OXFORD

ORIGINAL ARTICLE

Modulation of Theta Phase Synchrony during Syllable Processing as a Function of Interactive Acoustic Experience in Infancy

Silvia Ortiz-Mantilla(iD), Cynthia P. Roesler, Teresa Realpe-Bonilla and April A. Benasich

Center for Molecular and Behavioral Neuroscience, Rutgers University, Newark, NJ 07102, USA

Address correspondence to Silvia Ortiz-Mantilla, Center for Molecular and Behavioral Neuroscience, Rutgers University, Newark, 197 University Avenue Newark, NJ 07102, USA. Email: sortizma@newark.rutgers.edu

Abstract

Plasticity, a prominent characteristic of the infant brain, supports formation of cortical representations as infants begin to interact with and adapt to environmental sensory events. Enhanced acoustic processing efficiency along with improved allocation of attentional resources at 7 months and establishment of well-defined phonemic maps at 9 months have been shown to be facilitated by early interactive acoustic experience (IAE). In this study, using an oddball paradigm and measures of theta phase synchrony at source level, we examined short- and long-term effects of nonspeech IAE on syllable processing. Results demonstrated that beyond maturation alone, IAE increased the efficiency of syllabic representation and discrimination, an effect that endured well beyond the immediate training period. As compared with naive controls, the IAE-trained group at 7, 9, and 18 months showed less theta phase synchrony for the standard syllable and at 7 and 18 months for the deviant syllable. The decreased theta phase synchrony exhibited by the trained group suggests more mature, efficient, acoustic processing, and thus, better cortical representation and discrimination of syllabic content. Further, the IAE modulatory effect observed on theta phase synchrony in left auditory cortex at 7 and 9 months was differentially associated with receptive and expressive language scores at 12 and 18 months of age.

Key words: brain oscillations, infants, plasticity, speech perception, training

Introduction

Across development, acoustic mapping and oscillatory entrainment allow detailed syllabic encoding that enables fast, automatic responses to incoming language. This process is possible due to the significant brain plasticity observed during the first year of life that supports the creation and refinement of functional networks as infants begin to interact with and adapt to their natural environment. Over this crucial period of early ongoing brain plasticity, the foundations of language are established via the interplay among auditory processing abilities, brain maturation, and the infant's interactive experience with the surrounding environment.

Analysis of speech occurs in a progressive fashion, beginning with the acoustic signal, advancing to phonemic units, and then to meaningful words (Greenberg 1998; Hagoort and Poeppel 2013). It is thought that to facilitate this process, speech is organized into syllable-like segments that are the basic units for lexical access from the acoustic level (Greenberg 1998). During early language acquisition, infants must be able to rapidly segment the ongoing speech into these smaller acoustic bits/units that will be mapped within cortical temporal areas and activate phonological representations that will in turn trigger the mental lexicon (Bertoncini and Mehler 1981; Eimas 1999; Jusczyk 1999; Jusczyk et al. 1999; Hickok and Poeppel 2007; Hagoort and Poeppel 2013; Räsänen et al. 2018). Phonemes are the smallest

segments of speech, whereas syllables are the smallest pronounceable linguistic structures, which are able to hold the abstract representation of consonantal phonemes (Eulitz and Lahiri 2004; Hickok and Poeppel 2007). Syllables have fewer and less complex spectro-temporal content than longer, more complex speech forms allowing for faster retrieval and facilitating continuous mapping onto the lexicon (Lahiri and Marslen-Wilson 1991). Moreover, the syllabic structure is key, as it not only indicates the speaking rate but also the stress and tonal contrasts while carrying important prosodic features (Hagoort and Poeppel 2013; Peelle et al. 2013). All these characteristics of prelexical syllabic analysis are advantageous for speech perception particularly during the time over which early language is laid down.

The ability to proficiently accomplish fine-grained acoustic analysis within the tens of milliseconds (ms) range, known as rapid auditory processing (RAP), is an essential skill for decoding and discrimination of phonemic information carried by the syllable and is crucial for the cortical establishment of accurate phonemic representations (Eilers et al. 1981; Aslin 1989; Werker and Tees 2005). For example, studies have shown that RAP abilities differ between typically developing infants and infants at familial risk for developmental language disorders, and importantly, these differences in infant RAP abilities predict later language proficiency (Benasich et al. 2002; Tsao et al. 2004; Guttorm et al. 2005; Benasich et al. 2006; Leppänen et al. 2010; Choudhury and Benasich 2011; Maitre et al. 2013; Benasich et al. 2016; Cantiani et al. 2016, 2019). In a pioneering behavioral study, Eimas et al. (1971), using a voice onset time (VOT) contrast showed that 4-month-old infants discriminate phonemic information within the syllable. But it is between 6 and 12 months of age that infants narrow their linguistic perceptual abilities by constructing phonemic representations (acoustic maps) of the native sounds present in their environment, thus facilitating more efficient native language processing (Werker and Tees 1984; Rivera-Gaxiola et al. 2005; Ortiz-Mantilla et al. 2016). In addition, differing structural and functional maturational trajectories of the temporal lobes, with the left maturing faster than the right (Kinney et al. 1988; Paus et al. 2001; Pujol et al. 2006; Deoni et al. 2011; Adibpour et al. 2020), may support and facilitate RAP of incoming sensory information and phonemic mapping in the left auditory cortex (LAC).

Research in adults has shown that for successful syllabic decoding and for the syllable to be recognized as a perceptual unit, a resolution parameter in the spectral domain and an integration parameter in the temporal domain are required (Poeppel 2003). That explains why most of the syllabic information is encoded by changes occurring in the theta spectral range (Greenberg and Kinsbury 1997), as the theta oscillatory rate aligns with the syllabic temporal rate (Hickok and Poeppel 2000, 2007; Poeppel 2003; Giraud and Poeppel 2012; Ghitza 2013). Specifically, modulation of speech information is associated with temporal envelope fluctuations occurring between 4 and 7 Hz (Obleser et al. 2012; Peelle et al. 2013; Ding and Simon 2014), a spectral range compatible with the syllabic structure, with a peak at 5 Hz reflecting the most common syllabic duration of 200 ms (Greenberg 1998). The theta rhythm closely tracks syllabic information, facilitating the parsing required for successful decoding and intelligibility (Giraud and Poeppel 2012; Ghitza 2013; Ding and Simon 2014; Doelling et al. 2014) with envelope fluctuations in theta reflecting both the rhythmicity of speech and the temporal coherence between acoustic features (Ding and Simon 2014). In addition to encoding syllabic information,

theta oscillations have been shown to support syllable discrimination within native language in infants (Ortiz-Mantilla et al. 2013) and adults (Jin et al. 2014), lexical activation in preschoolers (Meyer et al. 2019), and semantic integration in school-age children (Schneider et al. 2016; Schneider and Maguire 2019). Furthermore, atypical theta power lateralization has been proposed as a neural mechanism that underlies deficits in children with developmental language disorders such as dyslexia (Spironelli et al. 2006; Goswami 2011).

Theta oscillatory phase also plays a critical role during speech perception in adults and contributes to development of semantic processing in preschool-age children (Kikuchi et al. 2011). For instance, a consistent relationship was observed between syllable identity and ongoing oscillatory phase within the theta range, suggesting syllable identity is contingent with theta phase (ten Oever and Sack 2015). Moreover, the phase of low-frequency theta oscillations determines perceptual sensitivity, facilitating the processing of sensory signals, tracking attentional shifts (Kerlin et al. 2010; Zion Golumbic, Cogan, et al. 2013; Zion Golumbic, Ding, et al. 2013), and encoding predictions during speech (syllabic) segmentation (Arnal and Giraud 2012). Thus, theta oscillations encode and discriminate syllabic information and track the acoustic amplitude envelope of the speech signal, while theta phase alignment shapes syllable perception (Kayser et al. 2012; Peelle and Davis 2012) and relates to syllable identity (ten Oever and Sack 2015).

Considerable evidence about the oscillatory underpinnings of speech processing has been obtained from adult populations, but only scant knowledge has been acquired concerning how oscillatory dynamics change over the critical period of early language acquisition. A few studies have shown that, like adults, young infants process syllables (Bosseler et al. 2013; Ortiz-Mantilla et al. 2013, 2016, 2019; Leong et al. 2017; Nacar Garcia et al. 2018) and track the speech envelope (Kalashnikova et al. 2018) within the theta range. One longitudinal study reported that during processing of a syllable contrast varying by 40 ms in voice-onset-time (VOT), phase synchrony of theta oscillations in auditory cortices decreased from 6 to 12 months of age as syllabic processing became more efficient and automatized (Ortiz-Mantilla et al. 2016). Over the same period, intertrial phase synchrony in the theta range also decreases when infants process tone pairs with a speech-like, 70-ms interstimulus interval (ISI) (Hämäläinen et al. 2019), suggesting that neural pathways and auditory perceptual abilities strengthen over development (Watson et al. 2014). However, exactly how the dynamics of phase synchrony in the theta band evolve as a function of increasingly more sophisticated syllable processing after 12 months of age has not been explored. Such information would help us to further understand the neural mechanisms underlying speech perception and discrimination over the developmental period in which young infants transition from processing sublexical syllables to acquiring/processing meaningful words. An earlier study, using a tone contrast containing a duration cue important for Italian language, found a correlation between the magnitude of theta phase synchrony at 6 months and expressive vocabulary at 20 months (Cantiani et al. 2019). The finding that infant theta phase synchrony in response to a speech-like stimuli relates to expressive language suggests that theta phase synchrony responses to syllables may also be a good predictor of later language outcome. The relevance of efficient syllabic processing for parsing, decoding, discrimination, and intelligibility of ongoing speech is well known (Giraud and Poeppel 2012; Ghitza 2013; Ding and Simon 2014; Doelling et al. 2014), along with theta

phase synchrony for syllable perception and identity (Kayser et al. 2012; Peelle and Davis 2012; ten Oever and Sack 2015). However, the association between changes in infant theta phase synchrony in response to syllables and later language abilities has not been systematically investigated.

Given that the infant brain is highly plastic and learning experiences induce cortical representational plasticity (Buonomano and Merzenich 1998; Kudo et al. 2020), a process that is enhanced by attention (Ahveninen et al. 2011), optimizing the earliest stages of acoustic mapping may facilitate a more robust establishment of phonemic representations (Zhang et al. 2001; Moucha et al. 2005). Previously, we demonstrated that over and above maturation alone, interactive acoustic experience (IAE) with spectro-temporally modulated nonspeech sounds significantly enhanced processing of key prelinguistic acoustic cues of known and novel nonspeech stimuli in 7-month-old infants (Benasich et al. 2014), promoting fine-tuning of early acoustic mapping (Musacchia et al. 2017). Moreover, at 9 months of age, the IAE modulatory effect extended to speech processing, fostering more efficient discrimination of phonemic content and facilitating the establishment of cortical phonemic representations (Ortiz-Mantilla et al. 2019). In this study, we propose that, in a similar way, IAE may modulate syllabic processing in the theta range and thus impact subsequent language outcomes. To examine this hypothesis, we use intertrial phase locking (ITPL) within theta oscillations, a measure of phase synchrony across trials, to explore whether experience-dependent effects of early IAE with nonspeech sounds generalize to syllable processing, and further, we examine the relationship between infant theta phase synchrony and language abilities. One group of typically developing 4-month-old infants, the Auditory Experience (AEx) group, a subset of a cohort from a larger longitudinal study, participated for 6 weeks in an IAE with temporally modulated nonspeech sounds in which attention was specifically driven to detect smaller and smaller changes in the auditory environment. These infants were subsequently followed longitudinally until 18 months of age. At 7, 9, 12, and 18 months of age, dense-array electroencephalography (EEG)/event-related potential (ERP) signals were recorded, while the AEx infants passively listened to a VOT phonemic contrast presented in a passive oddball paradigm. To tease apart maturational from experimental effects, four cross-sectional naive control (NC) groups (NC07, NC09, NC12 and NC18), that did not receive any sound training, served as age-matched controls. As the auditory system is highly plastic during these early sensitive periods, previous experience with the experimental sounds had the potential to influence the brain response; a cross-sectional design assured that every control group was naive to the syllable tokens used in the ERP paradigm. Using source localization techniques, which allow more precise analysis of developing lateralization patterns, we also examined event-related oscillatory phase dynamics in left and right auditory cortices. We were particularly interested in investigating whether the infants' amount of theta ITPL to the syllable contrast varied across age as a function of acoustic experience in the trained infants when compared at each age with NCs. It is known that: 1) theta oscillations support infant syllable processing (Bosseler et al. 2013; Ortiz-Mantilla et al. 2013, 2016, 2019; Leong et al. 2017; Nacar Garcia et al. 2018); 2) intertrial phase synchrony of theta oscillations in auditory cortices decreases from 6 to 12 months as processing of native syllables becomes more efficient (Ortiz–Mantilla et al. 2016); 3) IAE with nonspeech acoustic signals sharpens infants' abilities to detect and discriminate known and novel nonspeech stimuli (Benasich et al. 2014) and promotes

fine-tuning of acoustic representations (Musacchia et al. 2017) and phonemic mapping (Ortiz-Mantilla et al. 2019); and 4) theta phase encodes information important to establishing syllable perception (Kayser et al. 2012; Peelle and Davis 2012) and identity (ten Oever and Sack 2015). Based on those facts, we posit that theta phase synchrony for syllable processing will decrease over development and that attention-driven, experience-dependent effects may confer a syllabic processing advantage in the AEx group, and thus, the amount of theta phase synchrony at 7 and 9 months might be linked to later linguistic abilities at 12 and 18 months of age. Lastly, based on what we know about differing maturational trajectories, in particular, that the left temporal cortex matures faster than the right temporal area (Paus et al. 2001; Pujol et al. 2006; Deoni et al. 2011), we expected to see a larger and earlier modulatory effect with a more mature pattern emerging in LAC.

Materials and Methods

Participants

Our sample included 110 typically developing infants, who were participants in a mixed design, longitudinal, and cross-sectional developmental study; a subset of these infants were also included in published reports at 4, 7, and 9 months of age (Musacchia et al. 2013; Benasich et al. 2014; Musacchia et al. 2017; Ortiz-Mantilla et al. 2019). The 18 infants who formed the longitudinal group for this study were recruited at 4 months of age and were invited to participate in an interactive auditory experience (AEx group). All infants in the AEx group were followed across 7, 9, 12 and 18 months, although 2 participants did not return for the 18-month follow-up visit. In addition, to control for maturation effects, 92 cross-sectional infants were recruited, but 10 were later excluded due to data quality issues, thus 82 infants comprised the final cross-sectional sample. Those 82 infants, 19 at 7 months, 22 at 9 months, 20 at 12 months, and 21 infants at 18 months of age comprised the cross-sectional NC groups (NC07, NC09, NC12, and NC18). All 100 participants in the final sample (47 females/53 males; AEx: 8/10; NC07: 9/10; NC09: 9/13; NC12: 11/9; NC18: 10/11) had uneventful prenatal and perinatal circumstances and were born healthy, full-term (mean: 39.6 weeks; standard deviation [SD]: 1.1; range: 37–42 weeks), normal birth weight (mean: 3433 gr; SD: 461; range: 2551–5046 gr) into monolingual English families. Mothers of all infants have completed high school (data for six mothers was not available), and seven of the mothers did not have a college education (mean post-high school years of education: 4.84 years; SD: 2.42; range: 0–11 years). Infants came from urban and suburban communities and had no family history of specific language impairment, autism, hearing loss, no repeated episodes of otitis media, or other medical or neurological disorders and had passed the newborn hearing screening. Parents received compensation for their time and infants were given a toy after the visit. The study was conducted in accordance with the Declaration of Helsinki, and informed consent that was approved by our University Human Subjects Institutional Review Board was obtained before inclusion in the study.

Procedure

Behavioral Protocol for the IAE
After their initial evaluation at 4 months of age, infants in the AEx group visited the lab once a week for 6 consecutive weeks between 4.5 and 6 months of age (mean: 4.7 [SD: 0.3]—5.9 [0.3]

months). Via a go/no-go looking task, AEx infants learned an association between an auditory stimulus and the onset of a video reward (Nawyn et al. 2007). The procedure followed three phases: familiarization, training, and baseline. During all phases, a standard stimulus was repeatedly presented, interspersed with an experimenter-initiated target stimulus paired with a video reward presentation. Three different pairs of acoustic stimuli were used in the training as follows: weeks 1 and 2, complex tones (standard: 800–800 Hz; target: 800–1200 Hz); weeks 3 and 4, bandpass noise (standard: 400–1900 Hz and 400–1900 Hz; target: 400–1900 Hz and 800–1900 Hz); and weeks 5 and 6: simple sweeps (standard: 1600–1200 Hz and 1600–1200 Hz; target: 1600–1200 Hz and 1200–1600 Hz). Infants were conditioned to direct their gaze to a specified region on a computer screen in response to a "go trial" (i.e., target stimulus). The reward video was initiated automatically when the infant looked toward the reward area. Looking to the reward video is a measure of the infant's operantly conditioned auditory detection response that indexes whether the infant discriminated the change in the auditory stimuli. The sound stimuli were presented at varying ISIs, depending on the phase of the session, using an up-down staircase procedure (Trehub et al. 1986) and were continued for 7–9 min each session, or until the child fatigued. The ISI for each block of stimuli was increased or decreased according to infant performance. The IAE training protocol was designed to focus on inducing plasticity in auditory processing by supporting RAP abilities, optimizing acoustic mapping and discrimination of spectro-temporal cues, and enhancing attention on key prelinguistic cues that had relevance for subsequent linguistic mapping. A more comprehensive description of the training protocol can be found in Benasich et al. (2014).

Event-Related Potentials

Stimuli. The stimuli used for the ERP were computer-generated consonant-vowel syllables that differed in VOT. The standard (STD) stimulus was the syllable /da/ (VOT = 0 ms) and the deviant (DEV) stimulus was /ta/ (VOT = 40 ms). Both syllables were comprised of stop consonants, voiced for the STD and voiceless for the DEV, and followed by a 60 ms steady-state vowel. The duration of each syllable was 100 ms; the fundamental frequency was 120 Hz; and three formants F1, F2, and F3 were 750, 1200, and 2500 Hz, respectively. The stimulus onset-to-onset interval was 1020 ms, and the offset-to-onset interval was 920 ms. The stimuli were presented in a pseudo-randomized passive oddball paradigm by using a block design comprised of 85% STD and 15% DEV, with at least 3 and no more than 12 STD presented before each DEV. Stimuli were presented binaurally using E-Prime software (Psychology Software Tools) in a sound-attenuated free field environment at 75 dB SPL.

EEG recording and ERP data processing. Dense array EEG/ERP recordings were acquired at each age visit (7, 9, 12, and 18 months), while participants were seated on their parent's lap, watching a silent movie, or were entertained with silent toys to keep them calm and engaged (Musacchia et al. 2015). EEG was recorded from 125-channel EGI sensor nets (Electrical Geodesic, Inc.) for most of the participants (AEx: 18; NC07: 16; NC09: 17; NC12: 15; NC18: 13), while for a subset, it was recorded with 125-channel EGI sensor Hydrocel nets (NC07: 3; NC09: 5; NC12: 5; NC18: 8). The vertex channel was used as on-line reference, with a sample rate of 250 Hz, and 0.1 Hz high pass and 100 Hz low pass filters. Artifact correction of eye movements was conducted on the raw data using an automatic correction algorithm based

on principal component analysis (PCA) method, a standard routine included in the Brain Electrical Source Analysis (BESA GmbH) 6.1 software. For ERP processing, EEGs were band-pass-filtered off-line at 1–15 Hz and were rereferenced to an average reference. EEG/ERP data were subsequently segmented into epochs according to stimulus type (STD, DEV) with 300 ms prestimulus, 1020 ms poststimulus time, and 100 ms before stimulus onset was used as the baseline. Epochs with signals exceeding ±300 μV from the baseline were excluded. After data cleaning, a total of 152 recordings were accepted for analysis. Traditionally, the mismatch response (DEV − STD) has been used to investigate auditory discrimination (Peltola et al. 2003) but has also been considered to be a measure of central sound representation accuracy (Näätänen and Alho 1997; Näätänen and Winkler 1999). The mismatch response depends on two fundamental elements: 1) the presence of a short-term memory trace in the auditory cortex representing the regularity of a repetitive stimulus (STD) and 2) a change-detection (discrimination process) of a sporadic stimulus (DEV), incongruent with the memory representation of the preceding stimulus (for a review, see Näätänen et al. 2007). The mismatch is typically manifested in the deviant ERP as an amplitude increase in the P2 and/or N2 peak components (Näätänen 1992). The response to the deviant implicitly carries the change detection process, and thus, is a more focused surrogate of the mismatch response. Consequently, in this paper, effects of auditory experience on syllable representation were examined using the STD stimulus and on syllable discrimination using the DEV stimulus. The STD syllable was presented 566 times assuring a good signal-to-noise ratio, and thus, an increased probability of having a more reliable syllable representation, while the DEV syllable was presented 100 times. A minimum of 70% STD (396) and DEV (70) artifact-free epochs per infant were required for inclusion in ERP averaging. The groups did not differ in the number of STD and DEV trials accepted for ERP analysis at any of the ages examined. The mean averages of number of trials across the four age points were: STD: 482 for AEx, 485 for NC; DEV: 85 for AEx, 86 for NC.

Source Localization of ERP Generators

Source localization is a technique used to identify the loci of the neural activation registered at the scalp surface. This approach enables the high temporal resolution provided by EEG/ERP to be combined with structural images, thus allowing a closer spatial approximation as to where in the brain neural responses are being generated. To localize source generators of the response to the STD and DEV syllables, EEG/ERPs were mapped onto a 0- to 18-month head template provided by BESA Research 6.1 software. Based on PCA, the first positive peak (P1) for the STD and DEV responses was identified in the grand average ERP of each group at each age. A discrete dipole source model (Scherg and Von Cramon 1985) using a four-shell ellipsoidal head model was applied to the P1 within each STD and DEV condition for source modeling. Following a protocol developed for infant source modeling (Hämäläinen et al. 2011; Ortiz-Mantilla et al. 2012), a time window of ±20 ms around the peak was used for dipole fitting. A two-dipole model identified sources of activation in left (LAC) and right (RAC) auditory cortices. The source montage generated during discrete dipole fitting in the grand average of each condition (STD, DEV) and group (AEx, NC) at each age (7, 9, 12, and 18 months) was saved for further use during time-frequency analysis.

Time-Frequency Analyses in Source Space

Spectro-temporal changes in event-related oscillations during STD and DEV processing were examined in source space as follows: The previously saved two-source montage (that functions as a fixed spatial filter) from the grand average source model for each condition and age was applied to the raw 125-channel recording of each individual in the corresponding group to transform the continuous EEG into two-channel source space (Hoechstetter et al. 2004). To control for low-frequency activity while at the same time preserving as much of the frequency information as possible, only a low cutoff of 0.5–1 Hz was applied to the raw EEG data (Paul-Jordanov et al. 2017). A complex demodulation method with 1-Hz wide frequency bins and 50-ms time resolution, from −300 to 1020 ms in the range of 2–90 Hz was used next for decomposing the single-trial EEG data into time-frequency representations (Paul-Jordanov et al. 2017). As theta oscillations are particularly suited to processing syllabic information, for the purposes of this paper, event-related changes in the amount of intertrial phase synchrony (Tallon-Baudry et al. 1996; Hari and Salmelin 1997; Tallon-Baudry and Bertrand 1999) were investigated within the lower-frequency bands by using ITPL. ITPL is a measure of phase synchrony that determines the stability of the phase alignment across trials. ITPL is reported between 0 and 1 in which 0 represents random phase alignment and 1 represents total phase alignment. ITPL individual results, generated for the STD and DEV in the left and right auditory sources for each group at each age, were exported to MATLAB R2020a (MathWorks) for plotting graphics across subjects.

Standardized Behavioral Measures

Cognitive assessment. The Cognitive Subscale from the Bayley Scales of Infant and Toddler Development, Third Edition (BSID-III; Bayley 2006) was administered at each age. The Cognitive composite standard score is reported for this subscale (mean: 100, SD: 15).

Language assessment. The Preschool Language Scale-4th ed. (PLS-4; Zimmerman et al. 2002) that assesses receptive (auditory comprehension) and expressive (expressive communication) language skills in children from birth to 6 years and 11 months of age was also administered at each age. The test yields standard scores (mean: 100, SD: 15), percentile ranks, and age scores for the subscales as well as a total language score. For this study, standard scores for the Auditory Comprehension and Expressive Communication subscales were used.

Statistical Analyses

Descriptive statistics and demographic analyses were conducted using analysis of variance (ANOVAs) and Chi Square in SPSS Statistics 26 (SPSS, Inc.) software. All time-frequency analyses were conducted in BESA Statistics 2.0 (BESA GmbH) software. For comparisons between two groups, BESA Statistics uses a parameter-free permutation testing based on Student's *t*-test, and for within and between comparisons in which more than two comparisons are estimated, BESA Statistics uses parameter-free permutation testing based on ANOVA (Bullmore et al. 1999; Ernst 2004; Maris and Oostenveld 2007). In BESA Statistics, correction for multiple comparisons in post hoc tests is addressed by means of a Bonferroni-Holm correction. Furthermore, results are considered to be corrected for multiple comparisons as only those clusters, which have higher cluster values than 95% of all clusters derived by random permutation of data, will be identified. (For a more detailed description of these methods, refer to BESA Statistics website and user manual,

Bornfleth et al. 2020, http://www.besa.de/products/besa-statistics/besa-statistics-overview/; https://www.besa.de/wp-content/uploads/2014/05/BESA-Statistics-2.1-User-Manual.pdf).

In this study, there were no predefined clusters as after 1000 permutations, BESA Statistics automatically identifies time-frequency regions (clusters) with significant changes in the magnitude of ITPL within/between ages/groups. To determine short- and long-term effects of the IAE on syllable processing, intertrial phase synchrony in the theta range was examined. For the AEx longitudinal analysis, a 2×4 (source [LAC, RAC] by age [07, 09, 12, 18]) within-group permutation ANOVA was conducted, followed by post hoc permutation Scheffe's test to determine which pairwise comparisons were responsible for the age main effect. Effects of age were also examined cross-sectionally between the NC groups using a 2×4 (source [LAC, RAC] by group [NC07, NC09, NC12, NC18]) permutation ANOVA, followed by post hoc permutation Scheffe's test to determine which pairwise comparisons were responsible for the group main effect. Group differences in theta ITPL between the AEx and the age-matched NC group were examined at each age using permutation *t*-tests. To assess the association between intertrial phase synchrony and linguistic outcome, cluster permutation testing based on Pearson correlation analysis (Freedman and Lane 1983; O'Gorman 2012) was conducted in BESA Statistics 2.0. The *p* statistics values reported in this study correspond to the *P* value of the cluster for the correlation after permutation testing.

Results

Preliminary analysis did not find any differences in gestational age ($F(4, 95) = 1.66$, $P = 0.17$), birth weight ($F(4, 95) = 1.26$, $P = 0.29$), sex ($X^2 = 0.89$, $P = 0.93$), or maternal education ($F(4, 89) = 1.32$, $P = 0.27$), among the groups. Standardized behavioral assessments showed that the AEx and NC groups had cognitive abilities within the average range at 12 and 18 months of age. However, the NC18 group had significantly lower cognitive scores than the AEx group at 18 months ($F(1, 35) = 6.90$, $P = 0.013$), a result not found at 12 months ($F(1, 37) = 0.22$, $P = 0.643$). This difference could be explained by the fact that NC12 and NC18 are two different groups of children, as the study design required that the NCs be cross-sectional cohorts. Expressive and receptive language skills were found to be within the normative range at 12 and 18 months with no differences observed between the groups (means and SDs are provided in Table 1).

Source Localization of ERP Generators for the P1 Response

ERP responses to the STD and DEV stimuli at each age and group were characterized by a fronto-central positivity, followed by a negative deflection resembling those reported in other infancy studies using syllables differing in VOT (Rivera-Gaxiola et al. 2005; Ortiz-Mantilla et al. 2012, 2013, 2016). Generators of the positive ERP response to STD and DEV were identified from the grand average ERP waveforms and a two-dipole model freely fitted for each group and age. The free dipole fitting procedure placed dipoles in LAC and RAC explaining ~97% and ~96% of the variance for STD and DEV, respectively (Fig. 1, middle). The source waveforms in each group followed the positive–negative pattern observed in the original ERP waveforms, indicating a good model fit to the data (Fig. 1A–D).
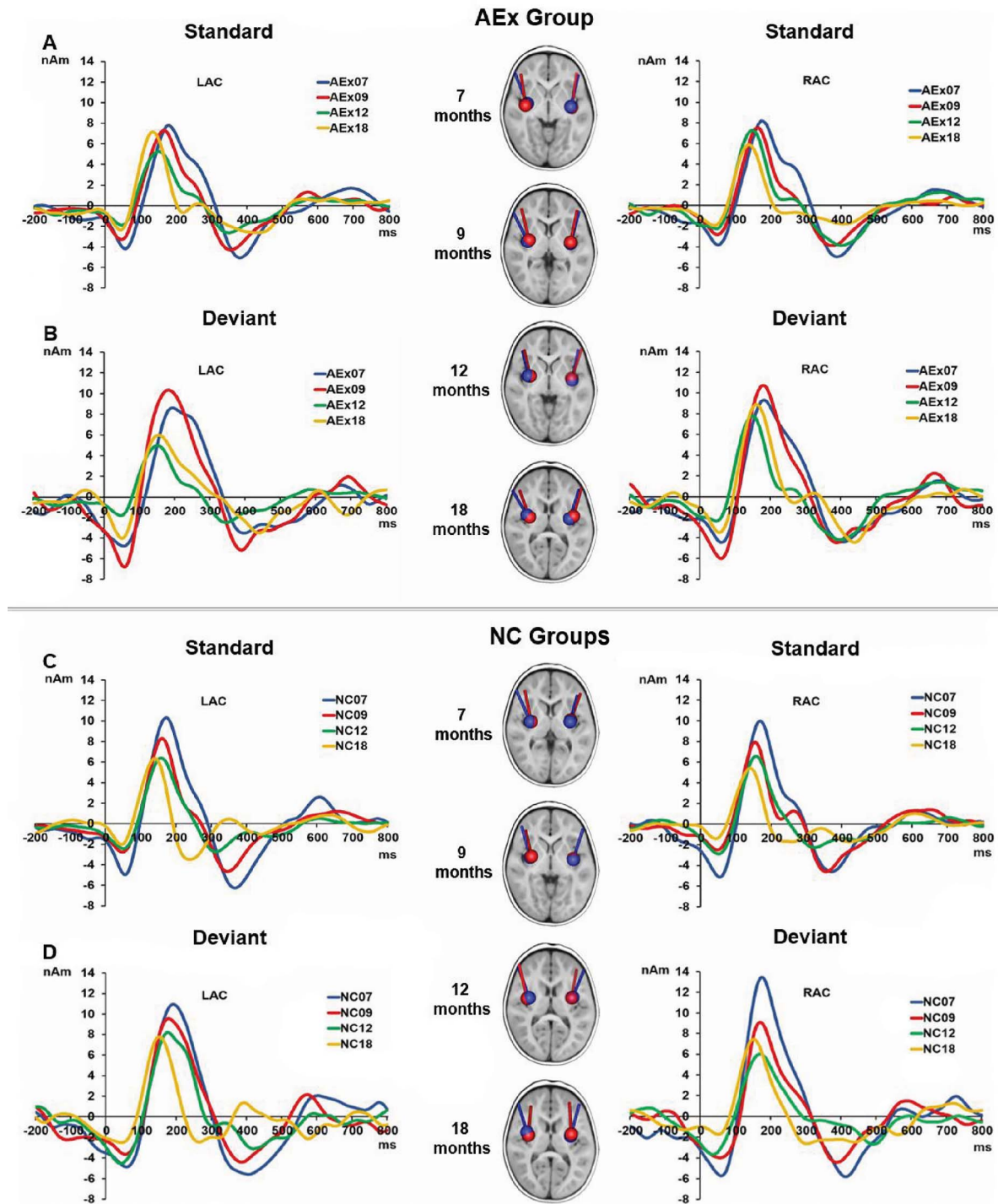
Figure 1. Source waveforms: Above: Grand average source waveforms of the response to the standard (*A*, first row) and deviant (*B*, second row) syllables at LAC and RAC depicted for the longitudinal IAE AEx group. The source waveform at 7 months is shown in blue, at 9 months in red, at 12 months in green, and at 18 months in yellow. Below: Grand average source waveforms of the response to the standard (*C*, first row) and deviant (*D*, second row) syllables at LAC and RAC depicted for the cross-sectional NC groups. The source waveform for the NC07 group is shown in blue, for the NC09 group in red, for the NC12 group in green, and for the NC18 group in yellow. Positivity is plotted up, time is shown in ms on the *x*-axis, and amplitude of the source response is given in nanoampere meters (nAm) on the *y*-axis. Middle: Source localization of the P1 generators is shown in transverse view in a 0–18-month infant brain template. Left and right dipoles in blue represent the response to the standard syllable /da/, and dipoles in red represent the response to the deviant syllable /ta/. Above: Sources for the longitudinal interactive experience (AEx) group at 7, 9, 12, and 18 months of age are displayed. Below: Sources for the cross-sectional, NC (NC07, NC09, NC12, and NC18) groups are displayed.

Table 1 Descriptive statistics (means and SD) for standardized cognitive and language assessments at 12 and 18 months of age

Cognitive and language assessments

Group	Number of participants	Cognitive score mean (SD)	Exp. language score mean (SD)	Rec. language score mean (SD)
AEx12	18	110.0 (9.5)	100.9 (8.8)	96.6 (14.9)
NC12	20	111.5 (10.1)	105.0 (7.6)	93.1 (7.3)
AEx18	16	111.6 (16.4)	105.3 (13.3)	99.4 (20.2)
NC18	21	100.2 (9.7)	104.3 (7.6)	97.3 (12.7)

AEx12, IAE group at 12 months; NC12, NC group cross-sectional age 12 months; AEx18, IAE group at 18 months; NC18, NC group cross-sectional age 18 months; Exp, expressive; Rec, receptive.

Time-Frequency analysis

Measures of ITPL were obtained via data clustering in combination with permutation testing. The P value reported in this series of analyses represents the P value of the permutation already corrected for multiple comparisons. Significant differences in ITPL dynamics during syllable representation were seen in the response to the STD stimulus. Cross-sectional analyses between the NC groups 2×4 (source by group [$n = 82$]) permutation ANOVA showed a group effect ($P = 0.024$) with a decrease in ITPL in older as compared with younger groups in the left source (significant cluster: 250–500 ms, 3–6 Hz). Pairwise comparisons determined that the NC07 group exhibited a larger amount of ITPL than seen in the NC18 group ($P = 0.007$). Likewise, longitudinal analyses within the AEx group [$n = 16$] 2×4 (source by age) permutation ANOVA showed an age effect with a bilateral decrease in ITPL with age (significant cluster: 200–350 ms, 2–5 Hz; LAC: $P = 0.026$; RAC: $P = 0.000$). Pairwise comparisons showed that the significance was driven by a decrease in ITPL from 7 to 12 months of age in LAC ($P = 0.005$) and RAC ($P = 0.001$) sources and from 7 to 18 months in RAC ($P = 0.002$). No significant difference in the amount of ITPL was seen for the DEV (which measures syllable discrimination) across age for either the AEx or NC groups.

Group comparison between the AEx and NC groups found larger amounts of ITPL for the STD in all NC groups as compared with the AEx group except for NC12 at 12 months. Specifically, larger ITPL was found for NC07 in the LAC source at 7 months, NC09 in both LAC and RAC sources at 9 months, and for NC18 group in the RAC source at 18 months (Table 2; Fig. 2). Similarly, for DEV processing, larger amounts of ITPL were displayed for the NC groups as compared with the AEX group. Precisely, larger ITPL was shown for NC07 in both the LAC and RAC sources at 7 months and for NC18 in the RAC source at 18 months (Table 2; Fig. 3). Thus, when compared with NC groups, the modulatory effects of IAE on ITPL were observed in both the short- (7 months) and long-term (18 months) follow-ups. Importantly, this ongoing modulation of ITPL supported more efficient processing of syllabic representation and discrimination in the LAC source at younger ages and in the RAC source at older ages.

Correlations

To examine if the IAE modulatory effect on ITPL at 7 and 9 months impacted language abilities at 12 and 18 months, correlational analyses were conducted in the AEx longitudinal group.

A consistent pattern of negative associations was found between the amount of ITPL for STD and DEV in the LAC source and expressive and receptive language abilities. Specifically, less ITPL for the STD at 7 months was related to better expressive language at 12 ($P = 0.023$) and 18 months ($P = 0.019$), while less ITPL at 9 months was associated with better receptive language at 18 months ($P = 0.033$). On the other hand, less ITPL for the deviant syllable (DEV) at 7 months was correlated with better expressive ($P = 0.025$) language abilities at 18 months, and less ITPL for DEV at 9 months was correlated with better receptive language abilities at both 12 ($P = 0.033$) and 18 ($P = 0.005$) months (Table 3).

Discussion

In this study, we examined the short- and long-term effects of nonspeech auditory training on syllable processing and the relationship between theta phase synchrony modulation and language abilities. Our results demonstrate that infant auditory perceptual abilities are associated with language outcomes and that supporting these emerging perceptual abilities during early development is specifically related to increases in syllabic processing efficiency. Notably, we see here that the plasticity effects of early acoustic experience (IAE) with spectro-temporally modulated nonspeech generalizes to speech and confers a significant processing advantage for syllabic processing well beyond the post-training period. In both LAC and RAC sources, IAE increases the efficiency of syllabic representation as captured by the STD stimulus and of syllabic discrimination as indexed by the DEV stimulus. Modulatory effects of IAE on theta phase synchrony also showed a prominent lateralization pattern across age. Short-term effects of IAE were more evident on the left auditory source, whereas long-term effects were predominantly observed on the right source. Furthermore, the early modulatory effect of IAE on theta phase synchrony in the LAC has a positive impact on later language abilities. Less phase synchrony (a sign of more efficient, automatized processing) during syllable representation at 7 months was associated with better expressive language at 12 and 18 months, while at 9 months, less phase synchrony was related to better receptive language. A similar differential pattern was seen during syllable discrimination with less phase synchrony, while processing the deviant syllable at 7 months associated with better expressive language at 18 months, and less 9-month phase synchrony related to better receptive language at 12 and 18 months.

One of the most important characteristics of the auditory cortex during early development is its high degree of plasticity, which allows rapid retuning to the acoustic changes in the environment (Buonomano and Merzenich 1998; Zhang et al. 2001; Froemke and Jones 2011; Kudo et al. 2020). Exposure-based plasticity facilitates establishment and ongoing modification of cortical representations as infants become familiar with the

Table 2 Group comparison on effects of IAE on theta intertrial phase synchrony in the AEx group as compared with naive age-matched control groups on syllable representation (standard) and syllable discrimination (deviant)

Differences in theta ITPL between the NC and AEx groups

G1 G2 NC AEx	N G1 G2	Stim	Time cluster	Freq. cluster	Source	P	Mean ITPLG1	Mean ITPLG2
NC07 > AEx07	19 18	STD	50–250	4–8	LAC	0.022	0.1369	0.1087
NC09 > AEx09	22 18	STD	50–300	3–10	LAC	0.000	0.1291	0.1016
					RAC	0.001	0.1371	0.1087
NC18 > AEx18	21 16	STD	100–400	2–6	RAC	0.020	0.1241	0.0992
NC07 > AEx07	19 18	DEV	50–250	2–6	LAC	0.016	0.2066	0.1704
					RAC	0.028	0.2221	0.1730
NC18 > AEx18	21 16	DEV	200–450	2–7	RAC	0.014	0.1603	0.1304

G1, group 1; NC, NC group (cross-sectional); G2, group 2; AEx, IAE group (longitudinal); the testing age is given in months; N, number of participants in the group; Stim, stimulus; Mean ITPLG1, mean ITPL for group 1; Mean ITPLG2, mean ITPL for group 2; NC07, NC group at 7 months; NC09, NC group at 9 months; NC18, NC group at 18 months; STD, standard stimulus; DEV, deviant stimulus; time frame for the cluster is given in ms; Freq, frequency range for the cluster is given in Hertz; LAC, source located in left auditory cortex; RAC, source located in right auditory cortex; p, statistical value for permutation analyses; >, larger than.
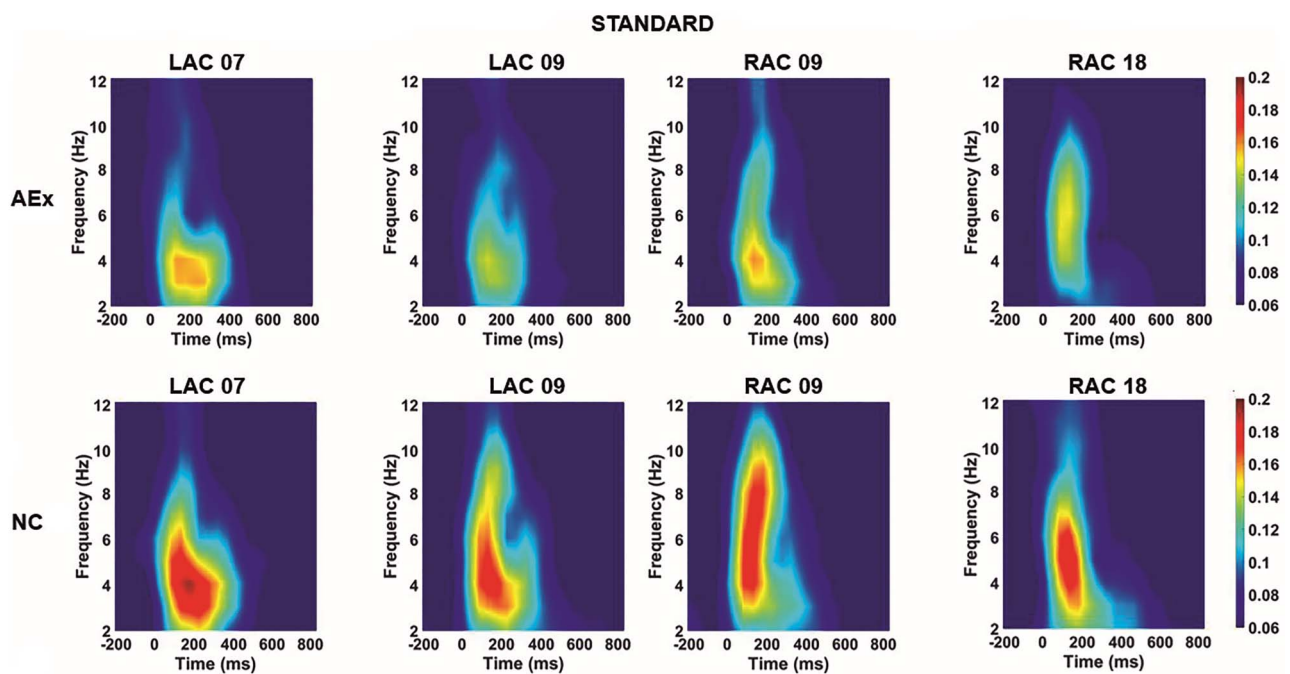
Figure 2. Time-frequency plots displaying significant changes between the groups in intertrial phase synchrony as a response to the standard syllable /da/ in theta frequency band. In the first row, phase synchrony plots are shown for the AEx group, and, in the second row, for the matched NC group. In the first column, AEx responses at 7 months in LAC are compared with NC07 responses; in the second and third columns, AEx responses at 9 months in LAC and right (RAC) sources are compared with corresponding NC09 responses; in the fourth column, AEx responses at 18 months in RAC are compared with NC18 responses. In all cases, phase synchrony was significantly smaller for AEx than matched NC group as reported in Table 2. Time is shown in ms on the x-axis and frequency is shown in Hz on the y-axis.

distinctive characteristics of their linguistic surroundings (Ortiz–Mantilla et al. 2016). In addition to bootstrapping developmental maturation and native language familiarization, an enriched auditory environment also increases the processing efficiency in the infant brain (Benasich et al. 2014; Musacchia et al. 2017; Ortiz-Mantilla et al. 2019). In this study, we demonstrated that experience-dependent effects of prelinguistic interaction with nonspeech between 4.5 and 6 months persist well past 9 months through at least 18 months of age, thus providing a long-term advantage in syllabic processing to the AEx infants exposed to this IAE early in the first year.

The AEx group showed less theta phase synchrony for the standard syllable than NCs at 7 months in the left source, at 9 months in left and right sources, and at 18 months in the right

source. A decrease in theta phase synchrony during processing of speech and nonspeech signals over the first year of life (Ortiz-Mantilla et al. 2016; Hämäläinen et al. 2019) is thought to represent the more competent processing achieved with brain maturation and native language exposure. But, in this study, the decrease in theta phase synchrony exhibited by acoustically trained infants was well above that observed for NC infants (Table 2). This fact implies that, compared with untrained infants, the AEx group allocated less neuronal resources to processing of the STD and suggests better, more specific, and easily accessible syllabic cortical representations and thus increased processing efficiency as a function of training. The IAE modulatory effect on phase synchrony was seen at an earlier time in the left (7 and 9 months) than in the right
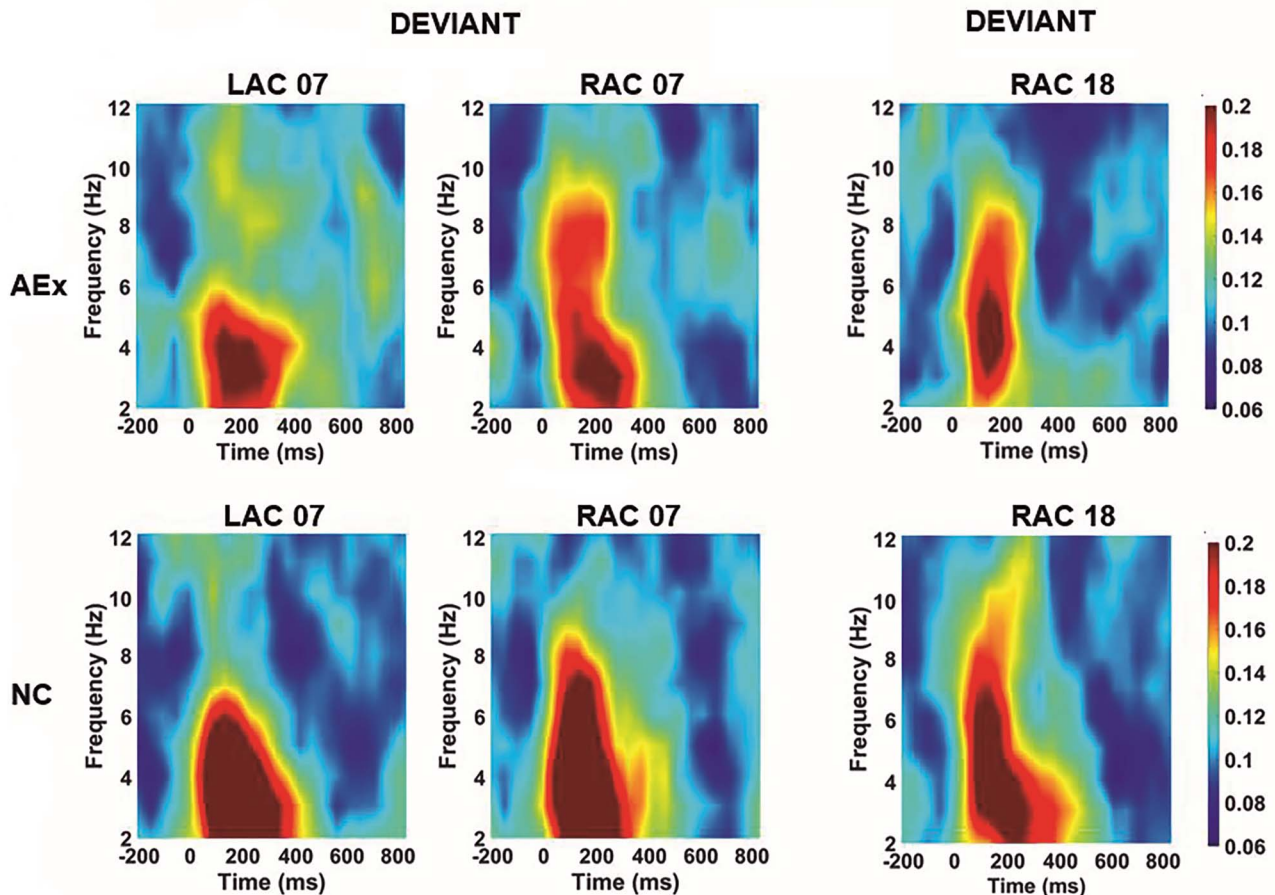
Figure 3. Time-frequency plots displaying significant changes between groups in intertrial phase synchrony as a response to the deviant syllable /ta/ in theta frequency band. In the first row, phase synchrony plots are shown for the AEx group, and, in the second row, for the matched NC group. In the first and second columns, AEx responses at 7 months in LAC and RAC are compared with NC07 responses; in the third column, AEx responses at 18 months in RAC are compared with corresponding NC18 responses. In all cases, phase synchrony was significantly smaller for AEx than matched NC group as reported in Table 2. Time is shown in ms on the x-axis and frequency in Hz on the y-axis.

Table 3 Correlations between early (7 and 9 months) theta phase synchrony and later (12 and 18 months) expressive and receptive language abilities for the AEx group

| ITPL | AEx group correlations | | | | | | | |
	Language	N	Stim	TW	Freq	Source	Direction	P
7 m	Expressive, 12 m	18	STD	50–400	2–5	LAC	Negative	0.023
7 m	Expressive, 18 m	16	STD	200–500	2–5	LAC	Negative	0.019
9 m	Receptive, 18 m	16	STD	350–450	2–4	LAC	Negative	0.033
7 m	Expressive, 18 m	16	DEV	300–500	2–4	LAC	Negative	0.025
9 m	Receptive, 12 m	18	DEV	150–300	2–4	LAC	Negative	0.033
9 m	Receptive, 18 m	16	DEV	200–400	2–4	LAC	Negative	0.005

AEx, IAE group; N, number of participants; STD, standard; DEV, deviant; TW, time window for the cluster is given in ms; m, months of age; Freq, frequency range for the cluster is given in Hz; LAC, left auditory cortex; p, statistical value for permutation analyses of the correlation.

auditory cortex (9 and 18 months) in line with studies during the first year of life that have reported a left-lateralized structural (Witelson and Pallie 1973; Gilmore et al. 2007; Choe et al. 2013; Adibpour et al. 2020) and functional pattern (Dehaene-Lambertz et al. 2002; Gervain et al. 2008; Minagawa-Kawai et al. 2011; Mahmoudzadeh et al. 2013; Ortiz-Mantilla et al. 2013, 2016) and greater myelination (Kinney et al. 1988; Paus et al. 2001; Pujol et al. 2006; Deoni et al. 2011) of the left hemisphere

auditory-related areas. Greater increases in myelination in the left as compared with right temporal areas suggest an asymmetrical maturational sequence in which the left cortical structures mature faster than the right (Deoni et al. 2011). These early left/right disparities in myelination rates might well have functional implications for speech processing. A more myelinated LAC may facilitate faster, synchronized neural conduction, thus providing optimal support for processing of

the rapid temporal and spectral transitions that characterize syllabic content (Zatorre and Belin 2001). Therefore, we hypothesized that by bootstrapping a normative developmental process, IAE might facilitate expeditious maturation of the LAC. However, further studies are needed to test this premise.

At 18 months, the AEx group also exhibited a larger decrease in phase synchrony than NCs, however, the effect was seen in the RAC source. Although as discussed above, short-term modulatory effects of IAE on ITPL were seen in the left auditory areas, known to mature faster than those on the right, these longer-term effects at 18 months were seen predominantly in right auditory areas, known to mature at a slower pace than those on the left. Thus, this pattern of IAE modulatory effects on theta phase synchrony seems to capture the known left first and then right developmental maturational patterns. In addition to these effects for the STD stimulus, IAE also modulated theta phase synchrony for the DEV stimulus. Less phase synchrony suggested more efficient syllabic discrimination was enabled for AEx than NCs at 7 months in LAC and RAC at 18 months in the RAC.

A more widespread form of training in the auditory domain is musical training. In infants, musical training has been shown to benefit among others, processing of auditory patterns (Trainor 1997), rhythm, including beat and meter perception (Gerry et al. 2010; Cirelli et al. 2016), and development of sensitivity to tonality (Gerry et al. 2012; Trainor et al. 2012). Although music and language have a remarkable resemblance in that they both rely on overlapping neural networks, depend on experience and learning, and involve interaction between spectral and temporal domains (Trainor and Hannon 2013; White et al. 2013), musical training in infants predominantly favors processing of suprasegmental elements. Conversely, by using a discrimination paradigm based on a frequency change, the IAE training presented in this study was specifically designed to: 1) support RAP abilities in the tens of ms range (critical to processing speech at the segmental level); 2) optimize acoustic mapping and acoustic discrimination of spectro-temporal cues in the period in which the infant brain is highly sensitive to environmental sensory experience; and 3) enhance attention to salient acoustic cues known to be important for language (Benasich et al. 2014). Training or environmental experience is associated with functional and/or structural neural changes (for a review, see Galvan 2010), and as shown here, this specific IAE training increased processing efficiency, which was reflected in the smaller magnitude of the theta phase coherence required for processing the syllabic content. Thus, decreases in cortical activity represent a more skilled, automatic processing as a function of training (Galvan 2010).

As reported above, we also found that increases in processing efficiency in the LAC at 7 and 9 months of age was consistently associated with better language outcomes. Interestingly, for both the deviant and the standard syllable, more efficient phase synchrony at 7 months was related to expressive language, while at 9 months, it was linked to receptive language. Based on previous knowledge, we can try to elucidate this age-sensitive differential pattern of associations. It is known that as infants acquire language, they must conduct fine-grained prelexical acoustic analysis of the speech signal in the tens of ms range, which allows the infant to learn about the sound structure of consonants and vowels. This is essential for further decoding and discrimination of phonemic information and is crucial for the establishment of cortical phonemic representations (Eilers et al. 1981; Aslin 1989; Werker and Tees 2005). The ability of EEG/ERPs measured

early in development to link to later language outcomes has been widely reported. ERP studies have shown that efficient early acoustic processing abilities in typically developing infants and in infants at high familial risk for developmental language disorders are robustly related to language proficiency as well as later reading (Tsao et al. 2004; Guttorm et al. 2005; Benasich et al. 2006; Leppänen et al. 2010; Choudhury and Benasich 2011; van Zuijen et al. 2012; Hämäläinen et al. 2013; Maitre et al. 2013; Benasich et al. 2016; Cantiani et al. 2016; Lohvansuu et al. 2018). When these basic steps are achieved, infants then progress to acquisition and linking of lexical/semantic and phonemic knowledge by learning words (Greenberg 1998; Swingley 2009). Accordingly, given this developmental timeline, we propose that, at 7 months, infants are focused on decoding the spectral and temporal information contained in native phonemes, a requisite for cortical mapping, and thus, important for emerging expressive language. At 9 months, infants are beginning to learn some words and their meanings; therefore, the association between more efficient phase synchrony and receptive language might be related to the emergence of semantic processing which is required for expanding receptive language. Oscillatory measures of theta spectral power and phase synchrony in infants processing nonspeech contrasts at 6 months of age were also recently found to be associated with expressive language (Cantiani et al. 2019). Aligned with those studies, the consistent negative correlation found in this study, between the magnitude of oscillatory phase synchrony in the theta range and later language abilities in the AEx group, further supports the premise that more efficient infant processing of syllabic information in the LAC may be a biomarker that predicts higher scores in the expressive and receptive language measures at later ages. It seems that for native language acquisition, the left-lateralized auditory processing pattern is a more mature pattern; consequently, efforts oriented to reinforcing leftward functional proficiency may be critical when attempting to improve atypical language acquisition, often seen in groups at risk for developmental language disorders.

Summary

Results from this study demonstrate the enduring effects of IAE on syllable processing. Those effects were characterized by increases in processing efficiency, which followed a lateralization pattern in accordance with known developmental left-/right asymmetries. Short-term effects of acoustic training on theta phase synchrony during processing of a syllable contrast were predominantly observed in the LAC, whereas longer-term effects were mostly reflected in the RAC. Furthermore, early modulatory effects on left auditory processing were linked to language outcome. At 7 and 9 months, for both syllables that differed in VOT, less theta phase synchrony in the LAC was related to better expressive and receptive language abilities at 12 and 18 months. To our knowledge, this is the first time that the magnitude of the intertrial phase synchrony to a syllable contrast has been examined in infants who had participated in an active acoustic training. Furthermore, this is also the first time that phase synchrony efficiency at 7 and 9 months has been shown to be differentially associated with later language abilities: at 7 months with expressive and at 9 months with receptive language. The results presented here are exciting as they show that use of an IAE can bootstrap a naturally unfolding process, improving syllabic processing efficiency even in typically developing infants with no demonstrated familial risk

for developmental language disorders. For children who are at higher risk for disorders, such as specific language impairment, dyslexia, and autism, the possibility of having an infant-friendly intervention protocol available, which facilitates syllabic processing early in development when acoustic maps are being established, seems even more appealing. Encouraged by these results, we recently began to administer the IAE to infants at familial risk for developmental language disorders. By intervening with infants at a preverbal stage when the foundations of language are being laid down, we hope to be able to ameliorate or perhaps even prevent the language deficits that frequently accompany those disorders, providing at-risk children with the opportunity to attain normative language skills.

Notes

Funding

References

Adibpour P, Lebenberg J, Kabdebon C, Dehaene-Lambertz G, Dubois J. 2020. Anatomo-functional correlates of auditory development in infancy. *Dev Cogn Neurosci*. 42:100752.

Ahveninen J, Hämäläinen M, Jääskeläinen IP, Ahlfors SP, Huang S, Lin FH, Raij T, Sams M, Vasios CE, Belliveau JW. 2011. Attention-driven auditory cortex short-term plasticity helps segregate relevant sounds from noise. *Proc Natl Acad Sci USA*. 108:4182–4187.

Arnal LH, Giraud AL. 2012. Cortical oscillations and sensory predictions. *Trends Cogn Sci*. 16:390–398.

Aslin RN. 1989. Discrimination of frequency transitions by human infants. *J Acoust Soc Am*. 86:582–590.

Bayley N. 2006. *Bayley Scales of Infant and Toddler Development, third edition (Bayley-III)*. San Antonio (TX): Harcourt Assessment.

Benasich AA, Thomas JJ, Choudhury N, Leppänen PHT. 2002. The importance of rapid auditory processing abilities to early language development: evidence from converging methodologies. *Dev Psychobiol*. 40:278–292.

Benasich AA, Choudhury N, Friedman JT, Realpe-Bonilla T, Chojnowska C, Gou Z. 2006. The infant as a prelinguistic model for language learning impairments: predicting from event-related potentials to behavior. *Neuropsychologia*. 44:396–411.

Benasich AA, Choudhury NA, Realpe-Bonilla T, Roesler CP. 2014. Plasticity in developing brain: active auditory exposure impacts prelinguistic acoustic mapping. *J Neurosci*. 34:13349–13363.

Benasich AA, Choudhury N, Peters S, Ortiz-Mantilla S. 2016. Early sreening and intervention of infants at high risk for developmental language disorders. The role of nonlinguistic rapid auditory processing. Perspectives on language and literacy. The official publication for the International Dyslexia Association (IDA). Winter issue, 24–28.

Bertoncini J, Mehler J. 1981. Syllables as units in infant speech perception. *Inf Behav Dev*. 4:247–260.

Bornfleth H, Cho JH, Spangler R. 2020. BESA Statistics 2.1 User Manual. https://www.besa.de/wp-content/uploads/2014/05/BESA-Statistics-2.1-User-Manual.pdf (accessed on July 27, 2021).

Bosseler AN, Taulu S, Pihko E, Mäkelä JP, Imada T, Ahonen A, Kuhl PK. 2013. Theta brain rhythms index perceptual narrowing in infant speech perception. *Front Psychol*. 4:690.

Bullmore ET, Suckling J, Overmeyer S, Rabe-Hesketh S, Taylor E, Brammer MJ. 1999. Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. *IEEE Trans Med Imaging*. 18(1):32–42.

Buonomano DV, Merzenich MM. 1998. Cortical plasticity: from synapses to maps. *Annu Rev Neurosci*. 21:149–186.

Cantiani C, Riva V, Piazza C, Bettoni R, Molteni M, Choudhury N, Marino C, Benasich AA. 2016. Auditory discrimination predicts linguistic outcome in Italian infants with and without familial risk for language learning impairment. *Dev Cog Neurosci*. 20:23–34.

Cantiani C, Ortiz-Mantilla S, Riva V, Piazza C, Bettoni R, Musacchia G, Molteni M, Marino C, Benasich AA. 2019. Reduced left-lateralized pattern of event-related EEG oscillations in infants at familial risk for language and learning impairment. *NeuroImage Clin*. 22:101778.

Choe M-S, Ortiz-Mantilla S, Makris N, Gregas M, Bacic J, Haehn D, Kennedy D, Pienaar R, Caviness VS, Benasich AA, *et al*. 2013. Regional infant brain development: an MRI-based morphometric analysis in 3 to 13 month olds. *Cer Cor*. 23: 2100–2117.

Choudhury N, Benasich AA. 2011. Maturation of auditory evoked potentials from 6 to 48 months: prediction to 3 and 4 year language and cognitive abilities. *Clin Neurophysiol*. 122: 320–338.

Cirelli LK, Spinelli C, Nozaradan S, Trainor LJ. 2016. Measuring neural entrainment to beat and meter in infants: effects of music background. *Front Neurosci*. 10:229. https://doi.org/10.3389/fnins.2016.00229.

Dehaene-Lambertz G, Dehaene S, Hertz-Pannier L. 2002. Functional neuroimaging of speech perception in infants. *Science*. 298:2013–2015.

Deoni SCL, Mercure E, Blasi A, Gasston D, Thomson A, Johnson M, Williams SCR, Murphy DGM. 2011. Mapping infant brain myelination with magnetic resonance imaging. *J Neurosci*. 31:784–791.

Ding N, Simon JZ. 2014. Cortical entrainment to continuous speech: functional roles and interpretations. *Fr Hum Neurosci*. 8:311.

Doelling KB, Arnal LH,Guitza O, Poeppel D.. 2014. Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage*. 85:761–768.

Eilers RE, Morse PA, Gavin WJ, Oller DK. 1981. Discrimination of voice onset time in infancy. *J Acoust Soc Am*. 70: 955–965.

Eimas PD, Siqueland ER, Jusczyk P, Vigorito J. 1971. Speech perception in infants . *Science*. 171:303–306.

Eimas PD. 1999. Segmental and syllabic representations in the perception of speech by young infants. *J Acoust Soc Am*. 105:1901–1911.

Ernst MD. 2004. Permutation methods: a basis for exact inference. *Stat Sci*. 19(4):676–685.

Freedman D, Lane D. 1983. A nonstochastic interpretation of reported significance levels. *J Bus Econ Stat*. 1(4):292–298.

Eulitz C, Lahiri A. 2004. Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition. *J Cogn Neurosci*. 16:577–583.

Froemke RC, Jones BJ. 2011. Development of auditory cortical synaptic receptive fields. *Neurosci Biobehav Rev*. 35:2105–2113.

Galvan A. 2010. Neural plasticity of development and learning. *Hum Brain Mapp*. 31:879–890.

Gerry DW, Faux AL, Trainor LJ. 2010. Effects of Kindermusik training on infants' rhythmic enculturation. *Dev Sci*. 13:545–551.

Gerry D, Unrau A, Trainor LJ. 2012. Active music classes in infancy enhance musical, communicative and social development. *Dev Sci*. 15:398–407.

Gervain J, Macagno F, Cogoi S, Peña M, Mehler J. 2008. The neonate brain detects speech structure. *Proc Natl Acad Sci USA*. 105:14222–14227.

Ghitza O. 2013. The theta-syllable: a unit of speech information defined by cortical function. *Front Psychol*. 4:138. 10.3389/fpsyg.2013.00138.

Gilmore JH, Lin W, Prastawa MW, Looney CB, Vetsa YSK, Knickmeyer RC, Evans DD, Smith JK, Hamer RM, Lieberman JA, *et al.* 2007. Regional gray matter growth, sexual dimorphism, and cerebral asymmetry in the neonatal brain. *J Neurosci*. 27:1255–1260.

Giraud AL, Poeppel D. 2012. Cortical oscillations and speech processing: emerging computational principles and operations. *Nat Neurosci*. 15:511–517.

Goswami U. 2011. A temporal sampling framework for developmental dyslexia. *Trends Cogn Sci*. 15:3–10.

Greenberg S, Kingsbury BED. 1997. The modulation spectrogram: in pursuit of an invariant representation of speech. Proceedings of the ICASSP-97 International Conference on Acoustics, Speech and Signal Processing, Munich, April, 1997. Vol. 3. Institute of Electrical and Electronics Engineers, p.1647–1650. doi: 10.1109/ICASSP.1997.598826.

Greenberg S. 1998. A syllable-centric framework for the evolution of spoken language. *Beh Brain Sci*. 21:499–546.

Guttorm TK, Leppänen PHT, Poikkeus AM, Eklund KM, Lyytinen P, Lyytinen H. 2005. Brain event-related potentials (ERPs) measured at birth predict later language development in children with and without familial risk for dyslexia. *Cortex*. 41: 291–303.

Hagoort P, Poeppel D. 2013. The Infrastructure of the Language-Ready Brain. In: Arbib MA, editor. *Language, music, and the brain: a mysterious relationship*. Cambridge (MA): MIT Press, pp. 233–255.

Hämäläinen JA, Ortiz-Mantilla S, Benasich AA. 2011. Source localization of event-related potentials to pitch change mapped onto age-appropriate MRIs at 6 months of age. *NeuroImage*. 54:1910–1918.

Hämäläinen JA, Salminen HK, Leppänen PHT. 2013. Basic auditory processing deficits in dyslexia: systematic review of the behavioral and event-related potential/field evidence. *J Learn Disabil*. 46:413–427.

Hämäläinen JA, Ortiz-Mantilla S, Benasich AA. 2019. Change detection to tone pairs during the first year of life—predictive longitudinal relationships for EEG-based source and time-frequency measures. *NeuroImage*. 198:83–92.

Hari R, Salmelin R. 1997. Human cortical oscillations: a neuromagnetic view through the skull. *Trends Neurosci*. 20:44–49.

Hickok G, Poeppel D. 2000. Towards a functional neuroanatomy of speech perception. *Trends Cog Sci*. 4:131–138.

Hickok G, Poeppel D. 2007. The cortical organization of speech processing. *Nat Rev Neurosci*. 8:393–402.

Hoechstetter K, Bornfleth H, Weckesser D, Ille N, Berg P, Scherg M. 2004. BESA source coherence: a new method to study cortical oscillatory coupling. *Brain Topogr*. 16:233–238.

Jin Y, Diaz B, Colomer M, Sebastián-Gallés N. 2014. Oscillation encoding of individual differences in speech perception. *PLoS One*. 9:e100901.

Jusczyk PW. 1999. How infants begin to extract words from speech. *Trends Cog Sci*. 3:323–328.

Jusczyk PW, Houston DM, Newsome M. 1999. The beginnings of word segmentation in English-learning infants. *Cogn Psychol*. 39:159–207.

Kalashnikova M, Varguese P, Di Liberto GM, Lalor EC, Burnham D. 2018. Infant-directed speech facilitates seven-month-old infants' cortical tracking of speech. *Sci Rep*. 8:13745.

Kayser C, Ince RAA, Panzeri S. 2012. Analysis of slow (theta) oscillations as a potential temporal reference frame for information coding in sensory cortices. *PLoS Comput Biol*. 8:e1002717.

Kerlin JR, Shahin AJ, Miller LM. 2010. Attentional gain control of ongoing cortical speech representations in a "cocktail party". *J Neurosci*. 30:620–628.

Kikuchi M, Shitamichi K, Yoshimura Y, Ueno S, Remijn GB, Hirosawa T, Munesue T, Tsubokawa T, Haruta Y, Oi M, *et al.* 2011. Lateralized theta wave connectivity and language performance in 2- to 5-year-old children. *J Neurosci*. 31:14984–14988.

Kinney HC, Brody BA, Kloman AS, Gilles FH. 1988. Sequence of central nervous system myelination in human infancy. II. Patterns of myelination in autopsied infants. *J Neuropathol Exp Neurol*. 47:217–234.

Kudo T, Morohashi Y, Yazaki-Sugiyama Y. 2020. Early auditory experience modifies neuronal firing properties in the Zebra Finch auditory cortex. *Front Neural Circuits*. 14:570174. https://doi.org/10.3389/fncir.2020.570174.

Lahiri A, Marslen-Wilson W. 1991. The mental representation of lexical form: a phonological approach to the recognition lexicon. *Cognition*. 38:245–294.

Leong V, Byrne E, Clackson K, Harte N, Lam S, de Barbaro K, Wass S. 2017. Infant's neural oscillatory processing of theta-rate speech patterns exceeds adults'. BioRxiv. https://doi.org/10.1101/108852. February 15, 2017, preprint: not peer reviewed.

Leppänen PHT, Hämäläinen JA, Salminen HK, Eklund KM, Guttorm TK, Lohvansuu K, Puolakanaho A, Lyytinen H. 2010. Newborn brain event-related potentials revealing atypical processing of sound frequency and the subsequent association with later literacy skills in children with familial dyslexia. *Cortex*. 46:1362–1376.

Lohvansuu K, Hämäläinen JA, Ervast L, Lyytinen H, Leppänen PHT. 2018. Longitudinal interactions between brain and cognitive measures on reading development from 6 months to 14 years. *Neuropsychologia*. 108:6–12.

Mahmoudzadeh M, Dehaene-Lambertz G, Fournier M, Kongolo G, Goudjil S, Dubois J, Grebe R, Wallois F. 2013. Syllabic discrimination in premature human infants prior to complete formation of cortical layers. *Proc Natl Acad Sci USA*. 110:4846–4851.

Maitre NL, Lambert WE, Aschner JL, Key AP. 2013. Cortical speech sound differentiation in the neonatal intensive care unit predicts cognitive and language development in the first 2 years of life. *Dev Med Child Neurol*. 55:834–839.

Maris E, Oostenveld R. 2007. Nonparametric statistical testing of EEG-and MEG-data. *J Neurosci Methods*. 164:177–190.

Meyer M, Endedijk HM, van Ede F, Hunnius S. 2019. Theta oscillations in 4-year-olds are sensitive to task engagement

and task demands. *Sci Rep.* 9:6049. https://doi.org/10.1038/s41598-019-42615-x.

Minagawa-Kawai Y, Cristià A, Dupoux E. 2011. Cerebral lateralization and early speech acquisition: a developmental scenario. *Dev Cogn Neurosci.* 1:217–232.

Moucha R, Pandya PK, Engineer ND, Rathbun DL, Kilgard MP. 2005. Background sounds contribute to spectrotemporal plasticity in primary auditory cortex. *Exp Brain Res.* 162:417–427.

Musacchia G, Choudhury NA, Ortiz-Mantilla S, Realpe-Bonilla T, Roesler CP, Benasich AA. 2013. Oscillatory support for rapid frequency change processing in infants. *Neuropsychologia.* 51:2812–2824.

Musacchia G, Ortiz-Mantilla S, Realpe-Bonilla T, Roesler CP, Benasich AA. 2015. Infant auditory processing and event-related brain oscillations. *J Vis Exp.* 101:e52420.

Musacchia G, Ortiz-Mantilla S, Choudhury N, Realpe-Bonilla T, Roesler C, Benasich AA. 2017. Active auditory experience in infancy promotes brain plasticity in theta and gamma oscillations. *Dev Cogn Neurosci.* 26:9–19.

Nacar Garcia L, Guerrero-Mosquera C, Colomer M, Sebastian-Galles N. 2018. Evoked and oscillatory EEG activity differentiates language discrimination in young monolingual and bilingual infants. *Sci Rep.* 8:2770.

Näätänen R. 1992. *Attention and brain function*. Hillsdale (NJ): Lawrence Erlbaum.

Näätänen R, Alho K. 1997. Mismatch negativity—the measure for central sound representation accuracy. *Audiol Neurootol.* 2:341–353.

Näätänen R, Winkler I. 1999. The concept of auditory stimulus representation in cognitive neuroscience. *Psychol Bul.* 125:826–859.

Näätänen R, Paavilainen P, Rinne T, Alho K. 2007. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin Neurophysiol.* 118:2544–2590.

Nawyn J, Roesler C, Realpe-Bonilla T, Choudhury N, Benasich AA. 2007. An operantly conditioned looking task for assessing infant auditory processing ability. Proceedings of the Ninth International ACM SIGACCESS Conference on Computers and Accessibility. ASSETS'07: October 15–17. Tempe, Arizona, p. 147–154.

Obleser J, Herrmann B, Henry MJ. 2012. Neural oscillations in speech: don't be enslaved by the envelope. *Front Hum Neurosci.* 6:250.

O'Gorman TW. 2012. *Adaptive tests of significance using permutations of residuals with R and SAS*. Hoboken (NJ): Wiley.

Ortiz-Mantilla S, Hämäläinen JA, Benasich AA. 2012. Time course of ERP generators to syllables in infants: a source localization study using age-appropriate brain templates. *NeuroImage.* 59:3275–3287.

Ortiz-Mantilla S, Hämäläinen JA, Musacchia G, Benasich AA. 2013. Enhancement of gamma oscillations indicate preferential processing of native over foreign phonemic contrasts in infants. *J Neurosci.* 33:18746–18754.

Ortiz-Mantilla S, Hämäläinen JA, Realpe-Bonilla T, Benasich AA. 2016. Oscillatory dynamics underlying perceptual narrowing of native phoneme mapping from 6 to 12 months of age. *J Neurosci.* 36:12095–12105.

Ortiz-Mantilla S, Realpe-Bonilla T, Benasich AA. 2019. Early interactive acoustic experience with non-speech generalizes to speech and confers a syllabic processing advantage at 9 months. *Cer Cor.* 29:1789–1801.

Paul-Jordanov I, Hoechstetter K, Bornfleth H, Waelkens A, Scherg M. 2017. BESA Research 6.1 User Manual. https://www.besa.de/downloads/quick-guides/Reference (accessed July 27, 2021).

Paus T, Collins DL, Evans AC, Leonard G, Pike B, Zijdenbos A. 2001. Maturation of white matter in the human brain: a review of magnetic resonance studies. *Brain Res Bull.* 54:225–266.

Peelle JE, Davis MH. 2012. Neural oscillations carry speech rhythm through to comprehension. *Front Psychol.* 3:320.

Peelle JE, Gross J, Davis MH. 2013. Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cer Cor.* 23:1378–1387.

Peltola MS, Kujala T, Tuomainen J, Ek M, Aaltonen O, Näätänen R. 2003. Native and foreign vowel discrimination as indexed by the mismatch negativity (MMN) response. *Neurosci Lett.* 352:25–28.

Poeppel D. 2003. The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. *Speech Comm.* 41:245–255.

Pujol J, Soriano-Mas C, Ortiz H, Sebastian-Galles N, Losilla JM, Deus J. 2006. Myelination of language-related areas in the developing brain. *Neurol.* 66:339–343.

Räsänen O, Doyle G, Frank MC. 2018. Pre-linguistic segmentation of speech into syllable-like units. *Cognition.* 171:130–150.

Rivera-Gaxiola M, Silva-Pereyra J, Kuhl PK. 2005. Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Dev Sci.* 8:162–172.

Scherg M, Von Cramon D. 1985. Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model. *Electroencephalogr Clin Neurophysiol.* 62:32–44.

Schneider JM, Abel AD, Ogiela DA, Middleton AE, Maguire MJ. 2016. Developmental differences in beta and theta power during sentence processing. *Dev Cogn NeuroSci.* 19:19–30.

Schneider JM, Maguire MJ. 2019. Developmental differences in the neural correlates supporting semantics and syntax during sentence processing. *Dev Sci.* 22(4):e12782.

Spironelli C, Penolazzi B, Vio C, Angrili A. 2006. Inverted EEG theta lateralization in dyslexic children during phonological processing. *Neuropsychologia.* 44:2814–2821.

Swingley D. 2009. Review: contributions of infant word learning to language development. *Phil Trans R Soc B.* 364:3617–3632.

Tallon-Baudry C, Bertrand O. 1999. Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn Sci.* 3:151–162.

Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J. 1996. Stimulus specificity of phase-locked and non-phase locked 40 Hz visual responses in humans. *J Neurosci.* 16:4240–4249.

ten Oever S, Sack AT. 2015. Oscillatory phase shapes syllable perception. *PNAS.* 112:15833–15837.

Trainor LJ. 1997. Effect of frequency ratio on infants' and adults' discrimination of simultaneous intervals. *J Exp Psychol Hum Percept Perform.* 23:1427–1438.

Trainor LJ, Marie C, Gerry D, Whiskin E, Unrau A. 2012. Becoming musically enculturated: effects of music classes for infants on brain and behavior. *Ann NY Acad Sci.* 1252:129–138.

Trainor LJ, Hannon EE. 2013. Musical development. In: Deutsch D, editor. *The psychology of music.* 3rd ed. Elsevier Academic Press, pp. 423–497. doi: 10.1016/B978-0-12-381460-9.00011-0.

Trehub SE, Bull D, Schneider BA, Morrongiello BA. 1986. PESTI: a procedure for estimating individual thresholds in infant listeners. *Infant Behav Dev.* 9:107–118.

Tsao FM, Liu HM, Kuhl PK. 2004. Speech perception in infancy predicts language development in the second year of life: a longitudinal study. *Child Dev*. 75:1067–1084.

Van Zuijen TL, Plakas A, Maasen BA, Been P, Maurits NM, Krikhaar E, van Driel J, van der Leij A. 2012. Temporal auditory processing at 17 months is associated with preliterate language comprehension and later word reading fluency: an ERP study. *Neurosci Lett*. 528:31–35.

Watson TL, Robbins RA, Best CT. 2014. Infant perceptual development for faces and spoken words: an integrated approach. *Dev Psychobiol*. 56:1454–1481.

Werker JF, Tees RC. 1984. Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behav Dev*. 7:49–63.

Werker JF, Tees RC. 2005. Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Dev Psychobiol*. 46: 233–251.

White EJ, Hutka SA, Williams LJ, Moreno S. 2013. Learning, neural plasticity and sensitive periods: implications for language acquisition, music training and transfer across the lifespan. *Front Syst Neurosci*. 7(90):1–18. doi.org/10.3389/fnsys.2013.00090.

Witelson SF, Pallie W. 1973. Left hemisphere specialization for language in the newborn. Neuroanatomical evidence of asymmetry. *Brain*. 96:641–646.

Zatorre RJ, Belin P. 2001. Spectral and temporal processing in human auditory cortex. *Cer Cor*. 11:946–953.

Zhang LI, Bao S, Merzenich MM. 2001. Persistent and specific influences of early acoustic environments on primary auditory cortex. *Nat Neurosci*. 4:1123–1130.

Zion Golumbic E, Cogan GB, Schroeder CE, Poeppel D. 2013. Visual input enhances selective speech envelope tracking in auditory cortex at "cocktail party". *J Neurosci*. 33:1417–1426.

Zion Golumbic EM, Ding N, Bickel S, Lakatos P, Schevon CA, McKhann GM, Goodman RR, Emerson R, Mehta AD, Simon JZ, et al. 2013. Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". *Neuron*. 77:980–991.

Zimmerman IL, Steiner VG, Pond RE. 2002. *Preschool Language Scale-4. Examiner's manual*. San Antonio (TX): The Psychological Corporation.