


ORIGINAL ARTICLE

Early Interactive Acoustic Experience with Non-speech Generalizes to Speech and Confers a Syllabic Processing Advantage at 9 Months

Silvia Ortiz-Mantilla , Teresa Realpe-Bonilla and April A. Benasich

Center for Molecular & Behavioral Neuroscience, Rutgers University-Newark, 197 University Avenue, Newark, NJ 07102, USA

Address correspondence to Silvia Ortiz-Mantilla, Center for Molecular & Behavioral Neuroscience, Rutgers University-Newark, 197 University Avenue, Newark, NJ 07102, USA. Email: sortizma@newark.rutgers.edu  orcid.org/0000-0002-8814-592X

Abstract

During early development, the infant brain is highly plastic and sensory experiences modulate emerging cortical maps, enhancing processing efficiency as infants set up key linguistic precursors. Early interactive acoustic experience (IAE) with spectrotemporally-modulated non-speech has been shown to facilitate optimal acoustic processing and generalizes to novel non-speech sounds at 7-months-of-age. Here we demonstrate that effects of non-speech IAE endure well beyond the immediate training period and robustly generalize to speech processing. Infants who received non-speech IAE differed at 9-months-of-age from both naïve controls and those with only passive acoustic exposure, demonstrating broad modulation of oscillatory dynamics. For the standard syllable, increased high-gamma (>70 Hz) power within auditory cortices indicates that IAE fosters native speech processing, facilitating establishment of phonemic representations. The higher left beta power seen may reflect increased linking of sensory information and corresponding articulatory patterns, while bilateral decreases in theta power suggest more mature automatized speech processing, as less neuronal resources were allocated to process syllabic information. For the deviant syllable, left-lateralized gamma (<70 Hz) enhancement suggests IAE promotes phonemic-related discrimination abilities. Theta power increases in right auditory cortex, known for favoring slow-rate decoding, implies IAE facilitates the more demanding processing of the sporadic deviant syllable.

Key words: auditory plasticity, development, infants, oscillations, phonemic mapping

Introduction

Early plasticity in the developing brain allows for the formation of functional networks that are gradually fine-tuned and individualized to the infant's surrounding environment—both sensory and experiential. This exquisite plasticity is an inherent characteristic of the infant brain, which allows critical early tasks to be accomplished and unfolds naturally over time as infants set up key linguistic and cognitive precursors. Much evidence has shown that ongoing plasticity permits the

developing brain to adapt and find the most optimal solution to environmental demands (Kilgard et al. 2001; Moucha et al. 2005; Threlkeld et al. 2009; Froemke and Jones 2011).

This plasticity-dependent neural process is achieved via the interplay between early auditory processing abilities, brain maturation and experience, leading young infants, in a hierarchically structured sequence, to assemble the foundations of their native language (Kuhl et al. 2006). From birth and perhaps even before (Mahmoudzadeh et al. 2013; Maitre et al. 2013), the

infant brain scans the surrounding environment to identify salient information, including acoustic cues that signal “this could be language”, as infants learn to recognize and map speech sounds into auditory cortex.

In the first months of life, typically developing infants are able to process speech signals in a universal manner, that is, all linguistic content is equally salient. However, by 12-months-of-age, infants’ perceptual abilities become more specific, and favor processing of native over non-native sounds (Best et al. 1995; Cheour et al. 1998; Rivera-Gaxiola et al. 2005; Kuhl et al. 2006; Ortiz-Mantilla et al. 2016). Linguistic perceptual narrowing, which occurs while infants acquire native speech, is a critical developmental step that supports creation of precise cortical representations of phonemic structure that ultimately allow efficient and automatic speech processing (Aslin et al. 1981; Werker and Tees 2005). Infants’ ability to proficiently perform fine-grained acoustic analysis in the tens of milliseconds range is essential for this process, permitting decoding and discrimination of critical phoneme information and subsequent establishment of acoustic phonemic maps. Importantly, early acoustic processing abilities are known to relate to language proficiency in typically developing children, as well as in infants at high familial risk for developmental language disorders (Benasich 2002; Benasich and Tallal 2002; Tsao et al. 2004; Guttorm et al. 2005; Leppänen et al. 2010; Choudhury and Benasich 2011; Maitre et al. 2013).

Electrophysiological studies in human adults have demonstrated clear neural responses to specific phonemic features (Steinschneider and Fishman 2011; Steinschneider et al. 2013; Khalighinejad et al. 2017) and selective spatial representation of phonemes’ temporal-spectral characteristics in the posterior superior temporal gyrus (P-STG) (Chang et al. 2010; Mesgarani et al. 2014; Hullett et al. 2016). In particular, bilateral increases in high-gamma power (75–150 Hz) in the P-STG during phonemic processing (Crone et al. 2001; Nourski et al. 2009; Steinschneider et al. 2011; Nourski 2017) distinctively characterized representation of temporal acoustic information such as voice-onset-time, an important feature for phoneme discrimination between stop consonants that share the same place of articulation. There is already evidence demonstrating high-gamma activation to phonemic processing in adults (Steinschneider et al. 2011) but to our knowledge, only one study has examined high-gamma power (>70 Hz) during the period in which infants are establishing their phonetic maps. In a longitudinal study from 6- to 12-months-of age, the 12-month-old infants showed an increase in high-gamma power in left auditory cortex to native but not to non-native phonemes, suggesting that enduring cortical representations of native language have been well established by the end of the first year of life (Ortiz-Mantilla et al. 2016).

It has been demonstrated that cortical effects of sensory input are continuously modified by experience (Kilgard et al. 2001; Cheour et al. 2002; Shestakova et al. 2003; Moucha et al. 2005; Threlkeld et al. 2009; Froemke and Jones 2011). During the first months of life, the brain, and in particular the auditory cortex (Froemke and Jones 2011) is highly plastic making infancy an invaluable period for optimizing language mapping. Further, it has been reported, using rat models, that auditory training can improve as well as remediate deficits in auditory processing (de Villiers-Sidani et al. 2007; Threlkeld et al. 2009). Benasich et al. (2014) demonstrated in human infants, that early acoustic experience with spectrotemporally modulated non-speech stimuli from 4 to 6 months-of-age significantly impacted the processing of known and new non-speech signals

at 7-months. The addition of attention and infant control (via an interactive acoustic experience [IAE]) induced even more striking advantages, increasing attention to environmental acoustic stimuli, enhancing acoustic mapping and sharpening discrimination in the tens-of-milliseconds range critical to phonemic perception. Examination of the neural mechanisms underlying these processing effects found that infants who received the IAE displayed a left-lateralized increase in amplitude of low-gamma oscillations during tone discrimination, whereas, infants who passively listened to the sounds and naïve controls showed less gamma activation and a bilateral pattern of response (Musacchia et al. 2017). Similarly, increased left-lateralized high-gamma power during mapping of native phonemes at 12-months has been reported as infants gain more experience with the language spoken in their environment (Ortiz-Mantilla et al. 2016).

Taken together, these studies demonstrate that across the first year of life, changes in the auditory cortex induced by speech or speech-like experience are supported by enhancement of gamma oscillations. And further, that active engagement with linguistically relevant acoustic cues within non-speech supports faster processing speed, as well as more accurate decoding of those cues known to be relevant for recognition of the distinctive characteristics that distinguish phonemes (Benasich et al. 2014; Musacchia et al. 2017). However, a larger question remains, specifically, does pre-linguistic experience with non-speech containing spectrotemporally modulated acoustic cues translate to better processing of speech information, more efficient discrimination abilities and/or enhanced phonemic mapping? And further, does that linguistic impact endure past the immediate training period, providing an ongoing advantage for speech?

In the present study, we examine the specific effects of early acoustic experience with non-speech stimuli on later speech processing. Two groups of typically developing infants received between 4- and 6-months-of age, 6 weekly sessions of auditory stimuli with temporally modulated non-speech sounds. One group participated in an interactive acoustic experience (AEx group) in which attention was specifically driven to detect changes in the auditory environment; the second group of infants (PEx group) were passively exposed to paired non-speech stimuli varying in interstimulus intervals (Benasich et al. 2014). At 9-months-of-age, dense-array EEG/ERP signals were recorded while AEx and PEx infants passively listened to a phonemic contrast varying in voice-onset-time (/da-/ta/), presented in an oddball paradigm. To tease apart maturational from experimental effects, a third, cross-sectional group (NC9), that did not have any sound training, served as naïve, age-matched controls. Using source localization techniques to determine the generators of the auditory response, we examined oscillatory dynamics in left and right auditory cortices, from 2 to 90 Hz. We were particularly interested in investigating whether, in response to the standard syllable, infants’ amount of spectral amplitude in high-gamma (>70 Hz) range varied as a function of acoustic experience (active or passive) as compared with naïve controls. Since: (1) the non-speech stimuli used in the early auditory experience contained speech-like temporal and spectral cues known to be relevant for speech recognition of consonant voicing features (Xu 2005), and (2) experience with those non-speech acoustic signals has been shown to sharpen infants’ abilities to detect and discriminate known and novel non-speech stimuli at 7-months-of age (Benasich et al. 2014), we posited that such experience-dependent effects, especially when attention is actively driven, may generalize to speech processing of familiar phonemic features, and possibly

enhance phonemic cortical representations above and beyond maturation alone.

Materials and Methods

Participants

Our sample included 53 (23 females/30 males) typically-developing infants, part of a mixed longitudinal and cross-sectional developmental study, 35 of which comprised a previous study cohort (Benasich et al. 2014). Those 35 infants were recruited at 4-months and invited to participate for 6 weeks in an acoustic exposure study. Infants were then assigned to either an interactive auditory experience (AEx group) or a passive auditory exposure (PEX group) and were followed longitudinally through 9-months-of-age. We retained all 18 AEx infants and all 17 PEX infants but during analysis, one PEX infant (male) was excluded from the time-frequency analysis due to a high-level of noise in the gamma range. In addition, 18 age-matched typically-developing infants were recruited at 9-months to participate as a cross-sectional naïve control group (NC9), however 2 infants (one male, one female) were later excluded due to poor data quality. All infants had uneventful prenatal and perinatal circumstances, were born healthy, full-term, with normal birth weight into monolingual English families (detailed sample information can be found in Table 1) and had passed the newborn hearing screening. Information about gestational age and birth weight as well as language spoken at home was collected via parental questionnaire at the 4-month visit for AEx and PEX infants, and at the 9-month visit for NC9 infants. Infants were recruited from urban and suburban communities in New Jersey, and had no family history of specific language impairment, autism, hearing loss, no repeated episodes of otitis media, or other medical, neurological, or psychiatric disorders. Parents were compensated for their time and infants received a toy after the visit. The study was conducted in accordance with the Declaration of Helsinki, and informed consent approved by the Rutgers University Human Subjects Institutional Review Board was obtained before inclusion in the study.

Procedure

Behavioral protocol for the Interactive Auditory Experience (IAE): Infants in the AEx group had visited the laboratory once a week for 6 consecutive weeks between 4 and 6 months-of-age (mean: 4.7 [SD: 0.3] – 5.9 [0.3] months). A go/no-go operantly-conditioned looking task was designed for the AEx group in which they learned an association between a series of auditory stimuli 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. Infants were

trained 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. The sound stimuli were presented at varying inter-stimulus-intervals (ISIs), depending on the phase of the session, using an up-down staircase procedure (Trehub et al. 1986). The ISI for each block of stimuli was increased or decreased according to infant performance. The task continued for approximately 7–9 min each session, until the child fatigued. This task was designed to focus attention on key pre-linguistic cues that had relevance for subsequent linguistic mapping while facilitating wider entrainment of auditory neurons (for a more detailed explanation of the AEx procedure, refer to Benasich et al. 2014).

Passive Auditory Exposure: Infants in the PEX group also visited the laboratory once a week for 6 consecutive weeks between the ages of 4 and 6 months-of-age (mean: 4.8 [SD: 0.3] – 6.0 [0.2], months). The PEX group was exposed to the same stimuli as the AEx group. The infant sat in an infant seat placed equidistant between left and right speakers in a sound-attenuated and electrically shielded sound booth (Industrial Acoustics Company). The stimuli were presented free field in an oddball fashion (80% standards [STD], 20% deviants [DEV], total: 665 stimuli per block), while the infant was silently entertained with puppets/toys to maintain alertness. Two blocks of stimuli were presented in random order at each session, 10 min at 40 ms ISI and 10 min at 70 ms ISI. This condition was designed to increase spectrotemporal processing efficiency through controlled background exposure.

Stimuli for interactive and passive auditory protocols: Infants in both groups were presented with 3 different types of acoustic stimuli, as follows: weeks 1 and 2, complex tone pairs (STD: 800–800 Hz; DEV: 800–1200 Hz); weeks 3 and 4, bandpass noise pairs (STD: 400–1900 Hz and 400–1900 Hz; DEV: 400–1900 Hz and 800–1900 Hz); and weeks 5 and 6 simple sweep pairs (STD: 1600–1200 Hz and 1600–1200 Hz; DEV: 1600–1200 Hz and 1200–1600 Hz).

Event-Related Potentials (ERP)

At 9 months-of-age, 3 months after the interactive or passive acoustic experience ended, the AEx and PEX infants again visited the laboratory to participate in an assessment that included an EEG/ERP session. A group of 16 naïve infants without any previous acoustic training serve as the 9-month cross-sectional control group (NC9). The three groups of infants participated in the EEG/ERP session which was recorded under identical conditions for all infants.

Stimuli

The stimuli for the ERP were computer generated consonant-vowel syllables differing in voice-onset-time (VOT). The

Table 1 Groups' characteristics

Group	Sex F/M	GA weeks (SD)	GA range	BW grams (SD)	BW range	Age months first session (SD)	Age months last session (SD)	Age weeks at 9-month ERP (SD)
AEx <i>n</i> = 18	8/10	39.7 (1.2)	38–41	3439.1 (450.4)	2551.4–4280.7	4.71 (0.3)	5.88 (0.3)	39.50 (0.9)
PEX <i>n</i> = 17	8/9	39.2 (0.9)	38–40	3435.2 (535.7)	2579.8–4592.5	4.77 (0.2)	5.98 (0.2)	39.64 (0.8)
NC9 <i>n</i> = 16	7/9	39.6 (1.0)	38–41	3430.2 (507.6)	2834.0–4507.5			39.63 (0.8)

AEx: interactive experience group; PEX: passive exposure group; NC9: naïve control group at 9-months; *n*: number of participants in the group; F/M: female/male; GA: gestational age; (SD): standard deviation; BW: birth weight; ERP: event-related potential.

standard (STD) stimulus was /da/ (VOT = 0 ms) and the deviant (DEV) stimulus was /ta/ (VOT = 40 ms). The consonant transition time of 40 ms was followed by a 60 ms steady-state vowel. The duration of each syllable was 100 ms; the fundamental frequency was 120 Hz and three additional formants F1, F2, and F3 were at 750 Hz, 1200 Hz, and 2500 Hz respectively. The stimulus onset-to-onset interval was 1020 ms, and the offset-to-onset interval 920 ms. The stimuli were presented in a pseudo-randomized passive oddball paradigm, using a block design comprised of 566 standards (85%) and 100 deviants (15%), 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 9-months of age (mean age 39.5 weeks [SD: 0.9] range: 37–41 weeks) while participants were seated in their parent's lap, watching a silent movie or entertained with quiet toys to keep them calm and engaged (Musacchia et al. 2015). EEG was recorded from a 125-channel EGI sensor net (Electrical Geodesic, Inc. Eugene, Oregon) with the vertex electrode used as on-line reference, sample rate of 250 Hz, and high/low pass filters of 0.1 and 100 Hz, respectively. Artifact correction of eye movements was completed on the raw data using an automatic correction algorithm based on Principal Component Analysis method (PCA) provided in Brain Electrical Source Analysis (BESA GmbH) 5.3 software. ERPs were processed using an off-line band-pass filter of 1–15 Hz and re-referenced to an average reference. EEG/ERP data were then segmented into epochs according to stimulus type (STD, DEV) with 300 ms pre-stimulus, 1020 ms post-stimulus time and 100 ms before stimulus onset used as the baseline. Epochs with signals exceeding $\pm 300 \mu\text{V}$ from the baseline were excluded. The STD stimulus corresponded to the syllable /da/ and was presented 566 times while the DEV /ta/ was available in only 100 of the trials resulting in reduced signal-to-noise ratio for DEV as compared with STD. Since we were interested in exploring effects of auditory experience in early sensory processing and phonemic mapping, and the repetition of the STD promotes a neural representation (sensory map) that is maintained in auditory memory for comparison with subsequent input (Näätänen and Winkler 1999), we included only the STD stimulus in the analyses related to cortical representation. Given the larger number of trials presented, the STD supports a higher signal-to-noise ratio (Heim et al. 2011), thus increasing the probability of having a more reliable phoneme representation. A minimum of 70% (396) artifact-free STD epochs per infant (STD mean: 494 trials, range: 413–536) was required for inclusion in ERP averaging. To specifically examine change-detection abilities we used the DEV stimulus. A minimum of 70% (70) artifact-free DEV epochs per infant (mean: 84 trials, 71–95) was required for inclusion in ERP averaging.

Source Localization of ERP Generators

Source localization is a technique used to identify the loci of the neural activation registered at the scalp surface. In that way, the high temporal resolution provided by EEG/ERP can be combined with structural images (individual or averaged) to give a closer spatial approximation as to where in the brain neural responses are being generated. Therefore, to localize source generators of the response to the STD and DEV syllables, EEG/ERPs were mapped at 9 months onto an age-appropriate brain template

(12-month template), using BESA 5.3 and Brain Voyager QX software programs (Scherg M, Berg P, Hoehstetter K. BESA research tutorial 2: EEG-fMRI coregistration, preprocessing, ERP and source analysis [2010a] <http://www.besa.de/downloads/training-material/tutorials/>) following an infant protocol (Hämäläinen et al. 2011; Ortiz-Mantilla et al. 2012, 2013, 2016; Musacchia et al. 2013, 2015, 2017). Based on a principal component analysis algorithm (PCA), and Global Field Power (GFP), peaks of interest for the STD and DEV responses were identified in the grand average of each group and in each individual waveform. PCA decomposes the data into mutually orthogonal topographies identifying the more dominant component that explains the largest fraction of the variance for the response of interest (Scherg M, Berg P, Hoehstetter K. BESA Tutorial 2: EEG-fMRI coregistration, preprocessing, ERP & source analysis [2010] <http://www.besa.de/downloads/training-material/tutorials/>). A discrete dipole source model (Scherg and Von Cramon 1985) using a 4-shell ellipsoidal head model and a confirmatory distributed source model calculated via Classic LORETA Recursively Applied (CLARA) method (Hoehstetter K, Berg P, Scherg M. BESA Research Tutorial 4: Distributed Source Imaging [2010] <http://www.besa.de/downloads/training-material/tutorials/>) were applied to the first positive peak (perceptual response) of each condition for source modeling. A time window of ± 20 ms around the peak was used for dipole fitting. A 2-dipole model identified sources of activation in left and right auditory cortices. The source montage generated during discrete dipole fitting in the grand average of each group (AEx, PEx, NC9) was saved for further use during time-frequency analysis. Individual P1 peak amplitude and latency were submitted to statistical analyses.

Time-Frequency Analyses in Source Space

Spectrotemporal changes in event-related oscillations during the STD and DEV syllable processing were examined in source space (Ortiz-Mantilla et al. 2013, 2016) as follows: The previously saved 2-source montage (that works as a fixed spatial filter) from the grand average source model was applied to the raw 125-channel recording of each individual in the corresponding group to transform the continuous EEG into 2-channel source space (Hoehstetter et al. 2004). 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 representation (Scherg M, Berg P, Hoehstetter K. BESA research tutorial 6: Time frequency analysis and source coherence [2010b] <http://www.besa.de/downloads/training-material/tutorials/>). To control for low frequency activity while at the same time preserving as much of the frequency information as possible, a low cutoff of 0.5 Hz was applied to the raw EEG. Event-related changes in oscillatory amplitude of frequency bands (Tallon-Baudry et al. 1996; Hari and Salmelin 1997; Tallon-Baudry and Bertrand 1999) were investigated using temporal spectral evolution (TSE) measures. The TSE measures the percentage of amplitude (spectral power) change as compared with the baseline of induced (random-phase/nonphase-locked) and evoked (phase-locked) oscillatory activity related to stimulus presentation (Tallon-Baudry et al. 1996). TSE individual results generated for the STD and DEV in left and right auditory cortices for each group were exported to MATLAB (MathWorks) for plotting graphics across subjects.

Statistical Analyses

The P1 source strength and latency and the residual variance not explained by the dipole model were examined separately

using source (left, right) by group (AEx, PEx, NC9) repeated measures ANOVAs in SPSS Statistics 24 (SPSS, Inc.) software. Source coordinates (x: medial-lateral; y: anterior-posterior; and z: superior-inferior) were examined separately for each direction also using repeated measures ANOVAs. Time frequency analysis were conducted using cluster identification and permutation testing. We detected time-frequency regions (clusters) with significant changes in the magnitude of amplitude (TSE) using BESA Statistics 1.0 (BESA GmbH, Gräfelfing, Germany) software. In a first step, BESA Statistics uses a preliminary (parametric) Student's t-test (Maris and Oostenveld 2007) per data point to determine if there is a significant difference between the groups/conditions means. To deal with the multiple comparison problem, BESA Statistics uses parameter-free permutation testing in combination with data clustering. There were no predefined clusters as BESA Statistics automatically identifies clusters in time and frequency between 2 groups (i.e., between AEx and NC9 or between PEx and NC9) that show a significant effect and calculates a cluster value from the sum of all t-values of all data points in the cluster. As a next step, a permutation procedure, in which data is randomly interchanged multiple times (1000 permutations in this case), is conducted to test if the initial data cluster survives the permutations. Results are considered corrected for multiple comparisons as only those clusters will be identified that have higher cluster values than 95% of all clusters derived by random permutation of data (For a more detailed description of the methods refer to BESA Statistics website: <http://www.besa.de/products/besa-statistics/besa-statistics-overview/>). To determine the specific pattern of the oscillatory dynamics in the two experimental conditions, each of the experimental groups (AEx and PEx), were separately compared with the NC9 group. In this way, we were able to identify "data clusters of significance" between the groups in the time-frequency domain. The P statistics values reported in here were derived from the permutation testing and cluster analyses.

Data Accessibility

Data files are securely stored per IRB guidelines at the Infancy Studies Lab at Rutgers University Rutgers-Newark. Access will be granted upon request.

Results

Our preliminary analysis did not reveal differences in gestational age ($F_{(2,49)} = 1.12$, $P = 0.3$), birth weight ($F_{(2,49)} = 0.01$, $P =$

0.9) or gender ($X^2 = 0.17$, $P = 0.9$) among the groups. No difference was found between the AEx and PEx groups in the age at which they received their first ($F_{(1,34)} = 0.38$, $P = 0.5$) and last (sixth) acoustic session ($F_{(1,34)} = 1.39$, $P = 0.2$) nor in the mean age at which the 9-month ERP was recorded ($F_{(2,49)} = 0.15$, $P = 0.8$). Characteristics of the sample are included in Table 1.

Source Localization of ERP Responses

ERP responses to the standard (STD) and deviant (DEV) stimulus closely resembled those reported in other studies using syllables differing in voice onset time (Rivera-Gaxiola et al. 2005; Ortiz-Mantilla et al. 2013, 2016). The ERP response was characterized by a fronto-central positivity followed by a negative deflection. Measured in the grand average ERP at 9 months-of-age, the peak of the positive response for the STD occurred at ~160 ms followed by a negative deflection at ~360 ms; for the DEV, the peak of the positive response was at ~172 ms and for the negative peak at ~380 ms with inversion of the polarity observed at the mastoids and posterior channels. The ERP morphology was similar among the groups. To explore differences at the perceptual level, generators of the positive (P1) ERP response were identified in the ERP waveforms. Differences among the groups in sound representation, that captures ongoing cortical mapping, were examined in the STD as this stimulus was presented 85% of the time and the probability of occurrence for the DEV results in reduced signal to noise ratio. Thus, we had more trials and therefore a cleaner signal, increasing the odds of success when examining cortical representation of the regularities of /da/. Group differences in change detection, that reflects stimulus discrimination, were examined via the response to the DEV. A 2-dipole model freely fitted in the grand average ERP waveform explained ~97% of the variance (see Table 2 for time windows chosen for dipole fitting and corresponding residual variance for each group and stimulus).

The free dipole fitting procedure placed dipoles in both auditory cortices for both STD and DEV (Fig. 1); bilateral sources of activation were confirmed by distributed source model, CLARA; no evidence of additional sources of activation was found with the distributed model. 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. 1). Differences among the groups in amplitude and latency of the P1 source component for STD and DEV were examined separately using a 2x3 (Source [left, right] x Group [AEx, PEx, NC9])

Table 2 Parameters chosen for dipole fitting of the First Positive response (P1) to the standard stimulus /da/ and to the deviant stimulus /ta/. Left and right P1 amplitude and latency were measured in the grand average source waveform

Group	STD TW (ms)	STD RV (%)	STD LAC Amp (nAm)	STD LAC Lat (ms)	STD RAC Amp (nAm)	STD RAC Lat (ms)
AEx	136–176	2.103	12.87	164	11.87	164
PEx	128–168	3.263	13.97	164	13.09	164
NC9	124–164	2.755	15.76	160	13.18	152
Group	DEV TW (ms)	DEV RV (%)	DEV LAC Amp (nAm)	DEV LAC Lat (ms)	DEV RAC Amp (nAm)	DEV RAC Lat (ms)
AEx	160–200	1.933	17.78	180	16.74	180
PEx	144–184	2.413	15.45	180	14.84	168
NC9	144–184	3.140	18.31	180	16.05	168

AEx: interactive experience group; PEx: passive exposure group; NC9: naïve control group recruited at 9-months-age; (SD): standard deviation; P1: first positive peak; STD: standard stimulus; TW: time window for dipole fitting; ms: milliseconds; RV: residual variance; %: percentage of residual variance non-explained by the dipole model; LAC: left auditory cortex; RAC: right auditory cortex; Amp: amplitude; nAm: nanoamperes; Lat: latency.

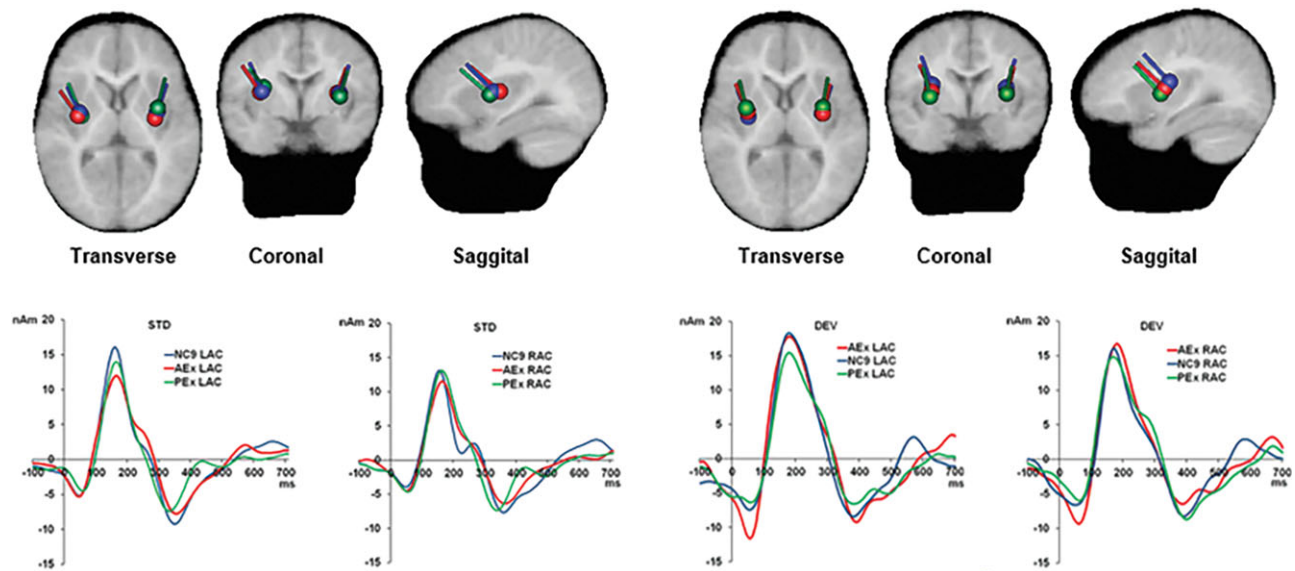


Figure 1. First row: Source localization of the Positive (P1) generators in response to the standard syllable /da/ (first row to the left) and to the deviant syllable /ta/ (first row to the right) are seen in transverse coronal and sagittal views in an age-appropriate brain template. The interactive acoustic experience (AEx) group is shown in red, the passive acoustic exposure (PEx) group in green and the naïve control at 9 months (NC9) group in blue. Discrete dipoles for each group are located in left (L) and right (R) auditory cortices. Second row: Grand average source waveforms of the response to the standard syllable /da/ (depicted on the left) and to the deviant syllable /ta/ (depicted on the right), at left (LAC) and right (RAC) auditory cortices. The interactive acoustic experience (AEx) group is shown in red, the passive acoustic exposure (PEx) group in green and the naïve control at 9 months (NC9) group in blue. Positivity is plotted up; time is shown in ms on the x-axis and amplitude of the source dipole moment is given in nanoampere meters (nAm) on the y-axis.

repeated measures ANOVA. No significant differences in amplitude or latency of the P1 for STD or DEV (Latency: STD $F_{(2,48)} = 2.906$, $P = 0.064$; DEV $F_{(2,48)} = 0.363$, $P = 0.697$; Amplitude: STD $F_{(2,48)} = 1.947$, $P = 0.154$; DEV $F_{(2,48)} = 1.987$, $P = 0.148$) were found. Similarly, no differences in the source locations on the x, y, and z coordinates or in the residual variance were significant among the groups (RV for STD: $F_{(2,48)} = 1.77$, $P = 0.18$; RV for DEV: $F_{(2,48)} = 1.01$, $P = 0.37$).

Time-Frequency Analysis

To examine if the oscillatory dynamics supporting cortical representation of the regularities of the standard syllable and discrimination of the deviant syllable were modulated by early acoustic experience, temporal spectral analyses were conducted from -300 to 1020 ms in the 2–90 Hz frequency range. Measure of temporal spectral evolution (TSE), that compute stimulus-related spectral amplitude change (spectral power) as compared with the baseline, was obtained via permutation analysis and cluster identification separately for AEx compared with NC9, for PEx compared with the NC9 and for AEx compared with PEx.

- A. Effects of early (between 4 and 6 months of age) interactive acoustic experience, in which attention was involved: Group comparison between AEx and NC9 groups at 9-months of age.

Our results showed group variations in the amount of spectral power in theta, beta and gamma ranges during STD and in theta and gamma ranges during DEV syllable processing.

- Differences in STD processing (syllable representation): We found that in the theta band, the AEx group showed less spectral power (4–9 Hz, 100–800 ms) than the NC9 group in both left ($P = 0.017$) and right ($P = 0.031$) auditory cortices (Fig. 2). The oscillatory dynamics also varied

between the groups in the beta band as the AEx group has more power (13–29 Hz, 0–800 ms, $P = 0.013$) than the NC9 group in left auditory cortex (Fig. 2). Similarly, the AEx group also displayed more bilateral spectral power in high-gamma (82–87 Hz, 100–850; left: $P = 0.008$; right: 0.033) than the NC9 group (Fig. 2).

- Differences in DEV processing (syllable discrimination): Differences between the groups were also found in the theta and gamma frequency bands during processing of the DEV. The AEx group showed more theta power (4–8 Hz, 50–500 ms, $P = 0.033$) in right auditory cortex than the NC9 group (Fig. 3). In the gamma band, both groups increased oscillatory activity in left auditory cortex (Fig. 3) but each group showed a particular pattern: whereas the AEx group had more spectral power in an early time-window in gamma frequencies around 50 Hz (150–450 ms, 44–52 Hz, $P = 0.006$) the NC9 group generated more power in a lower frequency range and at a later time-window (31–38 Hz, 600–900 ms, $P = 0.018$).

- B. Effects of early (between 4 and 6 months of age) passive acoustic exposure in which attention was not required: Group comparison between PEx and NC9 groups at 9-months of age.

- Differences in STD processing (syllable representation): Significant differences in the amount of spectral power allocated to STD processing were found between PEx and NC9 groups. In the theta range (5–9 Hz, 200–800 ms), the PEx group showed less power than the NC9 group ($P = 0.016$) in left auditory cortex (Fig. 2). No significant differences between the PEx and NC9 groups were detected in beta or high-gamma frequency ranges. However, a significant gamma cluster (30–37 Hz, 50–850 ms) indicated that the PEx group had significantly more power ($P = 0.017$) in left auditory cortex than the NC9 group.

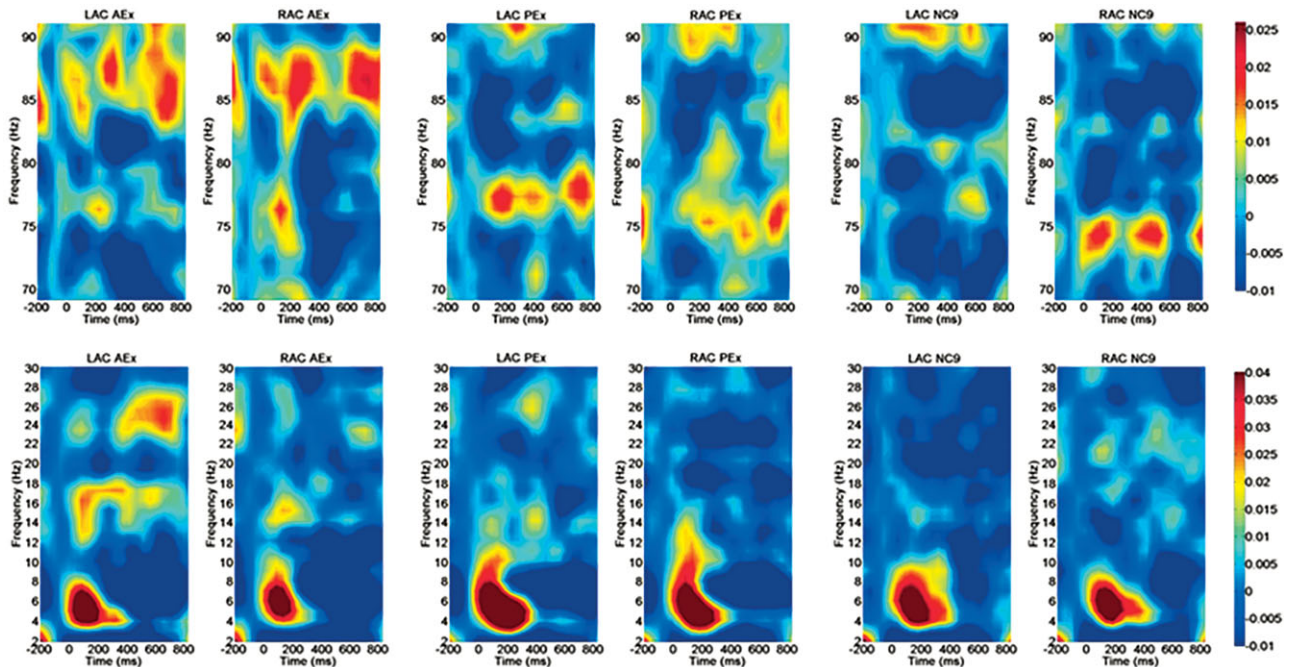


Figure 2. First row: Time Frequency plots showing changes in spectral power (temporal spectral evolution) as a response to the standard syllable /da/ in high-gamma (70 Hz) frequency band. The first two plots illustrate oscillatory activation in the left (LAC) and right (RAC) auditory cortices for the interactive acoustic experience (AEx) group, the two middle plots show activation for the passive acoustic exposure (PEx) group and the last two plots represent oscillatory power for the naïve control at 9 months (NC9) group. Bilateral high-gamma activity (82–87 Hz) is clearly seen for the AEx group in the upper part of the left and right auditory cortices. Note that a small activation is just beginning to appear for the PEx and NC9 groups at ~75 Hz but permutation analysis did not identify a cluster of significance at this level. Second row: Time Frequency plots for theta and beta frequency bands showing changes in spectral power as a response to the standard syllable /da/. The first two plots illustrate oscillatory activation in left (LAC) and right (RAC) auditory cortices for the interactive acoustic experience (AEx) group, the middle two plots, activation for the passive acoustic exposure (PEx) group and the last two plots for the naïve control at 9 months (NC9) group. Theta range (4–8 Hz) is shown at the bottom of the plots, while beta activity (13–30 Hz) is clearly and exclusively seen for the AEx group in the mid to upper portion of the first plot (left auditory cortex). Time is shown in milliseconds on the x-axis and frequency in Hz on the y-axis.

- Differences in DEV processing (syllable discrimination): differences between the groups during processing of the DEV syllable were found only in the theta range. As compared with NC9 group, the PEx group showed more theta power (2–8 Hz, 0–400 ms, $P = 0.018$) in right auditory cortex (Fig. 3).

C. Comparing effects of early interactive acoustic experience (attention involved) and passive acoustic exposure (attention not required) on syllable processing at 9-months-of-age:

Lastly, we examined if activity in the high-gamma range differed between infants who were passively exposed to non-speech stimuli (PEx group) and those that had an interactive acoustic experience (AEx group). We found that during early sensory processing of the STD the AEx group recruited more high-gamma power in both left (81–88 Hz, 0–200 ms; $P = 0.037$) and right (82–86 Hz, 0–300 ms; $P = 0.048$) auditory cortices (Fig. 2) than the PEx group. The AEx group also exhibited more left spectral power in the gamma range (45–60 Hz, 50–500 ms, $P = 0.001$) than the PEx group during processing of the DEV stimulus (Fig. 3).

Discussion

Our results provide evidence at the oscillatory level, which precisely captures the underlying neural response, that young infants who actively engage with spectrotemporally-modulated non-speech stimuli for short intervals each week, over a 6-

week period, show speech processing, change detection and phonemic mapping advantages at 9 months-of-age. Thus, the experience-dependent effects of non-speech training detailed in Benasich et al. (2014) not only generalize to linguistic stimuli but also extend beyond the immediate post-training period. Infants who received interactive acoustic experience (IAE) from 4- to 6-months processed the STD syllable differently at 9-months-of-age than both naïve controls and those with only passive acoustic exposure, demonstrating increases in high-gamma power within auditory cortices and higher beta power in left auditory cortex. The increases in high-gamma power to the STD syllable, seen exclusively in the AEx group suggest that early IAE with non-speech, that contains acoustic cues pertinent to linguistic decoding, can bootstrap processing of native speech and facilitate establishment of enduring phonemic representations. Moreover, we posit that the higher beta power observed in the AEx group reflects increased linking of sensory information and corresponding articulatory patterns, which would also benefit competent language acquisition. The fact that AEx and PEx groups showed less theta power than naïve controls suggests early acoustic intervention promotes a more mature and automatized processing of the frequently-presented STD syllable. The modulatory effects of early acoustic experience were not only seen for the processing of the STD, but also were evident for the DEV syllable. The AEx infants showed enhanced gamma power in the left auditory cortex at early stages of sensory processing, suggesting that interactive acoustic experience favored change-detection of the sub-lexical cues critical for phoneme identification. Furthermore, differing

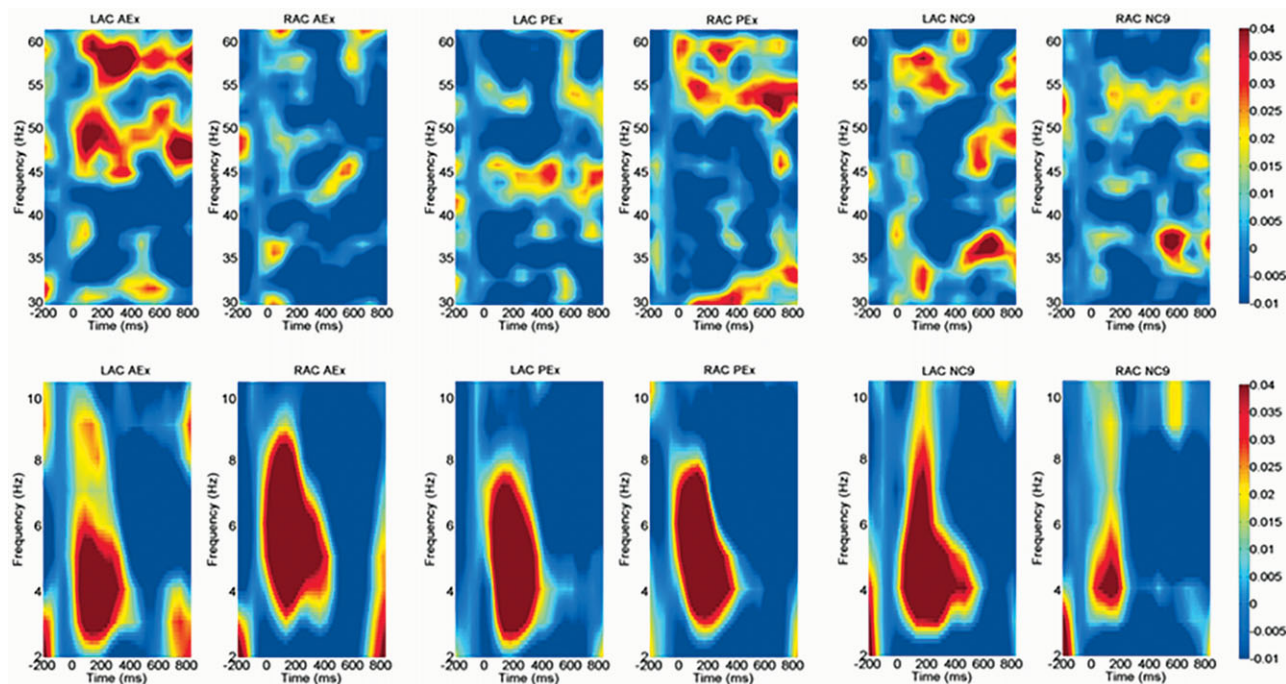


Figure 3. First row: Time Frequency plots for the gamma frequency band showing changes in spectral power as a response to the deviant syllable /ta/ in gamma (<70 Hz) frequency band. The first two plots illustrate oscillatory activation in left (LAC) and right (RAC) auditory cortices for the interactive acoustic experience (AEx) group, the two middle plots depict activation for the passive acoustic exposure (PEx) group and the last two plots represent oscillatory power for the naïve control at 9 months (NC9) group. Second row: Time Frequency plots for the theta frequency band showing changes in spectral power as a response to the deviant syllable /ta/. The first two plots illustrate oscillatory activation in left (LAC) and right (RAC) auditory cortices for the interactive acoustic experience (AEx) group, the middle two plots show activation for the passive acoustic exposure (PEx) group and the last two plots activation for the naïve control at 9 months (NC9) group. Time is shown in milliseconds on the x-axis and frequency in Hz on the y-axis.

from NC9 group, both AEx and PEx groups also showed theta power increases in right auditory cortex implying that acoustic experience may facilitate the more demanding processing of the sporadic DEV in the right auditory cortex, an area primarily involved in decoding slow-rate syllabic information. While the largest effects were seen for the AEx group, in gamma, beta, and theta ranges, it is clear that passive exposure also impacted subsequent processing in the PEx group, particularly for syllabic in the theta range.

Efficient auditory decoding abilities support a fundamental function of the brain, perceptual processing speed, which is critical to proficient brain function across the life span. Establishment of cortical representations of native phonemes is an essential step in this process as infants progress toward automatically and efficiently processing language input. For the majority of children, acquiring language appears to be an easy task. Some infants, however, have much difficulty setting up language, and as a consequence present with developmental language disorders (DLD) that impact their intellectual progress throughout their lives. Although we have shown that we can identify children who are poor processors of critical acoustic input as infants (Benasich 2002; Benasich and Tallal 2002), it has been the case that even when it is clear that processing is effortful, little has been done to improve such issues until a child exhibits difficulty in the classroom setting. The interactive baby-friendly intervention (IAE) we developed appears to bootstrap this early process of acoustic mapping, supporting optimal and efficient processing of basic pre-linguistic cues and providing scaffolding that can optimize the efficacy of this developmental process. Interestingly, it appears to do so even in typically developing infants without apparent familial risk

factors. Sensory/perceptual training is quite easy at this early age; however, it is much more challenging to effectively train categories of non-linguistic sounds that will allow generalization to novel spatiotemporally-organized acoustic cues. Nonetheless, if it is possible to improve early language acquisition by targeting basic pre-linguistic acoustic processing abilities, we may be able to ameliorate or even prevent the ongoing negative social and educational impact of DLDs.

Speech perception has been considered a multi-time resolution process involving coordinated oscillatory activity at different time scales. In this way, slow-rate oscillations such as delta (1–3 Hz) and theta (4–8 Hz) synchronically interact with fast-rate activity in the gamma (>30 Hz) range to resolve prosodic, syllabic and phonemic information (Poeppel et al. 2008; Giraud and Poeppel 2012). Gamma oscillations have been shown to correlate with the temporal sampling of fast duration cues, such as the formant transitions seen in the phonemic scale. Theta oscillations closely correlate with temporal sampling at slow modulation rates, corresponding to the syllabic scale, whereas delta activity relates to the even slower rate reflecting the prosodic elements contained in words and sentences. Theta and gamma generators may be weakly coupled at rest but become strongly coupled and nested in response to an incoming speech stream, supporting the concept that speech is processed in different but synchronized timescales (Giraud and Poeppel 2012; Hyafil et al. 2015). We have not yet examined phase-amplitude or cross-frequency coupling in infants, however, we do plan such analyses in the future, as we believe that investigating this, to date, unexplored domain in infants, may be critical to deepening our understanding of how speech information is coordinated among different frequency bands during

the early stages of language acquisition. It is important to note, however, that even without data in this domain, our results line up nicely with the adult literature in these not yet mature systems. We found that similar to adults, 9-month-old infants process phonemic cues that occur in the ten of milliseconds range, using gamma band (whose oscillatory rate aligns with the phonemic rate) and process syllables, occurring in the hundreds of milliseconds range, using theta band (whose oscillatory rate aligns with the syllabic rate) to extract different speech elements in a temporal multidimensional manner. Although still not at adult levels, it seems clear that young infants already have in place the temporally organized hierarchical structure necessary to process language. A few studies that have been conducted in the first year of life have reported involvement of low frequency oscillations during syllable processing (Bosseler et al. 2013; Ortiz-Mantilla et al. 2013) and of high-gamma oscillations signaling phonemic mapping (Ortiz-Mantilla et al. 2016).

In this study, we found that typically developing, 9-month-old (NC9) infants had not yet established mature cortical representations of phonemic features, given that enhanced high-gamma power to the STD syllable /da/ was not evident. Conversely, infants who had received early interactive acoustic experience (AEx) clearly displayed high-gamma power in auditory cortices, implying that phonemic mapping was already established. Similar to the NC9 group, the PEx infants, who were passively exposed to the same acoustic information as the AEx group, showed no significant increase in high-gamma power. This suggests that in addition to acoustic experience, interactive engagement with attention was a pivotal factor in inducing the high-gamma activation seen in the AEx group, which is held to be a signature of phonemic mapping (Steinschneider et al. 2011; Ortiz-Mantilla et al. 2016).

Early interactive acoustic experience also favored discrimination at the phonemic level. In response to the DEV syllable /ta/, the AEx group increased left gamma power, confirming the dominant role of gamma oscillations in supporting fast-rate discrimination of the sub-lexical information required for phoneme identification (Zatorre and Belin 2001; Hickok and Poeppel 2007; Giraud and Poeppel 2012). Discrimination abilities are critical for accurate phoneme identification, as they enable perception and decoding of rapid acoustic changes in incoming speech central to early language acquisition (Eimas et al. 1971; Jusczyk et al. 1980; Benasich and Tallal 2002). It is important to note that during the 9-month ERP session analyzed here (collected ~3 months after the acoustic experience ended), syllables were presented in a passive oddball paradigm, therefore, overt attention was not required. But it seems that the cortical plasticity in acoustic cortex induced at 7-months by the IAE (Benasich et al. 2014; Musacchia et al. 2017) has ongoing effects that both generalize to language and extend beyond the immediate post-training period.

In adults, direct intracranial recordings with surface arrays placed on auditory cortex, have shown that early (50–250 ms) high-gamma (70–150 Hz) activity on posteromedial Heschl's gyrus represents spectral-temporal features of phonemic sounds (Steinschneider et al. 2011) and is minimally modulated by task demands, context or attentional level (Nourski 2017). In infancy, attention may well be critical for cortical mapping but in adulthood, when phonemic representations are well established, and phonemic processing is automatized, this is not the case and overt attention is not required for phoneme identification. On the other hand, task-related increases in high-gamma activity (thus, involving attention) have been reported to occur

in adults, but in a later (after 250 ms) time window (Nourski et al. 2015). Our findings reveal a significant cluster of high-gamma enhancement from 100 to 800 ms, and this may well include both the initial activation responding to the spectral-temporal characteristics of the phoneme information contained in the syllable /da/ and later activation implying attentional modulation. One explanation for this extended processing in infants could be that over development, while cortical representations are not yet fully established in the infant brain, allocation of neuronal resources to process phonemic information is required for an extended time interval. An additional possibility is that attention plays an even more essential role and due to their previous interactive acoustic experience, AEx infants are more aware/vigilant of the auditory environment. Consequently, their attention might be more easily recruited, resulting in additional neuronal resources allotted to processing of the “passive” speech sounds played to them during the EEG/ERP recording (Nourski et al. 2015). As we do not yet fully understand how these various processes interact over infancy while phonemic cortical representation, change-detection of phonemic features and automatized processing evolve and mature, it is imperative to carefully consider all feasible explanations for our pattern of findings.

In addition to differences in high-gamma, the AEx group also showed a left-lateralized power increase in the beta range to the STD, which was not seen for the PEx or NC9 groups. Although much of the research on speech processing in adults and infants alludes to involvement of delta, theta, and gamma oscillations (Poeppel et al. 2008; Giraud and Poeppel 2012; Bosseler et al. 2013; Ortiz-Mantilla et al. 2013, 2016) several studies (Wang et al. 2012; Weiss and Mueller 2012; Bidelman 2015; Lewis et al. 2016; Mai et al. 2016) suggest a role for beta oscillations in speech processing, but none, to our knowledge, have discussed its role during syllable perception in infancy. In adults, oscillatory activity in the beta range has been related to higher-level language processing including language comprehension and detection of syntactic violations and semantic incongruence (Wang et al. 2012; Kiehl et al. 2015; Lewis et al. 2015, Lewis and Bastiaansen 2015, 2016). However, participants in this study were 9-month-old infants, processing the sub-lexical phonemic information contained in the syllable. Interestingly, it has also been suggested that beta oscillations increase if a language stimulus must be maintained in memory, as is the case for the standard sound presented in oddball paradigms and will decrease when a novel/deviant sound interrupts the regularity of the standard (Weiss and Mueller 2012). We speculate that the increase in beta power, shown by AEx infants, could also reflect a more “prepared” brain, ready to keep the memory trace of the standard stimulus in auditory memory. On the other hand, it has been reported that an increase in the beta range might well be associated with gamma band increases, given that beta oscillations occurring in lower frequency (13–20 Hz) ranges may have a functional role similar to alpha, while beta activation in the upper frequency ranges (20–30 Hz) may align with gamma activity (Spitzer and Haegens 2017).

Although it is not clear whether any (or perhaps all) of these theoretical possibilities might explain our findings, it's important to remember that all of the research cited above has been in adults. There is, however, a physiologically appealing and developmentally appropriate account that may explain the increase in beta power seen in the AEx group. This hypothesis references the modulatory effect of beta in linking motor and sensory information processing. Listening to speech involuntarily activates corresponding motor/articulatory patterns in

somatosensory cortex, particularly in the left hemisphere (Murakami et al. 2012; Bartoli et al. 2016). It has been proposed that modulation of beta activity to sound stimulation in auditory cortices and other motor-related areas may reflect auditory-motor communication even when movements are absent (Fujioka et al. 2012). During early development, not only are phonemic maps being established but also corresponding maps of paired correlated articulatory patterns (Bruderer et al. 2015). Therefore, the syllabic unit may play a pivotal role in evolving speech production as phonological knowledge is translated into motor gestures and babies begin babbling (Strauß and Schwartz 2017). The fact that only AEx infants showed left-lateralized oscillatory modulation in the beta band, a frequency range thought to represent a neural signature for motor activity, suggests advanced mapping of phonemic content and perhaps its corresponding motor pattern, a step required for speech production. Although the articulatory explanation for the increased beta observed in AEx infants is very appealing developmentally, it is clear that further research will be necessary to clarify just what role beta oscillations play during infant speech perception.

Oscillatory activity in the theta range also differed as a function of acoustic experience. Both groups, AEx bilaterally and PEx in left auditory cortex showed less theta power than NC9 infants when processing the STD. It is known that across age as infants become more automatized and efficient in processing their native language, oscillatory activity in the gamma range increases and less neuronal resources, indexed by amount of spectral power, are allocated in the theta range (Ortiz-Mantilla et al. 2016). These results, in which both experimental groups revealed what seems to be more mature processing as indexed by decreases in theta power, align with studies showing that over development, oscillatory activity at lower frequencies gradually shifts to higher frequency ranges (Koroleva et al. 2002; Marshall et al. 2002; Orekhova et al. 2006; Ortiz-Mantilla et al. 2016).

A different oscillatory pattern in the theta range was seen for DEV processing. As compared with NC9, both the AEx and PEx groups demonstrated increased theta power in right auditory cortex. Although in the opposite direction to what maturation usually induces in the theta range (i.e., a decrease in power), this finding might also be interpreted as more mature processing, as it accords with the proposed role of the right auditory cortex in sampling slow-rate information such as that contained in syllabic segments (Poeppel 2003; Abrams et al. 2008; Giraud and Poeppel 2012; Ghitza 2013; Vanvooren et al. 2014; Hyafil et al. 2015). Therefore, it seems that early (interactive or passive) acoustic experience may facilitate the more demanding processing required by a sporadically presented DEV, by using the right auditory cortex, an area primarily involved in decoding slow-rate syllabic information.

Lastly, we want to point out that the Benasich et al. (2014) study and the study detailed here are parallel and complementary given that the robust modulatory effects of early acoustic experience were demonstrated on auditory processing of non-speech complex tones (Benasich et al., 2014) and in this study on native speech syllables. Indeed, convergence between the two studies was particularly evident at the P1 sensory/perceptual level. When generalization effects to novel non-speech stimuli were examined at the high temporal resolution of ERPs, the Benasich et al. (2014) results included reports that the AEx group had significantly faster latencies for the P1 peak and

smaller, more mature amplitudes for the P1 and N1 peaks on the STD stimulus. In the present study, we used time-frequency analysis at the P1 source level, which increased the spatial resolution while maintaining the high temporal resolution, and further allowed us to expand the resolution of our analysis in the spectral domain. In line with our previous ERP findings using a non-speech paradigm to examine processing of the STD stimulus, in this study, temporo-spectral analysis of the STD syllable at the P1 source level revealed more mature processing for the AEx group, (i.e., allocation of less theta power in both left and right auditory sources) than their naïve controls at 9 months. Thus, these modulatory effects not only endured, but critically, generalized to speech.

We believe that by using both speech and non-speech stimuli while examining short-term as well as longer-lasting generalization effects of early acoustic experience, our findings from the Benasich et al. (2014) and Musacchia et al. (2017) studies are not only strengthened, but the impact of the original results is expanded by highlighting different but complementary aspects of infant brain plasticity.

Conclusions: Our results demonstrate that experience-dependent plasticity effects on processing of non-speech stimuli containing linguistically relevant acoustic cues, induced as a function of early interactive acoustic experience (IAE), generalize to speech and confer a significant processing advantage for syllabic processing. Specifically, 9-month-old infants that received early IAE in which attention was engaged and variation in stimulus speed and complexity was contingent on the infant's performance (AEx group), responded to a native STD phoneme with increased power in the high-gamma range, suggesting long-term phonemic representations were established in auditory cortex. This was not the case for infants with only passive exposure to the same sounds (PEx group) or for maturation alone (NC9 group). Thus, early interactive acoustic training may facilitate segmental processing and therefore, more efficient phonemic mapping. IAE also facilitated perception of rapid acoustic changes responding with increases in left gamma power to discrimination of the DEV syllable. Precise cortical representations and fine-grained discrimination abilities are essential as they favor faster encoding of speech information and efficient, more automatized processing of native language.

Early IAE modulated beta oscillations as well, perhaps suggesting advanced auditory-motor coupling, important to mapping phonemic content with the corresponding motor pattern, a step critical for speech production. Efficient, rapid and accurate language processing is a fundamental driver of strong linguistic and cognitive performance. Thus, the results presented here are very encouraging as they show that we can modulate, fine-tune and optimize acoustic processing even in typically developing infants with no demonstrated familial risk for developmental language disorders. For children who are at higher risk for DLD, access to an infant-friendly intervention that facilitates pre-linguistic acoustic mapping and generalizes to speech, raises the possibility of ameliorating or perhaps even preventing some DLDs and hopefully will translate into better language outcomes for all children.

Funding

This work was supported by the Elizabeth H. Solomon Center for Neurodevelopmental Research.

Notes

We would like to thank our colleagues at the Infancy Studies Laboratory at Center for Molecular and Behavioral Neuroscience Rutgers University-Newark, in particular Cindy P. Roesler, for their assistance during data collection and for their help revising and editing the manuscript. Special thanks go to all infants and families who participated in the study. *Conflict of Interest:* None declared.

References

- Abrams DA, Nicol T, Zecker S, Kraus N. 2008. Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. *J Neurosci.* 28:3958–3965.
- Aslin RN, Pisoni DB, Hennessy BL, Perey AJ. 1981. Discrimination of voice onset time by human infants: new findings and implications for the effects of early experience. *Child Dev.* 52:1135–1145.
- Bartoli E, Maffongelli L, Campus C, D'Ausilio A. 2016. Beta rhythm modulation by speech sounds: somatotopic mapping in somatosensory cortex. *Sci Rep.* 6:31182. doi:10.1038/srep31182.
- Benasich AA. 2002. Impaired processing of brief, rapidly presented auditory cues in infants with a family history of autoimmune disorder. *Dev Neuropsychol.* 22:351–372.
- Benasich AA, Choudhury N, Realpe-Bonilla T, Roesler CP. 2014. Plasticity in developing brain: active auditory exposure impacts prelinguistic acoustic mapping. *J Neurosci.* 34:13349–13363.
- Benasich AA, Tallal P. 2002. Infant discrimination of rapid auditory cues predicts later language impairment. *Behav Brain Res.* 136:31–49.
- Best CT, McRoberts GW, LaFleur R, Silver-Isenstadd J. 1995. Divergent developmental patterns for infants' perception of two non-native consonant contrasts. *Infant Behav Dev.* 18:339–350.
- Bidelman GM. 2015. Induced neural beta oscillations predict categorical speech perception abilities. *Brain Lang.* 141:62–69.
- 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. doi:10.3389/fpsyg.2013.00690.
- Bruderer AG, Danielson DK, Kandhadai P, Werker JF. 2015. Sensorimotor influences on speech perception in infancy. *Proc Natl Acad Sci.* 112:13531–13536.
- Chang EF, Rieger JW, Johnson K, Berger MS, Barbaro NM, Knight RT. 2010. Categorical speech representation in human superior temporal gyrus. *Nat Neurosci.* 13:1428–1432.
- Cheour M, Ceponiene R, Lehtokoski A, Luuk A, Allik J, Alho K, Näätänen R. 1998. Development of language-specific phoneme representations in the infant brain. *Nat Neurosci.* 1:351–353.
- Cheour M, Shestakova A, Alku P, Ceponiene R, Näätänen R. 2002. Mismatch negativity shows that 3–6-year-old children can learn to discriminate non-native speech sounds within two months. *Neurosci Lett.* 325:187–190.
- 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.
- Crone NE, Boatman D, Gordon B, Hao L. 2001. Induced electrocorticographic gamma activity during auditory perception. *Clin Neurophysiol.* 112:565–582.
- de Villers-Sidani E, Chang EF, Bao S, Merzenich MM. 2007. Critical period window for spectral tuning defined in the primary auditory cortex (A1) in the rat. *J Neurosci.* 27:180–189.
- Eimas PD, Siqueland ER, Jusczyk P, Vigorito J. 1971. Speech perception of rapid acoustic changes. *Science.* 171:303–306.
- Froemke RC, Jones BJ. 2011. Development of auditory cortical synaptic receptive fields. *Neurosci Biobehav Rev.* 35:2105–2113.
- Fujioka T, Trainor LJ, Large EW, Ross B. 2012. Internalized timing of isochronous is represented in neuromagnetic beta oscillations. *J Neurosci.* 32:1791–1802.
- Ghitza O. 2013. The theta-syllable: a unit of speech information defined by cortical function. *Front Psychol.* 4:138.
- Giraud AL, Poeppel D. 2012. Cortical oscillations and speech processing: emerging computational principles and operations. *Nat Neurosci.* 15:511–517.
- Guttorm TK, Leppänen PHT, Poikkeus A-M, Eklund K, 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.
- Hari R, Salmelin R. 1997. Human cortical oscillations: a neuro-magnetic view through the skull. *Trends Neurosci.* 20:44–49.
- Heim S, Friedman JT, Keil A, Benasich AA. 2011. Reduced sensory oscillatory activity during rapid auditory processing as a correlate of language-learning impairment. *J Neuroling.* 24:538–555.
- Hickok G, Poeppel D. 2007. The cortical organization of speech processing. *Nat Rev Neurosci.* 8:393–402.
- Hoehstetter K, Bornfleter 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.
- Hullett PW, Hamilton LS, Mesgarani N, Schreiner CE, Chang EF. 2016. Human superior temporal gyrus organization of spectrotemporal modulation tuning derived from speech stimuli. *J Neurosci.* 36:2014–2026.
- Hyafil A, Fontolan L, Kabdebon C, Gutkin B, Giraud AL. 2015. Speech encoding by coupled cortical theta and gamma oscillations. *eLife.* 2015(4):e06213. DOI:10.7554/eLife.06213.
- 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.
- Jusczyk PW, Pisoni DB, Walley A, Murray J. 1980. Discrimination of relative onset time of two-component tones by infants. *J Acoust Soc Am.* 67:262–270.
- Khalighinejad B, Cruzatto da Silva G, Mesgarani N. 2017. Dynamic encoding of acoustic features in neural responses to continuous speech. *J Neurosci.* 37:2176–2185.
- Kielar A, Panamsky L, Links KA, Meltzer JA. 2015. Localization of electrophysiological responses to semantic and syntactic anomalies in language comprehension with MEG. *Neuroimage.* 105:507–524.
- Kilgard MP, Pandya PK, Vazquez J, Gehi A, Schreiner CE, Merzenich MM. 2001. Sensory inputs direct spatial and temporal plasticity in primary auditory cortex. *J Neurophysiol.* 86:326–338.
- Koroleva NV, Nebera SA, Gutnik IN. 2002. Leading characteristics of maturity of the brain bioelectric activity in children 1–7 years of age. *Hum Physiol.* 28:687–692.
- Kuhl PK, Stevens E, Hayashi A, Deguchi T, Kiritani S, Iverson P. 2006. Infants show facilitation effect for native language phonetic perception between 6 and 12 months. *Dev Sci.* 9:F13–F21.

- 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.
- Lewis AG, Bastiaansen MCM. 2015. A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. *Cortex*. 68:155–168.
- Lewis AG, Schoffelen JM, Schriefers H, Bastiaansen M. 2016. A predictive coding perspective on beta oscillations during sentence-level language comprehension. *Front Hum Neurosci*. 10:85. Doi:10.3389/fnhum.2016.00085.
- Lewis AG, Wang L, Bastiaansen MCM. 2015. Fast oscillatory dynamics during language comprehension: unification versus maintenance and prediction? *Brain Lang*. 148:51–63. doi:10.1016/j.bandl.2015.01.003.
- 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. *PNAS*. 110:4846–4851.
- Mai G, Minett JW, Wang WSY. 2016. Delta, theta, beta and gamma brain oscillations index levels of auditory sentence processing. *Neuroimage*. 133:516–528.
- 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.
- Marshall PJ, Bar-Haim Y, Fox NA. 2002. Development of the EEG from 5 months to 4 years of age. *Clin Neurophysiol*. 113:1199–1208.
- Mesgarani N, Cheung C, Johnson K, Chang EF. 2014. Phonetic feature encoding in human superior temporal gyrus. *Science*. 343:1006–1010.
- 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.
- Murakami T, Restle J, Ziemann U. 2012. Effective connectivity hierarchically links temporoparietal and frontal areas of the auditory dorsal stream with the motor cortex lip area during speech perception. *Brain Lang*. 122:135–141.
- 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, 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.
- 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.
- Nawyn J, Roesler C, Realpe-Bonilla T, Choudhury N, Benasich AA. 2007. An operantly conditioned looking task for assessing infant auditory processing ability. In *Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility*. ASSETS'07 ACM, New York, NY, pp. 147–154.
- Nourski KV. 2017. Auditory processing in the human cortex: an intracranial electrophysiology perspective. *Laryngoscope Investig Otolaryngol*. 2:147–156.
- Nourski KV, Reale RA, Oya H, Kawasaki H, Kovach CK, Chen H, Howard MA 3rd, Brugge JF. 2009. Temporal envelope of time-compressed speech represented in the human auditory cortex. *J Neurosci*. 29:15564–15574.
- Nourski KV, Steinschneider M, Oya H, Kawasaki H, Howard MA III. 2015. Modulation of response patterns in human auditory cortex during a target detection task: an intracranial electrophysiology study. *Int J Psychophysiol*. 95:191–201.
- Näätänen R, Winkler I. 1999. The concept of auditory stimulus representation in cognitive neuroscience. *Psychol Bull*. 125:826–859.
- Orehova EV, Stroganova TA, Posikera IN, Elam M. 2006. EEG theta rhythm in infants and preschool children. *Clin Neurophysiol*. 117:1047–1062.
- 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.
- Poeppl D. 2003. The analysis of speech in different temporal integration windows: cerebral lateralization as ‘asymmetric sampling in time’. *Speech Commun*. 41:245–255.
- Poeppl D, Idsardi WJ, van Wassenhove V. 2008. Speech perception at the interface of neurobiology and linguistic. *Philos Trans R Soc Lond B Biol Sci*. 363:1071–1086.
- 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.
- Shestakova A, Huottilainen M, Ceponiene R, Cheour M. 2003. Event-related potentials associated with second language learning in children. *Clin Neurophysiol*. 114:1507–1512.
- Spitzer B, Haegens S. 2017. Beyond the status quo: a role for beta oscillations in endogenous content (re) activation. *eNeuro*. 4:ENEURO.0170–17.2017.
- Steinschneider M, Fishman YI. 2011. Enhanced physiological discriminability of stop consonants with prolonged format transitions in awake monkeys based on the tonotopic organization of primary auditory cortex. *Hear Res*. 271:103–114.
- Steinschneider M, Nourski KV, Fishman YI. 2013. Representation of speech in human auditory cortex: is it special? *Hear Res*. 305:57–73.
- Steinschneider M, Nourski KV, Kawasaki H, Oya H, Brugge JF, Howard MA III. 2011. Intracranial study of speech-elicited activity on the human posterolateral superior temporal gyrus. *Cereb Cortex*. 21:2332–2347.
- Strauß A, Schwartz JL. 2017. The syllable in the light of motor skills and neural oscillations. *Lang Cogn Neurosci*. 32:562–569.
- 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.
- Threlkeld SW, Hill CA, Rosen GD, Fitch RH. 2009. Early acoustic discrimination experience ameliorates auditory processing deficits in male rats with cortical developmental disruption. *Int J Dev Neurosci*. 27:321–328.

- 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.
- Vanvooren S, Poelmans H, Hofmann M, Ghesquiere P, Wouters J. 2014. Hemispheric asymmetry in auditory processing of speech envelope modulations in prereading children. *J Neurosci.* 34:1523–1529.
- Wang L, Jensen O, van den Brink D, Weder N, Schoffelen JM, Magyari L, Hagoort P, Bastiaansen M. 2012. Beta oscillations relate to the N400m during language comprehension. *Hum Brain Mapp.* 33:2898–2912.
- Weiss S, Mueller HM. 2012. “Too many betas do not spoil the broth”: the role of beta brain oscillations in language processing. *Front Psychol.* 3:201. doi:10.3389/fpsyg.2012.00201.
- 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.
- Xu L. 2005. Relative contribution of spectral and temporal cues for phoneme recognition. *J Acoust Soc Am.* 117:3255–3267.
- Zatorre RJ, Belin P. 2001. Spectral and temporal processing in human auditory cortex. *Cereb Cortex.* 11:946–953.