times, resulting in the total number of planned vowel productions in adult speech per participant as /i/ = 6, /a/ = 24, and /u/ = 6. Recordings were done in laboratory to control for noise, and thus create the intended number of acceptable recordings in less tries.

Acoustic Analyses. Each target word from both the task done at home, and the task done in the lab, were analyzed for the acoustic boundary dividing the first vowel from the remainder of the target word. The spectrograms and waveform were displayed simultaneously in PRAAT (Boersma & Weenink, 2018). To isolate the vowel, boundaries were manually inserted at the onset and offset of the vowel, excluding as much coarticulation as possible. Vowel length was calculated (offset boundary minus onset boundary) using PRAAT scripts that took formant

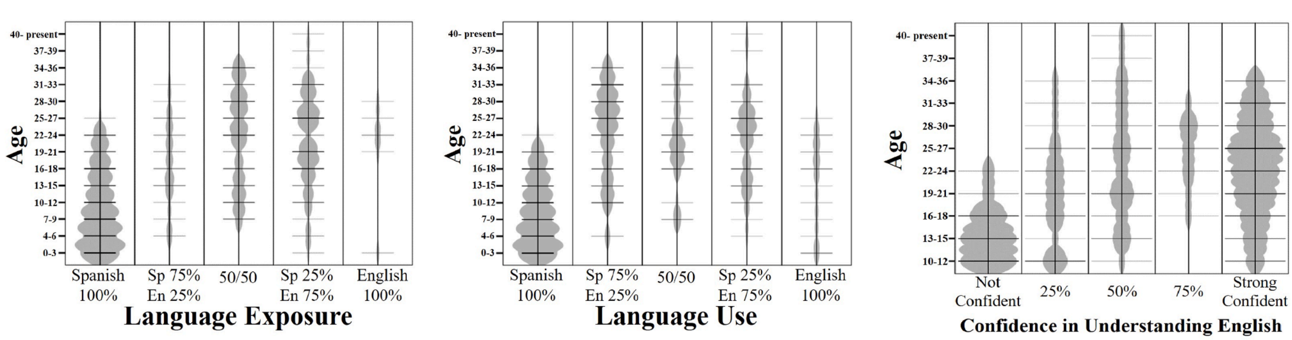

Fig. 1. Bilingual caregivers' violin plots of their self-reports for the questionnaire assessing amount of language exposure and language use in English and Spanish as a function of age.

values at 20%, 50%, and 80% time intervals. Since it has been shown that initial (20%), central (50%), and final (80%) vowel time intervals produce identical VSAs (Kuhl et al., 1997), we averaged our three measures to create a single measure associated with each formant (F1 and F2). Each formant measurement was visually inspected before being accepted in the final sample. The total number of vowels accepted in IDS from the reading task at home were /i/ = 97, /a/ = 373, and /u/ = 94, and for the reading task in the lab were /i/ = 89 /a/ = 354, and /u/ = 88. To compensate for the unequal number of productions across the three vowels and noisy data, an average vowel score was calculated for each respective vowel. Hence, each caregiver had one triad of vowels (i.e., average /i/, average, /a/, and average /u/) that was used to calculate the VSA for both the reading tasks.

Vowel Space Area: VSAs from this triad of point vowels were calculated like previous studies (e.g., Blomgren, Robb, & Chen, 1998; Sapir, Połczyńska, & Tobin, 2009). Namely, we used the first (F1) and second formants (F2) of the vowels /i/, /a/, /u/, to calculate the Euclidean distances (EDs) between each vowel (/i-u/, /i-a/, and /a-u/). The formant frequencies of these metrics were logarithmically scaled to reduce irrelevant interspeaker variability. These three Euclidian distances (EDiu, EDia, and EDau, respectively) define the VSA. The mathematical formula from which the VSA values were derived was:

$$\mathrm{VSA} = \mathrm{sqrt}(S^*(S - \mathrm{EDiu})(S - \mathrm{EDia})(S - \mathrm{EDau}))$$

where

$$S = (\mathrm{EDiu} + \mathrm{EDia} + \mathrm{EDau})/2$$

$$\mathrm{EDiu} = \mathrm{sqrt}((F1i - F1u)^2 + (F2i - F2u)^2)$$

$$\mathrm{EDia} = \mathrm{sqrt}((F1i - F1a)^2 + (F2i - F2a)^2)$$

$$\mathrm{EDau} = \mathrm{sqrt}((F1a - F1u)^2 + (F2a - F2u)^2)$$

4.5. Part 2: Electrophysiological recordings

Event Related Potentials Test. Infants were awake and tested inside a sound treated room. The child sat on the caregiver's lap. In front of them, a research assistant entertained the child with quiet toys while a muted movie played on a TV behind the assistant. The caregiver and the research assistant wore headphones with masking music during the testing phase. Two recording sessions were done on different days; one for the Spanish speech contrast and one for the Chinese speech contrast (counterbalanced). Each recording session lasted approximately 20 min. The electroencephalogram (EEG) was recorded using electro-caps (ECI, Inc.) incorporating 32 pre-inserted tin inverting electrodes. The EEG was referenced to the left mastoid from Fp1, Fpz, Fp2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T3, C3, Cz, C4, T4, CP5, CP1, CP2, CP6, T5, P3, Pz, P4, T6, O1 Oz, O2, and RM in the International 10/20 System. Infant eye-blinks were monitored by recording the electrooculogram from 1 infraorbital electrode placed on the infant's left cheek. The EEG data was collected in DC mode and it was re-referenced off-line to the right mastoid to obtain a more balanced reference distribution. The electroencephalogram was recorded using NeuroScan SynAmps RT amplifiers (24 bit A/D converter) using Scan 4.5 software. A 1 ms trigger was time-locked to the presentation of each stimulus to accomplish the ERP averaging process (Stim 2 Neuroscan Compumedics).

The impedances of all electrodes were kept below 5 kΩ. EEG segments with electrical activity +/−150 mV at any electrode site were omitted from the final average. EEG segments of 700 ms with a pre-stimulus baseline time of 100 ms were selected and averaged offline to obtain the ERPs. Baseline correction was performed in relationship to the pre-stimulus time. The ERP waveforms were band-pass filtered off-line (1–40 Hz with 12 dB roll off) using the zero phase shift mode function in NeuroScan Edit 4.5. The high- and low-cutoff filters used are reported elsewhere and do not produce attenuation of the ERP waveforms (see: Sabri & Campbell, 2002).

Design. Two different synthetic speech contrasts used by García-Sierra et al. (2016), one native (i.e., a Spanish speech contrast) and one non-native (i.e., a Chinese speech contrast), were tested using the classic odd-ball paradigm. Within a given speech contrast, one "standard" syllable is repeated for 85% of the trials (850 stimulus repetitions), and a different "deviant" sound is presented pseudo-randomly during the remaining 15% of the trials (150 stimulus repetitions). Here, the deviant sound did not occur successively, and at least three standard sounds were presented between deviant sounds. The time between the offset of a stimulus and the onset of the next stimulus (inter stimulus interval) was 705 ms.

At the end of our classic odd-ball paradigm, the deviant sound was presented 200 times to use as a control-deviant in further analyses. Specifically, the deviant sound presented within the paradigm (150 stimulus repetitions) was compared to this control-deviant, as opposed to the standard sound. Hence, the MMR reported in the present investigation is the difference between the response to a stimulus presented as a deviant in one condition and the response to this same stimulus presented alone in a separate measurement. Since these sounds are acoustically identical, analyses of brain activity intended to target auditory change detection were not confounded by brain activity associated with any physical acoustic differences between the stimuli, unlike if the traditional deviant vs. standard was used (e.g., King, McGee, Rubel, Nicol, & Kraus, 1995; Kraus, McGee, Sharma, Carrell, & Nicol, 1992; Kraus et al., 1995; May & Tiitinen, 2010).

The number of trials accepted for the Spanish Control-Deviant was 178.06 (SD = 54.53) and 108.13 (SD = 13.22) were accepted for the Spanish Deviant sound. The number of trials accepted for the Chinese Control-Deviant was 168.00 (SD = 52.39) and 100.76 (SD = 17.32) were accepted for the Chinese Deviant sound.

Event Related Potentials Statistical Analyses: ERP amplitude analyses and correlational analyses were done with BESA Statistics 2 (BESA GmbH, Gräfelfing, Germany) that uses data clustering in combination with permutation testing. This process is a data-driven approach that assumes that if a statistical effect is observed in an extended period of time in several neighboring channels, then it is unlikely that the effect occurred by chance. In the first step BESA, performs a parametric test to find data clusters that show pronounced effects. BESA calculates a cluster-value for each pronounced effect that represents the sum of the t-values in the time region(s) where p-values are below 0.05. Importantly, the cluster-value represents values in the time (ms) and spatial (electrode) domain. Therefore, a large cluster-value represents significant difference in the time domain across multiple neighboring electrodes, while small cluster values represent significant difference in one or few electrodes. In the present research, we used a channel neighbor distance of 4.5 cm. In the second step, BESA repeats step 1 but using a permutation test. This serves to test if the probabilities of the cluster-values across experimental conditions (or subjects) are exchangeable. Hence, for each of the calculated permutations (in our case 10,000), a new t-test is computed per data-point and new clusters are determined. Accordingly, each permutation will result in a new cluster-value derived for each cluster. Thus, a distribution of cluster-values can be established across all permutations and the α-error of the initial cluster-value in step 1 can be directly determined. In other words, it is determined if the initial cluster-value derived in step 1 is equally likely as any other cluster-values derived in each permutation step. This kind of analysis is performed to control for Type I error due to the large number of data points compared in ERP responses (see: Bullmore et al., 1999; Ernst, 2004; Maris & Oostenveld, 2007).

Importantly, since the analysis is done in both the time and space domains, the ERP time window in which a significant cluster is observed varies across channels. In other words, the significant time window observed in electrode "x" will be similar, but not the same, to the one observed in electrode "y".

Stimuli: Based on the fact that the acoustic properties of the 2 speech contrasts are very different, we analyzed each speech contrast independently.

Spanish Contrast. Here, the phonetic contrast among the deviant (/da/; prevoiced) and standard (/ta/; voiceless unaspirated) sounds represented two different stop consonants in Spanish, but not in Chinese or English. Namely, the 225 ms long syllables only differed in voice onset time (VOT), the primary acoustic cue for voicing distinction. The prevoiced syllable, /da/, had 50 ms of voicing before the vowel (-50 ms VOT + 175 ms vowel), while the voiceless unaspirated syllable, /ta/, had 0 ms of voicing before the vowel (0 ms VOT + 225 ms). Otherwise, both speech tokens were identical. Both syllables had a first formant (F1) of 500 Hz at the consonant release, and beginning F2, F3, and F4 values of approximately 1550, 2500, and 3800 Hz, respectively. Further, the steady state vowel formant frequencies for F1 to F4 were 800, 1280, 2200, and 3800 Hz, and bandwidths were 50, 60, 90, and 140 Hz, respectively. Pitch contours were identical during the vowel portion with a fundamental frequency of 110 Hz at the beginning of the vowel and tapering down to 95 Hz. Tokens were equalized in RMS amplitude and played to infants at a comfortable listening level of 67 dBA.

Chinese contrast. Here, the phonetic contrast among the deviant (/tɕhi/, alveolo-palatal affricate) and standard (/ɕi/, alveolo-palatal affricate) represented two different consonants in Mandarin Chinese, but not in Spanish or English. Namely, the 375 ms long syllables only differed in the point at which amplitude peaked in the initial 130 ms frication portion. The affricate consonant, /tɕhi/, had a fast amplitude rise reaching its maximum at 30 ms, while the fricative consonant, /ɕi/, had a slow amplitude rise reaching its maximum at 100 ms. Otherwise, both speech tokens were identical. The syllables had steady-state vowel formant frequencies of 293, 2274, 3186, and 3755 Hz; bandwidths of 80, 90, 150, and 350 Hz, respectively; and a fundamental frequency of 120 Hz (high-flat tone, Tone 1 in Mandarin). Tokens were equalized in RMS amplitude and played to infants at a comfortable listening level of 67 dBA.

5. Results

5.1. Do caregivers naturally use IDS at home?

We expected that some caregivers would naturally hyperarticulate vowels more than others while doing the reading task at home. In order to test this hypothesis, we assessed the VSA values from the tokens collected in natural environments at home (IDS-VSA) and the tokens collected in a controlled environment in the lab when reading as to another adult (adult speech-VSA). Table 1 shows the mean and standard deviations for F1 and F2 for formants on /i/, /u/, and /a/ for both IDS and adult speech from which VSA values were calculated. The IDS-VSA mean was 0.21 $LnHz^2$ (*SD* = *0*.04). The ED between /i-u/ was 0.85 LnHz (*SD* = *0*.13), between /i-a/ was 0.56 LnHz (*SD* = *0*.04), and between /a-u/ was 0.77 LnHz (*SD* = *0*.13). The adult speech-VSA mean was 0.17 $LnHz^2$ (*SD* = 0.06). The ED between /i-u/ was 0.70 LnHz (*SD* = 0.17), and between /i-a/ was 0.56 LnHz (*SD* = 0.14), between /a-u/ was 0.76 LnHz (*SD* = 0.21).

In order to observe if the caregivers' IDS-VSA scores significantly differed from their adult speech-VSA scores, the following steps were done: First, a median split on the IDS-VSA scores was done to define the high- and low-IDS-VSA groups. Then, for each participant, we paired their adult speech-VSA score. Finally, we carried out a repeated measures ANOVA examining speech style as a within factor (IDS vs. adult speech) and group as between factor (high-IDS-VSA vs. low-IDS-VSA). The results showed a trend towards significance for speech style $F(1, 15) = 3.60$, $p = .07$, $\eta p^2 = 0.43$), and a significant interaction between speech style and group $F(1,15) = 4.98$, $p = .041$, $\eta p^2 = 0.55$). Pair-wise comparisons with Bonferroni correction showed that the low-IDS-VSA group did not show a significant increase in vowel space area between IDS (Mean = 0.17 $LnHz^2$, SE = 0.01) and adult speech (Mean = 0.175 $LnHz^2$, *SE* = 0.02, p = .82; N = 8), while the high-VSA group showed a significant increase in vowel space area when producing IDS (Mean = 0.24 $LnHz^2$, *SE* = 0.01) than when producing adult speech (Mean = 0.18 $LnHz^2$, *SE* = 0.02, p = .009; N = 9). Fig. 2 shows the vowel triangles obtained from the low-IDS- and high-IDS-VSA groups in comparison with their vowel triangles obtained from adult speech.

Overall, these analyses indicate that there was a natural variability in the use of IDS at home indicating that in their natural environments some caregivers used IDS by hyperarticulating the vowels, while others used IDS speech that was equivalent to adult speech.

Table 1
Caregivers' means and standard deviations for F1 and F2 from the task done at home (IDS) and the reading task done in the lab (Adult Speech).

	/i/		/u/		/a/	
IDS	Mean	SD	Mean	SD	Mean	SD
F1 (Hz)	546.45	*43.98*	428.57	*39.37*	816.57	*54.53*
F2 (Hz)	2391.22	*161.81*	1102.85	*153.28*	1609.63	*125.35*
	/i/		/u/		/a/	
Adult Speech	Mean	SD	Mean	SD	Mean	SD
F1 (Hz)	509.33	*74.90*	389.23	*41.04*	832.25	*91.93*
F2 (Hz)	2290.01	*188.27*	1188.72	*107.37*	1605.23	*106.40*

5.2. Do both speech contrasts generate a positive-MMR?

The comparison between Control-Deviant ERP vs. Deviant ERPs was done independently for each speech contrast using BESA Statistics 2 (BESA GmbH, Gräfelfing, Germany). The analyses were done in the time interval between –100 and 700 ms in 29 scalp electrodes.

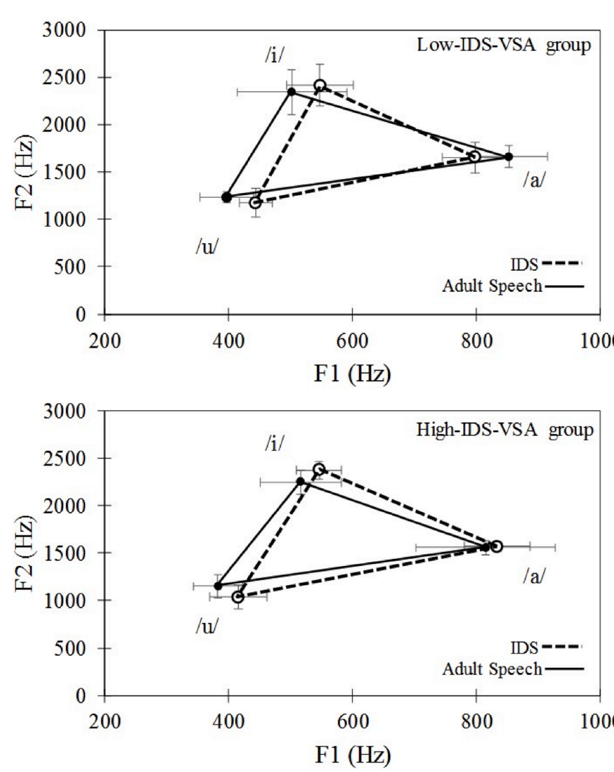

Fig. 2. Vowel triangles formed by the average of the F1 and F2 values for vowels /i/, /u/, and /a/. The upper panel shows the vowel space obtained from caregivers in the low-IDS-VSA group when producing IDS and adult speech. The lower panel shows the vowel space obtained from caregivers in the high-IDS-VSA group when producing IDS and adult speech. Error bars represent ± 1 Standard Deviation from the mean.

Positive-MMR analysis for the Spanish contrast. The analysis showed a cluster-value of 5342.41 between 166 and 348 ms. The cluster-value showed a different probability distribution between Deviant and Control-Deviant (p = .0003). Hence, the results showed a significant difference between Deviant (Mean = 4.16 μV, SD = 2.80) and Control-Deviant (Mean = 2.10 μV, SD = 2.63). Fig. 3 shows a more positive ERP response for the Deviant sound than the Control-Deviant sound in left and right frontal electrodes.

Positive-MMR analysis for the Chinese contrast. The analysis showed a cluster-value of 1128.63 between 180 and 289 ms. The cluster-value showed a different probability distribution between Deviant and Control-Deviant (p = .05). Hence, the results showed a significant difference between Deviant (Mean = 6.73 μV, SD = 2.63) and Control-Deviant (Mean = 4.98 μV, SD = 3.40). Fig. 3 shows a more positive ERP response for the Deviant sound than for the Control-Deviant sound in the right frontal electrodes.

Spanish positive-MMR vs. Chinese positive-MMR. In order to learn if the infants showed an enhanced neural sensitivity to the native speech contrast compared to the non-native contrast, we compared the positive-MMR to the Spanish contrast to the positive-MMR to the Chinese contrast. Using BESA statistics, we defined the time interval of interest from 150 to 350 ms (positive-MMR) in 29 scalp electrodes. The analysis showed a cluster-value of 1377.54 between 156 and 327 ms. The cluster-value showed a different probability distribution between both conditions (p = .04). Hence, the results showed a significant difference between the positive-MMR to the Spanish contrast (Mean = 2.106 μV, SD = 1.51) and the positive-MMR to the Chinese contrast (Mean = -0.600 μV, SD = 2.36). Fig. 4 depicts a more positive response to the Spanish contrast than to the Chinese contrast in left central electrodes (electrodes: T7, C3, CP5, CP1, P7, P3, and Pz).

5.3. Do caregivers' IDS correlate with infants' positive-MMRs?

We examined the correlation between caregivers' VSAs and infants' positive-MMRs.

The correlation was done independently for each speech contrast using BESA Statistics 2 (BESA GmbH, Gräfelfing, Germany). The analyses were done in the time interval between 150 and 350 ms in 29 scalp electrodes.

The analysis for the Spanish contrast showed a cluster-value of 2278.41 between 150 and 270 ms. The cluster-value showed a different probability distribution; hence, the cluster-value is not exchangeable across participants (p = .008). This means that there was a significant positive correlation between caregivers' VSA and infants' positive-MMR to the Spanish contrast (Mean = 1.07 μV, SD = 2.51). Fig. 5 depicts that caregivers' VSA correlated with left central electrodes.

The analysis for the Chinese contrast showed a cluster-value of 190.24 between 287 and 350 ms. The cluster-value showed a similar probability distribution; hence the cluster-value is exchangeable across participants (p = .22). This means that there was no significant correlation between caregivers' VSA and infants' positive-MMR to the Chinese contrast (Mean = −1.54 μV, SD = 2.65).

5.4. Do caregivers' adult speech correlate with infants' positive-MMRs?

We examined the correlation between caregivers' adult speech and infants' positive-MMRs. The correlation was done independently for each speech contrast using BESA Statistics 2 (BESA GmbH, Gräfelfing, Germany). The analyses were done in the time interval between 150 and 350 ms in 29 scalp electrodes.

The analysis for the Spanish contrast showed a cluster-value of −260.58 between 175 and 236 ms. The cluster-value showed a similar probability distribution; hence the cluster-value is exchangeable across participants (p = .21). This means that there was no significant correlation between caregivers' VSA adult speech and infants' positive-MMR to the Spanish contrast (Mean = 1.74 μV, SD = 2.73).

The analysis for the Chinese contrast showed a cluster-value of −30.78 between 338 and 350 ms. The cluster-value showed a similar probability distribution; hence the cluster-value is exchangeable across participants (p = .56). This means that there was no significant correlation between caregivers' VSA in adult speech and infants' positive-MMR to the Chinese contrast (Mean = −1.66 μV, SD = 3.31).

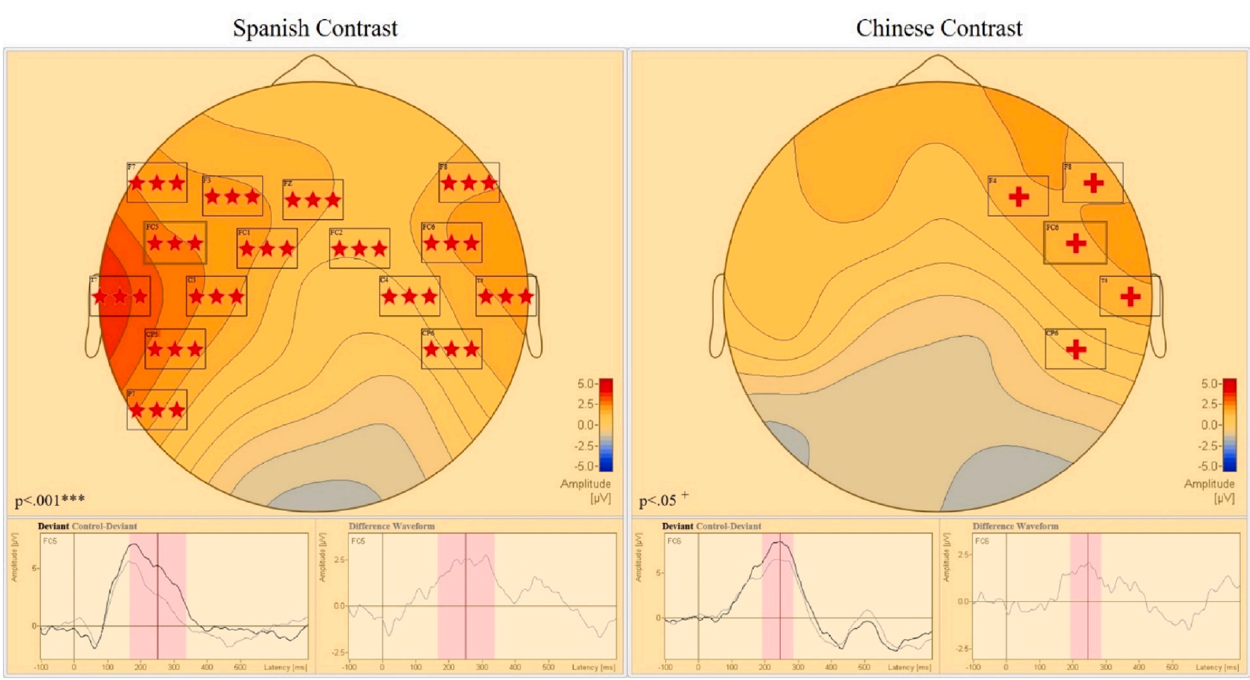

Fig. 3. Voltage maps and Event Related Potentials obtained for the Deviant and Control-Deviant sounds. In each panel, the top voltage map depicts the electrodes that showed a significant difference between both ERPs. Voltage maps are time-locked to 250 ms after stimulus onset The left-bottom box shows the amplitude (in microvolts) for both ERP responses for electrode FC5. The right-bottom box shows the difference waveform (positive-MMR). The highlighted areas depict the significant differences between Control-Deviant and Deviant for electrode FC6 using a point-by-point comparison for the Spanish contrast and Chinese contrast.

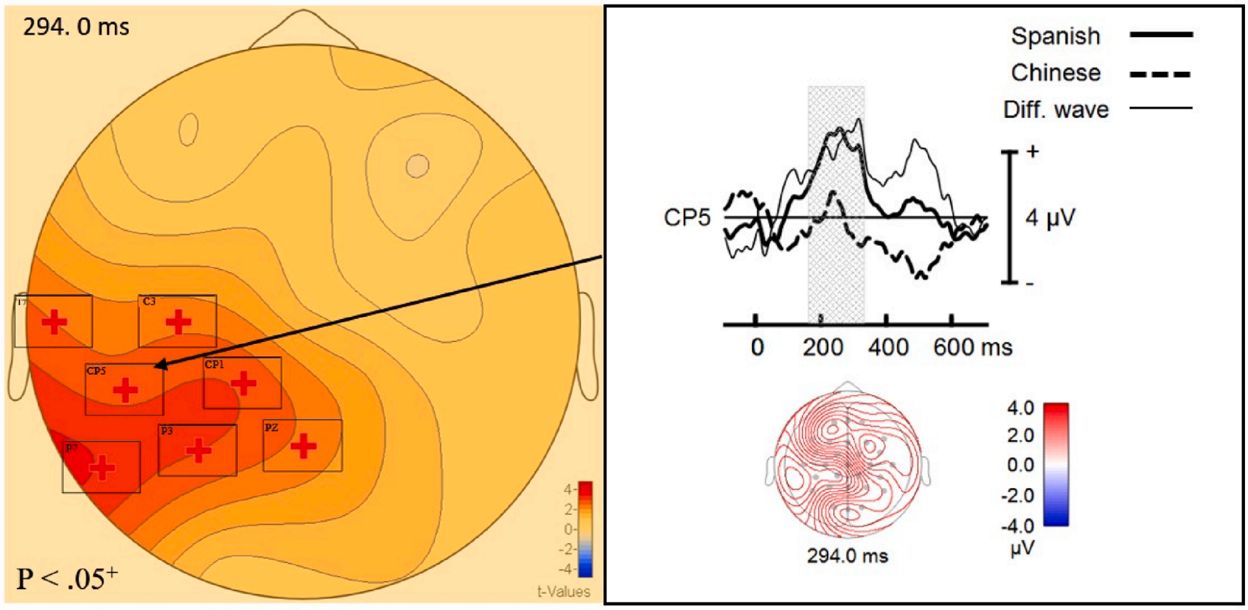

Fig. 4. Positive-MMR to the Spanish and to the Chinese contrast. Significant electrode cluster obtained by BESA statistics. The left side of the figure shows the t-values map at the maximum amplitude difference between the Spanish and the Chinese responses ($p = .04$). The right side of the figure shows the positive-MMRs and difference waveform (Spanish minus Chinese). The dark bar in the ERP responses represents the time window in which both responses differed in a significant way (156–327 ms). The voltage map is time-locked to the maximum amplitude difference observed in the difference waveform.

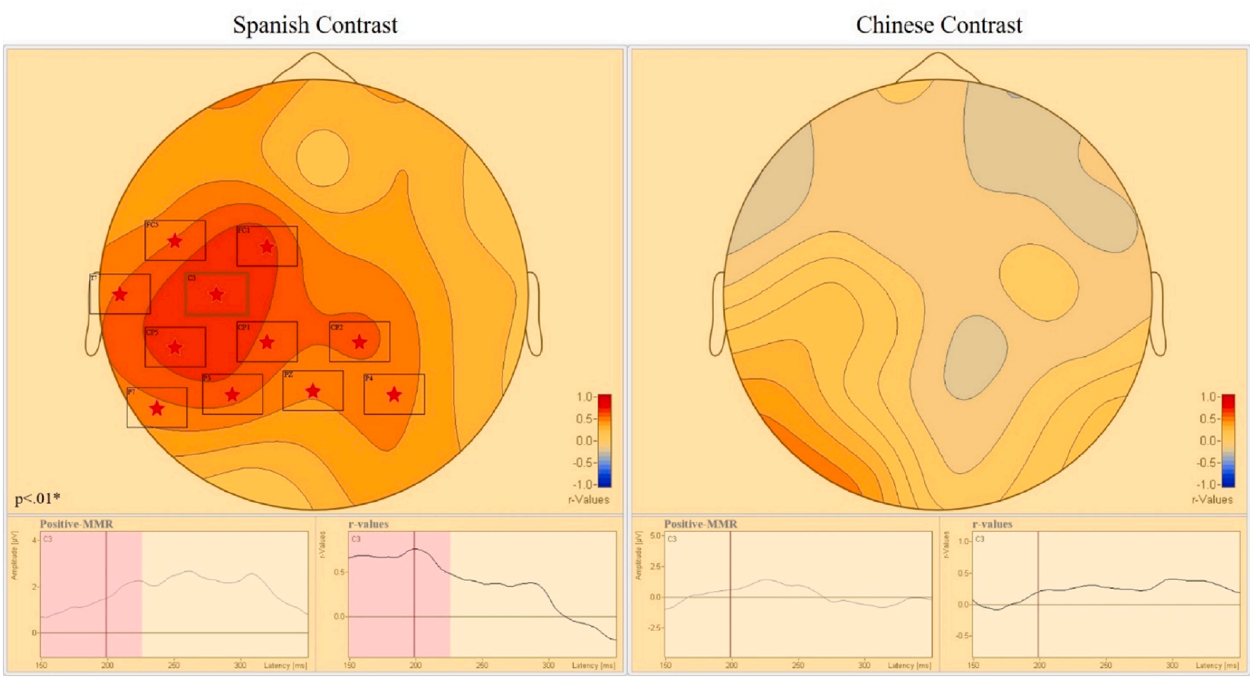

Fig. 5. Pearson correlation maps between caregivers' VSA and infants' positive-MMR. The upper part of the figure depicts the r-value maps. Only the electrodes showing a significant difference between the variables of interest are display. Cluster analysis for the Spanish contrast revealed one significant electrode cluster (left-central-parietal) between 150 and 270 ms after stimulus onset. The Chinese contrast did not show significant clusters. In each panel, the left-bottom box shows the positive-MMR (Deviant minus Control-Deviant) for C3 between 150 and 350 ms after stimulus onset, and the right-bottom box show the r-values for the point-by-point correlation between the variables of interest. The highlighted areas depict the significant r-values using point-by-point comparisons. The cursor is at 199 ms after stimulus onset.

5.5. Does caregivers' IDS affect infants' positive-MMRs, considering other predictors?

In order to learn if there were other predictors associated with the amplitude change of the positive-MMR (Spanish and Chinese), we carried out a linear mixed model. First, we extract the *fundamental frequency (F0)*. In order to observe the degree to which F0 is associated with brain responses, we extracted the fundamental frequency (F0) from each vowel of interest within the boundaries reported in the VSA values obtained from the reading task at home. We obtained F0 values at 5-time intervals of the predefined vowel segment (10%, 30%, 50%, 70%, and 90%). The F0 values were submitted to a 3 (vowel: /i/, /u/, and /a/) by

5 (Measures: 10%, 30%, 50%, 70%, and 90%) repeated measures ANOVA using Greenhouse-Geisser epsilon (ε) for non-sphericity correction. There was no significant interaction between vowel and measurement ($F (3.55, 56.89) = 1.56$, $p = .202$, $\eta p^2 = 0.080$); therefore, each vowel's F0 value was averaged across the 5 measurements. These values were used as a predictor in the mixed linear model (/i/ Mean = 250.93 Hz; *SE* = 12.18; /u/ Mean = 268.01 Hz; *SE* = 12.21; /a/ Mean = 241.14 Hz; *SE* = 9.34).

Second, we selected 5 representative electrodes (FC5, T7, C3, CP5, and CP1) that showed the most significant correlations across caregivers' IDS and infants' positive-MMRs to the Spanish contrast (182–227 ms after stimulus onset; see Fig. 5). In order to include the positive-MMR associated with the Chinese contrast in the model, we used the same electrodes and same time interval as the Spanish contrast. The positive-MMR for the Spanish contrast had a mean of 2.00 μV (*SD* = 1.86) and the mean for the Chinese contrast was −0.501 μV (*SD* = 3.25).

Third, the brain measures were treated as a bivariate response, while VSA along with SES, infants' age, and F0 (for the 3 vowels of interest) were treated as predictors in which the regression coefficients were allowed to be different for each response. The model was fit using maximum likelihood estimation and with heterogeneous compound symmetry correlation structure due to the high variance associated within the estimates of covariance parameters (Chinese contrast Estimate = 7.90; *SE* = 2.711 and Spanish contrast Estimate = 1.904; SE = 0.653).

The results showed no significant estimates, except for VSA. That is, VSA had a significant effect in the positive-MMR associated with the Spanish contrast $t(17) = 2.48$, $p = .024$; 95% CI [3.31, 40.55], $d = 0.61$. The estimate showed that the positive-MMR for the Spanish contrast increased 22 μV (SE = 8.8) per unit increase in VSA (LnHz^2), holding all other predictors constant. In contrast, VSA did not predict the positive-MMR to the Chinese contrast $t(17) = 0.74$, $p = .470$; 95% CI [−24.61, 51.24], $d = 0.18$.

6. Discussion

The overall goal of this investigation was to understand how the quality of infants' language exposure impacts early neural representations of language learning across diverse language environments. Thus, we assessed caregivers' IDS in natural settings (i.e., VSA as a quality of language exposure), and infants' speech sound discrimination ability between native sounds and non-native sounds using ERPs (i.e., positive-MMRs as an indication of early language learning). We also accounted for caregivers' SES, infants' age and F0 by including these variables as predictors in the analyses to see a more straightforward relationship between IDS and infants' language commitment. We found that VSA had a significant effect on the positive-MMR only for the native speech contrast, even after holding the other predictors constant.

6.1. The positive-MMR as a component of language commitment

Although the association between IDS and brain responses has been scarcely explored in the literature, there is an agreement that IDS facilitates attentional resources associated with phonetic processing and word learning (Bosseler et al., 2016; Peter et al., 2016; Uther et al., 2012; Zangl & Mills, 2007; Zhang et al., 2011). Most studies interested in IDS and brain responses have distinguished IDS from ADS by manipulating the acoustic properties of speech, primarily by formant exaggeration (i.e., hyperarticulation). Those research studies have shown that formant-exaggerated IDS, when compared to its non-formant exaggerated counterpart ADS, facilitates infants' ability to distinguish between vowel sounds and familiar words (Peter et al., 2016; Zangl & Mills, 2007; Zhang et al., 2011). Thus, this evidence suggests that the rich acoustic properties of IDS alter infants' brain activity for phonetic processing, which then can optimize word learning. This investigation goes further not only by correlating IDS measures observed in a naturalistic setting (VSA) with an ERP component associated with enhanced attentional processes during early language learning (positive-MMR), but also by extending assessments of this relationship beyond the English language in monolingual infants to the Spanish language in Spanish-English bilingual infants.

The results found in the present investigation support the idea that the positive-MMR to native speech contrasts reflect early stages on language specialization. Namely, we recorded infants' brain responses associated with both a native speech contrast (Spanish), and a non-native speech contrast (Mandarin Chinese) presented in an ADS-style. This was done to differentiate whether the positive-MMR reflects attentional processes in native language learning or in discriminating complex acoustic properties aside from native language learning. Given how adult directed speech is hypothesized to recruit enhanced attentional processes in infants, and that bilinguals at about one year of age begin to show positive-MMRs (García-Sierra et al., 2016), we speculated that bilinguals would show enhanced attentional processes (positive-MMRs) to both the native and non-native speech contrasts in an ADS-style (Peter et al., 2016). Indeed, our results support this hypothesis: a positive-MMR was found for both speech contrasts. However, the positive-MMR response to the native contrast showed substantial differences when compared with the non-native contrast. First, we found a widespread pattern of brain activation for the native speech contrast (frontal areas of both hemispheres), whereas there was a more restricted pattern of activation in the right-frontal area to the non-native speech contrast (see Fig. 3). These findings align with Ferjan Ramírez, Ramírez, Clarke, Taulu, and Kuhl (2017) who reported a larger, widespread MMR activation in response to a native compared to a non-native speech contrast (frontal areas of both hemispheres) in 11-month old monolingual infants.

Second, we observed a larger positive-MMR response to the native compared to the non-native contrast in left-central electrodes (see Fig. 4). We interpret the larger positive-MMR amplitude as neuronal sensitivity to the native speech contrast, which aligns with the neural language commitment hypothesis (Kuhl, 2000, 2004). Accordingly, our bilingual infants showed more complex attentional mechanisms associated with language specialization for the native speech contrast, but not for the non-native speech contrast. In other words, bilingual infants' brains "tune" into ambient speech sounds of their native language as language exposure increases over time, and then goes onto facilitate attentional processes in language learning (Jusczyk et al., 1993, 1994). We argue that perceptual routines that are not fully learned require increased attention and are manifested in the form of positive-MMRs (García-Sierra et al., 2016; Strange, 2011). Furthermore, we can infer that they will show neural commitment in the form of a negative-MMR as they get older (Kuhl et al., 2008). This neural pattern aligns with research showing that infants eventually develop selective perceptual routines to promote efficient and automatic detection of native-language speech sound contrasts (Zhang et al., 2005). Consequently, fully learned selective perceptual routines result in explicit cortical representations (i.e., adult-like MMRs) with low attentional processes.

It should be noted that it is hard to tease apart the underlying mechanisms behind the positive-MMR. For example, there is evidence favoring the idea that the positive-MMR reflects attentional mechanisms associated with general auditory discrimination abilities, as well as the idea that it reflects attentional mechanisms associated with early language learning. In favor of a general auditory discrimination ability, Marklund et al. (2019) reported positive-MMRs to native speech sounds and non-speech sounds (rotated speech) in 4–8 months-old infants. However, other researchers have shown that the positive-MMR depends on the characteristics of the stimuli tested and language experience. Namely, Cheng et al. (2015) showed that native speech contrasts that are hard to discriminate elicit positive-MMRs, while native speech contrasts that are easy to discriminate elicit negative-MMRs in 6-month-old infants. Even more, Cheng's et al. results showed a transition from a positive- to negative-MMRs for the hard to discriminate speech contrasts as a function of age. Similarly, Peter et al. (2016) reported positive-MMRs in 9-month-old infants to a native speech contrast presented in ADS-style (hard to discriminate), but negative-MMRs to the same speech contrast when presented in IDS-style (easy to discriminate). The above results suggest that the generators behind the positive-MMR interact with brain maturation and language experience. In line with that idea, García-Sierra et al. (2016) collected brain responses from 11 to 14 month-old infants to native and non-native speech contrasts presented in ADS-style. The results showed that infants with low amounts of language input elicited positive-MMRs, while infants with large amounts of language input elicited negative-MMRs. Importantly, this relation was observed only with the native speech contrasts. Hence, the attentional mechanisms underlying the elicitation of the positive-MMR to native speech sounds seem to be related with language specialization, while the attentional mechanisms behind the positive-MMR to non-native speech and non-speech seem not to be related with language specialization.

At first, our interpretation of the positive-MMR seems to be at odds with the early literature describing the positive-MMR as an "acoustic form of analysis" (Rivera-Gaxiola, Silva-Pereyra et al., 2005). However, as expressed next, we believe our claims are not at odds, and in fact adds valuable information concerning the neuronal mechanisms behind language specialization. In previous research, Rivera-Gaxiola, Klarman et al. (2005) showed that infants displaying large positive-MMRs to a non-native speech contrast produced more words later in development than infants showing negative-MMRs to the same non-native speech contrast. The authors concluded that the positive-MMR to the non-native contrast reflected infants' ability to track the acoustics of the speech signal without integrating it to the native phonology (acoustic form of analysis), while the negative-MMR reflected integration of the acoustics of a non-native sound into the native phonology.

We believe that our claims regarding the positive-MMR complement Rivera-Gaxiola et al. (2005) idea. Specifically, in accordance with Jusczyk et al. (1993, 1994) we propose that infants' attentional demands are reduced with increasing language experience and brain maturation, and that infants develop attentional mechanisms (selective perceptual routines), to process speech sounds efficiently and automatically (Strange, 2011). With this in mind, it is not surprising to find that infants lack efficient perceptual routines during the first years of language learning to both native and non-native speech contrasts. Thus, lacking perceptual routines to non-native speech contrasts is an indirect measure of language commitment since it reflects, as expected, lack of specialized perceptual routines. On the other hand, exposure to the native language enhances the attentional mechanisms needed to form efficient perceptual routines to native speech contrasts. Therefore, infants showing evidence of language commitment should show better perceptual routines to native, as opposed to non-native speech contrasts. This pattern has been observed such that infants show a transition from a positive- to a negative-MMR to native speech contrasts, reflecting the development of more efficient processing (Rivera-Gaxiola, Klarman, et al., 2005; García-Sierra et al., 2016). Finally, in the same manner that the lack of perceptual routines to non-native speech contrasts (positive-MMR) predicts later language abilities, it should be expected that the presence of efficient perceptual routines (negative-MMR) to native speech contrasts predict future language outcomes. This relationship has also been observed (Kuhl et al., 2008; García-Sierra et al., 2011).

Overall, our results are in accordance with the idea that more efficient neuronal processing of speech sounds results in explicit cortical representation (i.e., negative-MMR) with low attentional demands, while less efficient neuronal processing of speech sounds results in increased attentional demands that are manifested in the form of positive-MMRs. In addition, our results add information to the infant positive-MMR. Specifically, the present investigation shows positive-MMRs to both, native and non-native speech contrasts, but the positive-MMR to the native contrast was broader and stronger in left central electrodes. Finally, only the positive-MMR to the native contrast correlated with the quality of speech (in the form of the size of VSA) received from the primary caregiver.

6.2. *Vowel space area as a component of quality of speech*

Ramírez-Esparza et al. (2014; 2017a) demonstrated that caregivers' natural use of IDS is correlated with language development. The authors qualitatively measured caregivers' IDS from sample files of language activity, and reported that percentages of IDS used by the caregivers during infancy was positively associated with infants' vocabulary size at 2 years of age. Here, we complemented Ramírez-Esparza and colleagues' studies by rather quantifying IDS as a function of vowel hyperarticulation. Specifically, we used VSA because it tests the idea that the hyperarticulation of the vowels is what leads children to listen to clear tokens and learn language faster (Kuhl et al., 1997). Further, VSA has been shown to be a good indicator of language learning in Tawainese infants in head turn procedures (Liu & Kuhl, 2003), word recognition (Song et al., 2010), and expressive and receptive child language outcomes (Hartman et al., 2017).

However, there is research that has questioned the hyperarticulation hypothesis. Specifically, studies have demonstrated that IDS does not necessarily produce clear vowel tokens (Martin et al., 2015), and that the benefits of the hyperarticulation can be confounded with other prosodic mechanisms (McMurray et al., 2013) as well as the affectivity placed in IDS (Miyazawa et al., 2017). The mixed findings could be the result of methodological approaches, including analyses of tokens from a language that is mostly monophthongal (i.e., Japanese) and from spontaneous speech (e.g., Martin et al., 2015; Miyazawa et al., 2017). In this investigation, we introduced several methodological approaches to address these mixed findings on the effects of the hyperarticulation used in IDS. Specifically, we used a semi-spontaneous approach where caregivers read a storybook (in which the vowels /i/, /u/, and /a/ were used) in their dominant language (i.e., Spanish, a monophthongal language) to their child at home during a moment and time of their choice. This approach, which did not explicitly instruct the caregivers to use IDS, allowed us to obtain a controlled assessment of IDS in natural environments while preserving a normal distribution of IDS used among caregivers. Indeed, the results demonstrated that there are caregivers with lower VSA scores, and those with higher VSA scores as measured from the reading task at home. Further analyses demonstrated that in their natural environments, some caregivers use IDS, while others use speech that is comparable to adult speech, which suggests that infants' everyday exposure to IDS varies from family to family (Ramírez-Esparza et al., 2016; 2017a).

6.3. *Does IDS promote language commitment?*

In order to answer this question, we first assessed the relationship between IDS and the positive-MMR associated with a native Spanish speech contrast, and a non-native Chinese speech contrast. Our speculation was that if Spanish-English bilingual infants' positive-MMR depicts the enhanced attentional processes that support language commitment, then IDS would only correlate with the positive-MMR associated to the native Spanish speech contrast. Indeed, this association was found, markedly in the central-left electrodes. Although a significant positive-MMR was found for the Chinese contrast, it had no significant correlation with caregivers' VSA. Importantly, we did not find significant correlations between the positive-MMRs to each of the contrasts and adult speech VSA scores. These results indicate that IDS has a direct impact on language development, and presumably brain function early in life. These results are in line with previous findings showing that the positive-MMR to native speech contrasts correlates with the amount of language input received at home (García-Sierra et al., 2016).

We also carried out a linear mixed model to explore if other predictors known to influence language processing also influenced the amplitude of the positive-MMR in this study. Specifically, we included

SES because of its association with infants' language input (Hoff, 2006; Gilkerson et al., 2017; Rowe, 2008; Rowe & Goldin-Meadow, 2009), in which low amounts of language input can affect brain function (specialization), and thus language development (Raizada et al., 2008; Rowe, 2008). Furthermore, SES has been tied with IDS, such that caregivers with higher SES tend to use more IDS in natural environments (Ramírez-Esparza et al., 2014; 2017a). Age was also considered as a predictor in the model given that our sample included infants belonging to two age groups and language commitment has been associated with age (e.g., García-Sierra et al., 2016; Cheng et al., 2015). Finally, F0 was considered as a predictor given that F0 facilitates infants' neural processing of linguistic information (Bosseler et al., 2016; Saito et al., 2007; Zhang et al., 2011; Zangl & Mills, 2007; Peter et al., 2016). Even after including SES, infants' age and F0, we found that VSA was the only predictor that significantly influenced the increment of the positive-MMR. Overall, these analyses suggest that the quality of the speech as measured with VSA facilitates early language commitment.

6.4. Limitations

It is important to highlight some of the limitations of this study. First, the sample size is small and therefore findings should be taken cautiously. It is possible that other predictors could have also affected the increment of the positive-MMR with increased power. Second, we developed a methodology to extract natural variations of VSA, by using a semi-structured approach (i.e., storybooks) with the LENA recorder. Although we gained some controllability in the variation of the use of tokens, we lost the richness of spontaneous speech that caregivers use when talking to their infants in their everyday natural environments (Ramírez-Esparza et al., 2014; 2017a; 2017b). Furthermore, in a lab setting with advanced recording technologies in a soundproof booth, we would have been able to extract cleaner tokens for IDS and ideally be able to compare them to adult *directed* speech (e.g., Burnham et al., 2015; Inoue et al., 2011; McMurray et al., 2013). In this investigation, our tokens of adult speech were limited in the sense that they were not extracted from an interaction between the caregiver and the child. Finally, research with a purely Spanish monolingual sample is necessary to observe how the VSA hyperaticulation in Spanish compares to the bilingual sample. In sum, future studies should complement this study by doing VSA analyses with similar methodologies in other languages with larger samples and with infants with different degrees of language commitment (positive- and negative-MMRs). This can help to better understand the interplay of quality of input and early representations of language commitment in the brain.

7. Conclusion

We showed evidence that caregivers' infant directed speech (IDS) is associated with infants' brain function. We found that caregivers' VSA correlated with a larger positive-MMR that was associated with a native speech sound, even after holding other predictors constant. Our results add information to the big picture of language development by showing that the increased attentional processes that facilitate speech processing during the first years of life can be fostered by caregivers' hyper-articulated speech. The findings of this investigation have implications for understanding how bilinguals learn their two languages, and the importance of having interactions of quality to foment language learning.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The research reported here was supported by a National Science Foundation Science of Learning Program grant to the LIFE Center at the University of Washington, Seattle WA. (SBE-0354453). We thank Dr. Elizabeth Schifano at the University of Connecticut for her guidance in the implementation of the linear mixed model.

References

Blomgren, M., Robb, M., & Chen, Y. (1998). A note on vowel centralization in stuttering and nonstuttering individuals. *Journal of Speech, Language, and Hearing Research, 41*(5), 1042–1051. https://doi.org/10.1044/jslhr.4105.1042.

Boersma, P., Weenink, D. (2018). Praat: doing phonetics by computer [Computer program]. Version 6.0.37, retrieved 14 March 2018 from http://www.praat.org/.

Bosseler, A. N., Teinonen, T., Tervaniemi, M., & Huotilainen, M. (2016). Infant directed speech enhances statistical learning in newborn infants: An ERP study. *PLoS ONE, 11*(9), Article e0162177. https://doi.org/10.1371/journal.pone.0162177.

Bullmore, E. T., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E., & Brammer, M. J. (1999). Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. *IEEE Transactions on Medical Imaging, 18*(1), 32–42. https://doi.org/10.1109/42.750253.

Burnham, E. B., Wieland, E. A., Kondaurova, M.v., McAuley, Devin, J., Bergeson, T. R., & Dilley, L. C. (2015). The development of English as a second language with and without specific language impairment: Clinical implications. *Journal of Speech, Language, and Hearing Research, 58*, 241–253. https://doi.org/10.1044/2015.

Cheng, Y.-Y., Wu, H. C., Tzeng, Y. L., Yang, M. T., Zhao, L. L., & Lee, C. Y. (2013). The development of mismatch responses to mandarin lexical tones in early infancy. *Developmental Neuropsychology, 38*(5), 281–300. https://doi.org/10.1080/87565641.2013.799672.

Cheng, Y.-Y., Wu, H.-C., Tzeng, Y.-L., Yang, M.-T., Zhao, L.-L., & Lee, C.-Y. (2015). Feature-specific transition from positive mismatch response to mismatch negativity in early infancy: Mismatch responses to vowels and initial consonants. *International Journal of Psychophysiology, 96*(2), 84–94. https://doi.org/10.1016/j.ijpsycho.2015.03.007.

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998a). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience, 1*, 351–353. https://doi.org/10.1038/1561.

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998b). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience, 1*(5), 351–353. https://doi.org/10.1038/1561.

Ernst, M. D. (2004). Permutation methods: A basis for exact inference. *Statistical Science, 19*(4), 676–685. https://doi.org/10.1214/088342304000000396.

Ferjan Ramírez, N., Ramírez, R. R., Clarke, M., Taulu, S., & Kuhl, P. K. (2017). Speech discrimination in 11-month-old bilingual and monolingual infants: A magnetoencephalography study. *Developmental Science, 20*(1), Article e12427. https://doi.org/10.1111/desc.12427.

Fernald, A., & Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development, 10*(3), 279–293. https://doi.org/10.1016/0163-6383(87)90017-8.

Fish, M. S., García-Sierra, A., Ramírez-Esparza, N., & Kuhl, P. K. (2017). Infant-directed speech in English and Spanish: Assessments of monolingual and bilingual caregiver VOT. *Journal of Phonetics, 63*. https://doi.org/10.1016/j.wocn.2017.04.003.

Friederici, A. D., Friedrich, M., & Christophe, A. (2007). Brain responses in 4-month-old infants are already language specific. *Current Biology, 17*(14), 1208–1211. https://doi.org/10.1016/j.cub.2007.06.011.

Friedrich, M., Herold, B., & Friederici, A. D. (2009). ERP correlates of processing native and non-native language word stress in infants with different language outcomes. *Cortex, 45*(5), 662–676. https://doi.org/10.1016/j.cortex.2008.06.014.

García-Sierra, A., Ramírez-Esparza, N., & Kuhl, P. K. (2016). Relationships between quantity of language input and brain responses in bilingual and monolingual infants. *International Journal of Psychophysiology, 110*, 1–17. https://doi.org/10.1016/j.ijpsycho.2016.10.004.

García-Sierra, A., Rivera-Gaxiola, M., Percaccio, C. R., Conboy, B. T., Romo, H., Klarman, L., Ortiz, S., & Kuhl, P. K. (2011). Bilingual language learning: An ERP study relating early brain responses to speech, language input, and later word production. *Journal of Phonetics, 39*, 546–557. https://doi.org/10.1016/j.wocn.2011.07.002.

Gilkerson, J., Richards, J. A., Warren, S. F., Montgomery, J. K., Greenwood, C. R., Oller, D. K., Hansen, J. H. L., & Paul, T. D. (2017). Mapping the early language environment using all-day recordings and automated analysis. *American Journal of Speech-Language Pathology, 26*(2), 248–265. https://doi.org/10.1044/2016_AJSLP-15-0169.

Gomes, H., Molholm, S., Ritter, W., Kurtzberg, D., Cowan, N., & Vaughan, H. G. (2000). Mismatch negativity in children and adults, and effects of an attended task. *Psychophysiology, 37*(6), 807–816. https://doi.org/10.1111/1469-8986.3760807.

Hartman, K. M., Ratner, N. B., & Newman, R. S. (2017). Infant-directed speech (IDS) vowel clarity and child language outcomes. *Journal of Child Language, 44*(5), 1140–1162. https://doi.org/10.1017/S0305000916000520.

Hisagi, M., Shafer, V. L., Strange, W., & Sussman, E. S. (2010). Perception of a Japanese vowel length contrast by Japanese and American English listeners: Behavioral and electrophysiological measures. *Brain Research, 1360*, 89–105. https://doi.org/10.1016/j.brainres.2010.08.092.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review, 26*(1), 55–88. https://doi.org/10.1016/j.dr.2005.11.002.

Hollingshead, A. B. (2011). Four factor index of social status. *Yale Journal of Sociology., 8*, 21–52.

Inoue, T., Nakagawa, R., Kondou, M., Koga, T., & Shinohara, K. (2011). Discrimination between mothers' infant- and adult-directed speech using hidden Markov models. *Neuroscience Research, 70*(1), 62–70. https://doi.org/10.1016/j.neures.2011.01.010.

Jusczyk, P. W., Friederici, A. D., Wessels, J. M., Svenkerud, V. Y., & Jusczyk, A. M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory & Language, 32*(3), 402–420. https://doi.org/10.1006/jmla.1993.1022.

Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory & Language, 33*(5), 630–645. https://doi.org/10.1006/jmla.1994.1030.

King, C., McGee, T., Rubel, E. W., Nicol, T., & Kraus, N. (1995). Acoustic features and acoustic changes are represented by different central pathways. *Hearing Research, 85* (1–2), 45–52. https://doi.org/10.1016/0378-5955(95)00028-3.

Kraus, N., McGee, T., Carrel, T. D., King, C., Tremblay, K., & Nicol, T. (1995). Central auditory system plasticity associated with speech discrimination training. *Journal of Cognitive Neuroscience, 7*(1), 25–32. https://doi.org/10.1162/jocn.1995.7.1.25.

Kraus, N., McGee, T., Sharma, A., Carrell, T., & Nicol, T. (1992). Mismatch negativity event-related potential elicited by speech stimuli. *Ear and Hearing, 13*(3), 158–164. https://doi.org/10.1097/00003446-199206000-00004.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., … Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science, 277*(5326), 684–686. https://doi.org/10.1126/science.277.5326.684.

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences, 97*(22), 11850. https://doi.org/10.1073/pnas.97.22.11850.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5*, 831.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & 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 B-Biological Sciences, 363*(1493), 979–1000. https://doi.org/10.1098/rstb.2007.2154.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by six- months of age. *Science, 255*, 606–608. https://doi.org/10.1126/science.1736364.

Lee, C.-Y., Yen, H.-L., Yeh, P.-W., Lin, W.-H., Cheng, Y.-Y., Tzeng, Y.-L., & 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.

Liu, H.-M., Chen, Y., & Tsao, F.-M. (2014). Developmental changes in mismatch responses to mandarin consonants and lexical tones from early to middle childhood. *PLoS ONE, 9*(4). https://doi.org/10.1371/journal.pone.0095587.

Liu, H.-M., Kuhl, K., P, & Tsao, F.-M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. Developmental Science, 6(3), F1-F10. Retrieved from http://dx.doi.org/10.1111/1467-7687.00275.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.

Marklund, E., Schwarz, I.-C., & Lacerda, F. (2019). Amount of speech exposure predicts vowel perception in four- to eight-month-olds. *Developmental Cognitive Neuroscience, 36*, Article 100622. https://doi.org/10.1016/j.dcn.2019.100622.

Martin, A., Schatz, T., Versteegh, M., Miyazawa, K., Mazuka, R., Dupoux, E., & Cristia, A. (2015). Mothers speak less clearly to infants than to adults: A comprehensive test of the hyperarticulation hypothesis. *Psychological Science, 26*(3), 341–347. https://doi.org/10.1177/0956797614562453.

May, P. J. C., & Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology, 47*(1), 66–122. https://doi.org/10.1111/j.1469-8986.2009.00856.x.

McMurray, B., Kovack-Lesh, K. A., Goodwin, D., & McEchron, W. (2013). Infant directed speech and the development of speech perception: Enhancing development or an unintended consequence? *Cognition, 129*(2), 362–378. https://doi.org/10.1016/j.cognition.2013.07.015.

Mehl, M. R., & Holleran, S. E. (2007). An empirical analysis of the obtrusiveness of and participants' compliance with the Electronically Activated Recorder (EAR). *European Journal of Psychological Assessment, 23*(4), 248–257. https://doi.org/10.1027/1015-5759.23.4.248.

Mills, D. L., Coffey-Corina, S., & Neville, H. J. (1997). Language comprehension and cerebral specialization from 13 to 20 months. *Developmental Neuropsychology, 13*(3), 397–445. https://doi.org/10.1080/87565649709540685.

Miyazawa, K., Shinya, T., Martin, A., Kikuchi, H., & Mazuka, R. (2017). Vowels in infant-directed speech: More breathy and more variable, but not clearer. *Cognition, 166*, 84–93. https://doi.org/10.1016/j.cognition.2017.05.003.

Morr, M. L., Shafer, V. L., Kreuzer, J. A., & Kurtzberg, D. (2002). Maturation of mismatch negativity in typically developing infants and preschool children. *Ear and Hearing, 23* (2), 118–136. https://doi.org/10.1097/00003446-200204000-00005.

Näätänen, R. (1992). *Attention and brain function*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., … Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature, 385*, 432–434. https://doi.org/10.1038/385432a0.

Oller, D. K., Niyogi, P., Gray, S., Richards, J. A., Gilkerson, J., Xu, D., Yapanel, U., & Warren, S. F. (2010). Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences of the United States of America, 107*(30), 13354–13359. https://doi.org/10.1073/pnas.1003882107.

Peter, V., Kalashnikova, M., Santos, A., & Burnham, D. (2016). Mature neural responses to infant-directed speech but not adult-directed speech in pre-verbal infants. *Science Reports, 6*, 34273. https://doi.org/10.1038/srep34273.

Ramírez-Esparza, N., García-Sierra, A., & Kuhl, P. K. (2014). Look who's talking: Speech style and social context in language input to infants are linked to concurrent and future speech development. *Developmental Science, 17*(6), 880–891. https://doi.org/10.1111/desc.12172.

Ramírez-Esparza, N., García-Sierra, A., & Kuhl, P. K. (2017a). The Impact of early social interactions on later language development in Spanish-English bilingual infants. *Child Development, 88*(4), 1216–1234. https://doi.org/10.1111/cdev.12648.

Ramírez-Esparza, N., García-Sierra, A., & Kuhl, P. K. (2017b). Look who's talking NOW! parentese speech, social context, and language development across time. *Frontiers in Psychology, 8*(1008). https://doi.org/10.3389/fpsyg.2017.01008.

Ramírez-Esparza, N., García-Sierra, A., Rodríguez-Arauz, G., Ikizer, E. G., & Fernández-Gómez, M. J. (2019). No laughing matter: Latinas' high quality of conversations relate to behavioral laughter. *PLoS ONE, 14*(4), 1–23. https://doi.org/10.1371/journal.pone.0214117.

Raizada, R. D. S., Richards, T. L., Meltzoff, A., & Kuhl, P. K. (2008). Socioeconomic status predicts hemispheric specialisation of the left inferior frontal gyrus in young children. *Neuroimage, 40*(3), 1392–1401. Retrieved from: //000254728200038.

Rivera-Gaxiola, M., Klarman, L., García-Sierra, A., & Kuhl, P. K. (2005). Neural patterns to speech and vocabulary growth in American infants. *Neuroreport: For Rapid Communication of Neuroscience Research, 16*(5), 495–498. https://doi.org/10.1097/00001756-200504040-00015.

Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2005). Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Developmental Science, 8*(2), 162–172. https://doi.org/10.1111/j.1467-7687.2005.00403.x.

Rowe, M. L. (2008). Child-directed speech: Relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language, 35*(1), 185–205. https://doi.org/10.1017/S0305000907008343.

Rowe, M. L., & Goldin-Meadow, S. (2009). Differences in early gesture explain SES disparities in child vocabulary size at school entry. *Science, 323*(5916), 951–953. https://doi.org/10.1126/science.1167025.

Sabri, M., & Campbell, K. B. (2002). The effects of digital filtering on mismatch negativity in wakefulness and slow-wave sleep. *Journal of Sleep Research, 11*(2), 123–127. https://doi.org/10.1046/j.1365-2869.2002.00292.x.

Saito, Y., Aoyama, S., Kondo, T., Fukumoto, R., Konishi, N., Nakamura, K., Kobayashi, M., & Toshima, T. (2007). Frontal cerebral blood flow change associated with infant-directed speech. *Archives of Disease in Childhood - Fetal and Neonatal Edition, 92*(2), F113–F116. https://doi.org/10.1136/adc.2006.097949.

Sapir, S., Połczyńska, M., & Tobin, Y. (2009). Why does the vowel space area as an acoustic metric fail to differentiate dysarthric from normal vowel articulation and what can be done about it? *Poznan Studies in Contemporary Linguistics, 45*(2), 301–311. https://doi.org/10.2478/v10010-009-0018-2.

Song, J. Y., Demuth, K., & Morgan, J. (2010). Effects of the acoustic properties of infant-directed speech on infant word recognition. *The Journal of the Acoustical Society of America, 128*(1), 389–400. https://doi.org/10.1121/1.3419786.

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.

Swingley, D., & Alarcón, C. (2018). Lexical learning may contribute to phonetic learning in infants: A corpus analysis of maternal Spanish. *Cognitive Science, 42*(5), 1618–1641. https://doi.org/10.1111/cogs.12620.

Trainor, L., McFadden, M., Hodgson, L., Darragh, L., Barlow, J., Matsos, L., & Sonnadara, R. (2003). Changes in auditory cortex and the development of mismatch negativity between 2 and 6 months of age. *International Journal of Psychophysiology, 51*(1), 5–15. https://doi.org/10.1016/S0167-8760(03)00148-X.

Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2006). Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants. *Journal of the Acoustical Society of America, 120*(4), 2285–2294. https://doi.org/10.1121/1.2338290.

Uther, M., Giannakopoulou, A., & Iverson, P. (2012). Hyperarticulation of vowels enhances phonetic change responses in both native and non-native speakers of English: Evidence from an auditory event-related potential study. *Brain Research, 1470*, 52–58. https://doi.org/10.1016/j.brainres.2012.06.041.

Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology, 37*, 278–286. https://doi.org/10.1037/h0080725.

Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition, 103*(1), 147–162. https://doi.org/10.1016/j.cognition.2006.03.006.

Zangl, R., & Mills, D. L. (2007). Increased brain activity to infant-directed speech in 6- and 13-month-old infants. *Infancy, 11*(1), 31–62. https://doi.org/10.1207/s15327078in1101_2.

Zhang, Y., Koerner, T., Miller, S., Grice-Patil, Z., Svec, A., Akbari, D., Tusler, L., & Carney, E. (2011). Neural coding of formant-exaggerated speech in the infant brain. *Developmental Science, 14*(3), 566–581. https://doi.org/10.1111/j.1467-7687.2010.01004.x.

Zhang, Y., Kuhl, P. K., Imada, T., Kotani, M., & Tohkura, Y. (2005). Effects of language experience: Neural commitment to language-specific auditory patterns. *Neuroimage, 26*(3), 703–720. https://doi.org/10.1016/j.neuroimage.2005.02.040.