

Developmental changes in auditory-evoked neural activity underlie infants' links between language and cognition

Kali Woodruff Carr^{1,2}  | Danielle R. Perszyk¹ | Elizabeth S. Norton^{2,3} |
Joel L. Voss^{3,4} | David Poeppel^{5,6} | Sandra R. Waxman^{1,3,7} 

¹ Department of Psychology, Northwestern University, Evanston, Illinois, USA

² Department of Communication Sciences and Disorders, Northwestern University, Evanston, Illinois, USA

³ Institute for Innovations in Developmental Sciences, Northwestern University, Chicago, Illinois, USA

⁴ Department of Medical Social Sciences, Ken and Ruth Davee Department of Neurology, Department of Psychiatry and Behavioral Sciences, Interdepartmental Neuroscience Program, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA

⁵ Department of Neuroscience, Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany

⁶ Department of Psychology and Center for Neural Science, New York University, New York, New York, USA

⁷ Institute for Policy Research, Northwestern University, Evanston, Illinois, USA

Correspondence

Sandra R. Waxman, Department of Psychology, Northwestern University, 2029 Sheridan Road, Evanston, IL 60208-2710.

Email: s-waxman@northwestern.edu

Present address

Danielle R. Perszyk, Google, 1600 Amphitheatre Parkway, Mountain View CA, USA.

Abstract

The power and precision with which humans link language to cognition is unique to our species. By 3–4 months of age, infants have already established this link: simply listening to human language facilitates infants' success in fundamental cognitive processes. Initially, this link to cognition is also engaged by a broader set of acoustic stimuli, including non-human primate vocalizations (but not other sounds, like backwards speech). But by 6 months, non-human primate vocalizations no longer confer this cognitive advantage that persists for speech. What remains unknown is the mechanism by which these sounds influence infant cognition, and how this initially broader set of privileged sounds narrows to only human speech between 4 and 6 months. Here, we recorded 4- and 6-month-olds' EEG responses to acoustic stimuli whose behavioral effects on infant object categorization have been previously established: infant-directed speech, backwards speech, and non-human primate vocalizations. We document that by 6 months, infants' 4–9 Hz neural activity is modulated in response to infant-directed speech and non-human primate vocalizations (the two stimuli that initially support categorization), but that 4–9 Hz neural activity is not modulated at either age by backward speech (an acoustic stimulus that doesn't support categorization at either age). These results advance the prior behavioral evidence to suggest that by 6 months, speech and non-human primate vocalizations elicit distinct changes in infants' cognitive state, influencing performance on foundational cognitive tasks such as object categorization.

KEYWORDS

conceptual development, developmental tuning, infant cognition, infant EEG, infant ERP, language acquisition

1 | INTRODUCTION

The power of human language derives, at least in part, from its inextricable relationship with human cognition. Recent behavioral evidence reveals that well before they begin to speak, infants have begun to establish the link between language and cognition. For infants as young as 3 and 4 months of age, pairing human language with images of objects from a shared category supports infant's ability to form object categories in a way that many other sounds do not (Ferry et al., 2010;

Fulkerson & Waxman, 2007; Perszyk & Waxman, 2019; Woodruff Carr et al., 2021). Still, speech is not the only acoustic stimulus that boosts very young infants' object categorization. Vocalizations produced by non-human primates (e.g., blue-eyed black lemur; *Eulemur macaco flavifrons*) also support object categorization for 3- and 4-month-olds (Ferry et al., 2013). However, in contrast to speech, which supports categorization consistently throughout infancy, the advantage conferred by listening to non-human primate vocalizations is short-lived. By 6 months of age, infants have tuned this link: non-human primate

vocalizations no longer support infant categorization (Ferry et al., 2013).

The evidence that human speech and non-human primate vocalizations boost cognition in very young infants comes primarily from a behavioral task designed to assess infants' success in forming object categories, a fundamental building block for cognition (Aslin, 2007; Colombo, 2002; Hirsh-Pasek et al., 1987; Murphy, 2004; Quinn et al., 2019; Waxman & Gelman, 2009). The categorization task includes two phases. First, during a familiarization phase, infants view a series of distinct objects (e.g., images of eight discriminably different dinosaurs), all members of the same object category (e.g., dinosaur), each presented in conjunction with an auditory signal (e.g., speech). Next, in a test phase, infants view two new objects, one a new member of the now-familiar category (for this example, a different dinosaur) and the other a member of a different category (e.g., a fish). Infants' looking to the test images serves as the dependent measure: if infants detect the category-based commonalities among the familiarization objects, then at test their looking behavior should distinguish between the two test objects; if they fail to detect the object category, then they should perform at chance levels.

This behavioral paradigm has lent itself to investigations of whether and how speech and other acoustic stimuli influence infants' object categorization. By holding constant the objects infants view, while systematically manipulating which sounds accompany the familiarization objects, researchers have been able to conclude that differences in infants' object categorization success must reflect differences elicited by different types of sounds. This collection of work has revealed that both human speech and non-human primate vocalizations support very young infants' object categorization in a way that other acoustic stimuli, such as sine-wave tone sequences (Ferry et al., 2010; Fulkerson & Waxman, 2007), time-reversed ("backwards") speech (Ferry et al., 2013), and other animal vocalizations (e.g., birdsong; Woodruff Carr et al., 2021), do not.

What remains unknown is how human speech and non-human primate vocalizations exert their initial influence on infant cognition, and the mechanism by which this behavioral link is rapidly tuned to human speech. Certainly by 12 months, language may initiate a search to discover a word's meaning (Carey, 2011; Waxman & Markow, 1995). But in the first few months of life—when infants neither reliably glean distinct words from the ongoing stream of speech (Jusczyk, 1999; Saffran et al., 1996), nor infuse acoustic representations of words with meaning (Bergelson & Swingle, 2012; Tincoff & Jusczyk, 2012)—a different, perhaps coarser mechanism must be at play.

One possibility is that for very young infants, human speech and non-human primate vocalizations elicit heightened attention, effectively highlighting the objects and events in their current surroundings and, in this way, supporting performance on fundamental cognitive tasks including object categorization [for review of this idea, see (Perszyk & Waxman, 2018)]. Attention is a multifaceted set of processes which can be difficult to measure precisely and disentangle, even in adults. This challenge is even greater in infants, whose rudimentary attentional capacities develop rapidly [for reviews, see (de Diego-Balaguer et al., 2016; Gomes et al., 2000; Oakes & Amso, 2018)].

Research Highlights

- For infants as young as 3 months, both human speech and non-human primate vocalizations facilitate successful object categorization.
- By 6 months, non-human primate vocalizations no longer exert this advantageous effect on infant cognition, and only human speech supports successful object categorization.
- Here, we provide evidence of emerging differences in 4–9 Hz neural activity in response to recordings of human speech and non-human primate vocalizations.
- These findings suggest that these sounds differentially engage early components of infants' attention, offering a potential mechanism that supports developmental tuning of the language-cognition link.

In the first months, when infants' attention is largely under exogenous control, salient stimuli capture their attention automatically. This strong exogenous attentional pull is likely valuable, directing infants towards key stimuli in their environment, including the communicative cues of their caregivers [for review, see (de Diego-Balaguer et al., 2016)]. Following this line of evidence, we have hypothesized that when human speech and non-human primate vocalizations are presented in conjunction with objects, very young infants' attention to the objects is heightened, supporting the process of object categorization (Perszyk & Waxman, 2018). In time, as infants increasingly develop endogenous control over their focus of attention, they increasingly ignore salient distractors [for visual attention, this develops during the second half of the first year: (e.g., Althaus & Mareschal, 2012; Kwon et al., 2016; Tummeltshammer et al., 2014)]. This increase in endogenous control of attention may guide infants to devote less attention to objects that co-occur with less familiar and less informative acoustic stimuli, such as non-human primate vocalizations, than those presented in conjunction with human speech.

In the current study, we focus on three different kinds of sounds (human speech, non-human primate vocalizations, and backwards speech) for which the influence on infant categorization has been established by prior work (Ferry et al., 2010, 2013). Our goal here is to identify how infants' neural responses to these sounds change between 4- and 6-months, when the behavioral influence of non-human primate vocalizations fades. Our analyses focus especially on how these sounds modulate markers of infant attention. Electroencephalographic (EEG) activity was recorded in healthy, full-term 4- and 6-month-olds as they were presented with 3-s recordings of human speech (infant-directed speech; IDS), non-human primate vocalizations (LEMUR), and backwards speech (BW-IDS) over a single 20-min session. Three different tokens of each stimulus type (matched for mean frequencies and durations; see Figure 1 for exemplar tokens of each stimulus type; see Figure S1 and Table S1 for details on all nine tokens) were presented in

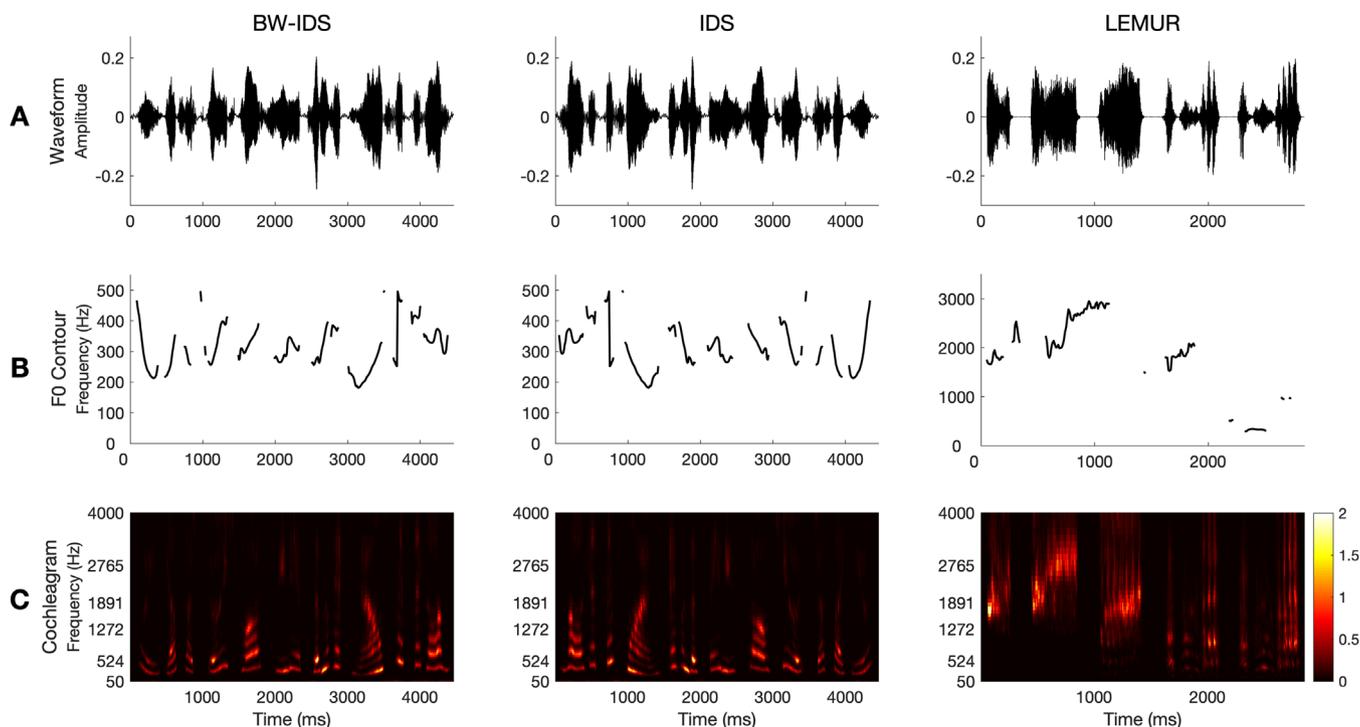


FIGURE 1 Neural responses were collected to three types of sounds: backwards infant-directed speech (BW-IDS), infant-directed speech (IDS), and non-human primate vocalizations (LEMUR). Nine different recordings were presented (3 different tokens of each stimulus type). Plotted here are example tokens of each stimulus type: (A) time-domain waveforms; (B) Fundamental frequency (F0) contours extracted using *Praat* (Boersma & Weenink, 2001) with frequency ranges optimized for F0 range of each stimulus type (75–500 Hz for human speech; 100–3000 for non-human primate vocalizations); and (C) cochleograms revealing changes in frequency across time, extracted using *gammatonegram.m* (Ellis, 1998) to construct the weighting matrix to convert the time-frequency spectrograms into gammatone-filter approximations to match the ear's frequency sub-bands. See Figure S1 for plots of all 9 stimuli tokens

a pseudo-randomized order to ensure that no two recordings of the same stimulus type were presented successively.

There are several methodological advantages of implementing this listening paradigm, rather than collecting neural responses as infants participate in the behavioral task. For one, this listening paradigm allows us to collect enough trials for each stimulus to maximize the signal-to-noise ratio (the object categorization task includes only 8 trials per infant, which would be problematic for event-related EEG comparisons) and to compensate for lost trials (e.g., due to movement-related artifacts, common in young infants). Second, we are able to minimize head-turn- and eye-movement-related artifacts that are elicited by the object categorization task. And finally, this paradigm also permits within-subjects comparisons of infants' neural responses to different acoustic stimuli, as compared to the object categorization task in which only one acoustic stimulus can be presented per infant.

Our analyses focus primarily on slow neural activity, below 10 Hz, in the frequency band interpreted as infant alpha. We focus on this frequency band because alpha activity is the predominant rhythm in the infant brain, it is detectable in infants as young as 3 months (Davidson & Fox, 1982; Lindsley, 1939; J. R. Smith, 1938, 1939), and it is thought to index infant cognitive engagement [e.g., attention (Orekhova et al., 2001; Stroganova et al., 1999; Xie et al., 2018) and working memory (Bell, 2002; Libertus et al., 2009)]. Importantly, a decrease in alpha power is typically observed under conditions of increased task

demands in children and adults and is associated with attention and functional inhibition [e.g., (Berger, 1929; Klimesch, 2012; Thut et al., 2006; Van Diepen et al., 2019)]. A decrease in activity below 10 Hz has also been observed in infants during conditions with increased attentional demands, including those requiring sustained visual attention (Stroganova et al., 1999; Xie et al., 2018). In contrast, an increase in activity below 10 Hz has been observed under conditions requiring infants to suppress attention to task-irrelevant information (Orekhova et al., 2001). This pattern of change in alpha activity has even been observed during passive listening conditions, like the listening task used here, that require no overt response. For example, a decrease in neural activity below 10 Hz was observed in 2-month-olds in response to speech, relative to silence (Mai et al., 2014), and in school-aged children and adults in response to speech or musical instrument sounds, compared to an increase in alpha activity in response to backwards speech (Fujioka & Ross, 2008; Krause et al., 1997). Together, these findings led us to predict that acoustic stimuli that support infant categorization would elicit a decrease in neural activity below 10 Hz, while acoustic stimuli that fail to support categorization would elicit an increase infants' neural activity below 10 Hz.

Auditory-evoked event-related potentials (ERPs) were also analyzed. Early components of the auditory-evoked ERP waveform have been shown to index developmental changes in neural processing of speech in infants (Dehaene-Lambertz & Dehaene, 1994; Nacar Garcia

et al., 2018; Peña et al., 2010) and are thought to reflect gains in neural maturation and experience (Kushnerenko et al., 2002). Based on these findings, we predicted that developmental changes in speech processing may also be evident in the early components of the ERP due to neural maturation and infants' increasing experience with speech.

Together, these spectral and temporal neural measures will provide insight into how infants' cognitive and attentional states may be modulated by acoustic stimuli that exert differing effects on infants' object categorization at 4 and 6 months of age. If low-frequency brain activity does indeed prove a robust neural marker for developmental changes with compelling cognitive consequences, it may be useful for investigating a range of hypothesized mechanisms that underlie the acquisition of our species' tight link between language and cognition.

2 | METHODS

2.1 | Participants

34 full-term infants were included in the final analyses (17 4-month-olds, range = 4.11–5.0 [6F]; 17 6-month-olds, range = 5.95–6.98 [9F]) [c.f. similar EEG studies with infants in this age range that report data from: 15 infants from 3 to 6 months old (Peña et al., 2010); 13–17 infants per age group, from 6 to 12 months old (Xie et al., 2018); 17 6- and 12-month-olds (Bosseler et al., 2013); 15 4-month-olds (Purhonen et al., 2005)]. 28 additional infants participated but were excluded because of contributing too few (< 15) trials (8 4-month-olds, 13 6-month-olds) or technical difficulties ($N = 7$). All participants were recruited from the Chicago area, using the same recruitment strategies as in prior behavioral work with infants in this age range (Ferry et al., 2010, 2013). As in prior work, this resulted in a sample of infants of predominantly white, college-educated families with high-to middle-socioeconomic status. Parents completed a questionnaire about language exposure, and the average parent-estimated exposure to English was 87% (SD: 21.4%). Informed consent was obtained from legal guardians in accordance with procedures approved by the Northwestern University Institutional Review Board, and participants were compensated for their participation with either monetary payment or a children's book and t-shirt.

2.2 | Stimuli

Acoustic stimuli consisted of nine unique recordings (three per stimulus type) of infant-directed speech (IDS), backward IDS, and non-human primate vocalizations (LEMUR). All recordings are available upon request; exemplar recordings are visualized in Figure 1 and all 9 recordings are visualized in Figure S1. Each IDS recording comprised two English sentences, spoken by a different native American English-speaking woman. Backward IDS (BW-IDS) recordings were the same 3 recordings of IDS, played in reverse. Recordings of lemur vocalizations were selected from recordings made by Chris Mercer at the Duke University Lemur Center to match the mean duration and frequency range

of the IDS recordings (see Table S1 and Figure S1 for further information and analyses of all recordings). Recordings were presented with an inter-stimulus interval (ISI) of 1800–2200 ms in sound field at 65 dB SPL, in a pseudorandomized order such that no two sounds of the same stimulus type (IDS, BW-IDS, or LEMUR) occurred twice in a row. Each recording was presented 24 times for a total of 216 trials (72 trials per stimuli type).

2.3 | Electroencephalography (EEG) recording parameters

All EEG recordings were made in a quiet room. Infants sat on a caretaker's lap facing two speakers placed 4 ft. in front of the infant and 2 ft. from one another. During the recording, an experimenter kept the infant calm and engaged by blowing soap bubbles and/or silently playing with puppets in front of the infant. The study was stopped at the completion of all 216 trials, or when the infant exhibited discomfort (e.g., crying, fussiness). The total duration of the full experiment was about 25 min.

2.4 | Data acquisition and pre-processing

Cortical responses were recorded with a sampling rate of 250 Hz using a Brain Products actiCAP EEG system (actiCAP, Brain Products GmbH, Gilching, Germany). Seventeen Ag/AgCl active electrodes were placed according to the 10–20 system, with an additional grounding electrode at Fpz. Data were online referenced to the electrode at Cz per the hardware configuration of the Brain Products actiCAP. Individual electrode impedances were ascertained to be below 10 k Ω prior to commencement of recording. Recordings were amplified (V-amp, Brain Products GmbH), online band-pass filtered (high-pass cutoff: 0.05 Hz; low-pass cutoff: 100 Hz), and notch filtered (centered at 60 Hz). Electrophysiological data were processed offline in MATLAB (The Mathworks, Inc., Natick, MA) using EEGLAB v.14.1.2 (Delorme & Makeig, 2004) and ERPLAB v.7.0.0 (Lopez-Calderon & Luck, 2014). The online reference electrode (Cz) was returned to the recording mathematically and re-referenced offline via linear transformation [see (Cohen, 2014)]. Due to hardware constraints, it was not possible to record from additional earlobe or mastoid references. Therefore, to avoid the confounding factor of the location of the reference electrode, all electrodes were re-referenced offline to the average of all electrodes to obtain an unbiased estimate of noise across the scalp. Although we had only 17 active electrode locations, 4- and 6-month-old infants' heads are relatively small, and the inter-electrode distance of the actiCAP (~4–5 cm, depending on head size) was similar to recommendations for inter-electrode spatial coverage to permit an whole-head average reference [c.f. recommended ~3–4 cm inter-electrode distance per (DeBoer et al., 2007)]. Continuous data were high-pass filtered at 0.1 Hz using a second order IIR Butterworth filter and segmented into –800 to 3000 ms epochs relative to stimulus onset (at 0 ms). Epochs containing artifacts such as eyeblinks or excessive noise were first detected using an



automated moving 200 ms window peak-to-peak amplitude threshold of 200 μV . Then all data were visually inspected, and artifacts confirmed before rejection. Additional artifact-contaminated epochs were manually detected, and all contaminated epochs were excluded from further analysis. Participants with fewer than 15 artifact-free trials per stimulus type (BW-IDS, IDS, LEMUR) were excluded from analyses (based on similar studies with infants in this age range that required a minimum of 10 trials (Friedrich & Friederici, 2011; Nacar Garcia et al., 2018) or 14 trials (Friedrich & Friederici, 2017; Friedrich et al., 2015) per stimulus type or condition). The mean number of artifact-free trials per stimulus type in the included participant groups was 27–30. See Table S2 for further information regarding artifact-free trials.

2.5 | Event-related spectral perturbation (ERSP)

To compute the total (evoked and induced) event-related spectral perturbations (ERSPs), the EEG time series of each epoch, at each electrode, was convolved with a set of complex Morlet wavelets, defined as a Gaussian-windowed complex sine wave: $e^{-i 2\pi t f} e^{-\frac{t^2}{2\sigma^2}}$, where t is time, f is frequency (which increased from 3–55 Hz centered over 53 linear-spaced 1 Hz-wide frequency bins, centered at each single-integer frequency), and σ defines the width (or “cycles”) of each frequency band, beginning with 3 cycles at the lowest frequency, and increasing by a factor of 0.92, to 4.4 cycles at the highest frequency (using eeglab’s *newtimef* function). The lower frequency bound was set at 3 Hz to allow for 3 cycles within the length of the epoch to adequately capture power at that frequency using the wavelet approach. Due to the temporal width of the wavelet at the lowest frequency and the sampling rate, the time-frequency analytical window was restricted to –242 ms to 2500 ms with respect to the stimulus onset. The pre-stimulus baseline of 242 ms was subtracted on a frequency-by-frequency basis for each trial. The power time series were normalized by conversion to a decibel (dB) scale ($10 \cdot \log_{10}[\text{power}(t)/\text{power}(\text{baseline})]$) to allow for direct comparison of effects across frequency bands. Epochs were then averaged for each participant across trials and stimulus tokens for each stimulus type at each electrode, and then averaged across participants for each age group.

2.6 | Identifying the alpha response in 4- and 6-month-old infants

While the canonical alpha band in the adult brain is 8–12 Hz, alpha activity in the infant brain is predominantly reflected in lower frequencies. For infants in the first year of life, there is little consensus regarding the frequency cutoffs for discriminating alpha from lower (theta) and higher (beta) activity. This is because the alpha EEG rhythm, which first emerges around 3- to 4-months, undergoes substantial changes in its frequency range with development: from ~3–5 Hz in 3–4 month-olds to ~6–10 Hz by 12-months (Lindsley, 1939; Marshall et al., 2002; Saby & Marshall, 2012; Stroganova et al., 1999). Therefore, in order

to assess changes in neural activity across the dynamic developmental period between 4 and 6 months, we compared responses averaged across a comprehensive range (4–9 Hz) based on the literature for both age groups, rather than selecting different frequency ranges for each age group. Power at each frequency can be observed in Figures 2 & 3, confirming that stimulus-driven changes in infants’ neural responses were largely captured by this literature-based frequency range.

Alpha in infants is typically maximal over central-parietal sites (Lindsley, 1939; J. R. Smith, 1939, 1941), so analyses were based on the average of the recordings from the two midline centroparietal channels, Cz and Pz, as many of the recordings from the occipital electrodes were quite noisy due to poor cap fit at those channels.

Given the consistency of the response at each individual 1 Hz frequency bin over the entire 4–9 Hz range (see Figures 2 & 3 and Table S3), and the concentration of the effect over centroparietal regions, we interpret this activity as largely representing an alpha-like response. Although this 4–9 Hz range may also include theta-range activity, theta activity typically exhibits the opposite pattern of activity than alpha, and it has a different topographical distribution, with activity maximal over frontal regions in infants.

Visual inspection of event-related spectral perturbation (ERSP) plots and permutation t-tests (computed for each group in response to each stimulus type compared to pre-stimulus baseline; Figure S2) confirmed that power within this frequency band was largely similar for each stimulus type over the majority of the response period (0–2500 ms), so subsequent analyses were computed on the average power over the entire time window of the response.

2.7 | Event-related potential (ERP)

In order to determine the appropriate time windows and electrodes for analysis, we first averaged across all stimulus types. Artifact-free epochs were averaged across all artifact-free trials of all stimuli types, baseline corrected to a pre-stimulus period of 200 ms, low-pass filtered at 20 Hz, then averaged across participants for each age group, and visualized using ERPLAB (Figure 4). A large anterior positivity and posterior negativity were observed at ~400 ms for both ages. Following prior work (Nacar Garcia et al., 2018; Peña et al., 2010), electrodes over central-anterior (F3, Fz, F4) and central-posterior (P3, Pz, P4) regions were selected *a priori* as spatial regions of interest (ROIs) for further comparisons. ERPs were then re-averaged separately for each stimulus type at each electrode for each participant, baseline corrected to a pre-stimulus period of 200 ms, and low-pass filtered at 20 Hz. Stimulus-specific averages were then averaged across electrodes of each ROI, and then averaged across participants for each age group. The mean amplitude of the ERP component was computed for each ROI centered around the peak observed in the grand average, over the 200–600 ms time window. Permutation t-tests were also computed to compare developmental changes in the response to each stimulus and confirm the time window of interest (Figure S3).

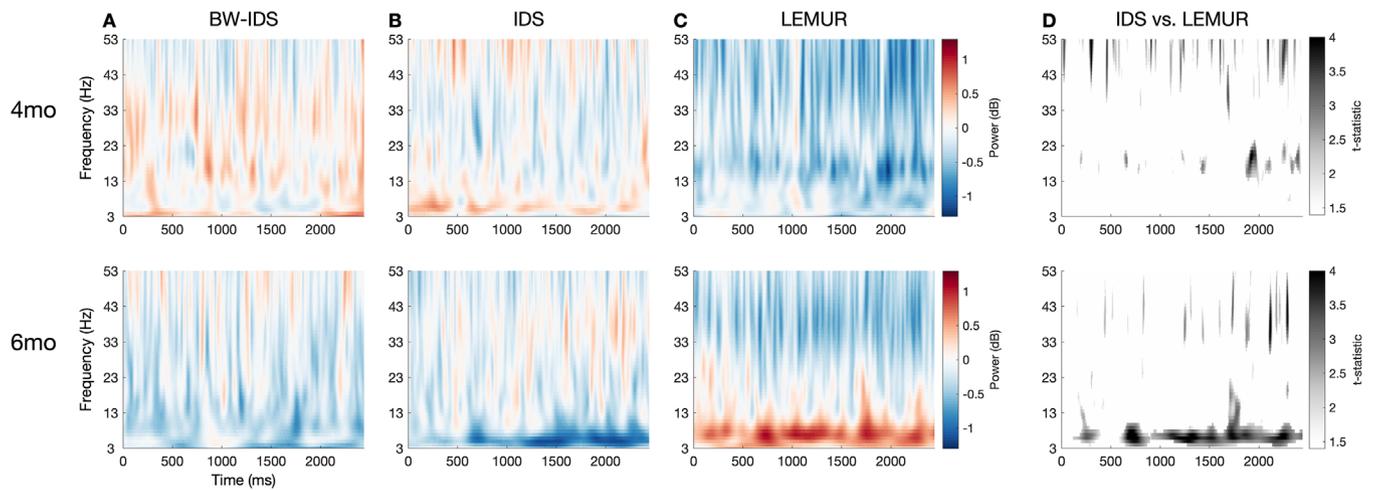


FIGURE 2 Event-related spectral perturbations (ERSPs) reveal the mean total (evoked and induced) changes in event-related EEG spectral power (in dB), in 4- ($N = 17$; upper panels) and 6-month-olds ($N = 17$; lower panels) elicited by (A) backwards speech (BW-IDS), (B) infant-directed speech (IDS), and (C) non-human primate vocalizations (LEMUR). ERSPs have been baseline corrected to the pre-stimulus period on a trial-by-trial basis for each frequency, such that 0 dB represents no change from baseline. ERSPs at Cz and Pz were comparable, so mean power at each time-frequency point were averaged. (D) Permutation t-tests were computed on each time-frequency point of the ERSPs elicited by IDS and LEMUR and masked to plot only significant ($p < 0.05$) differences

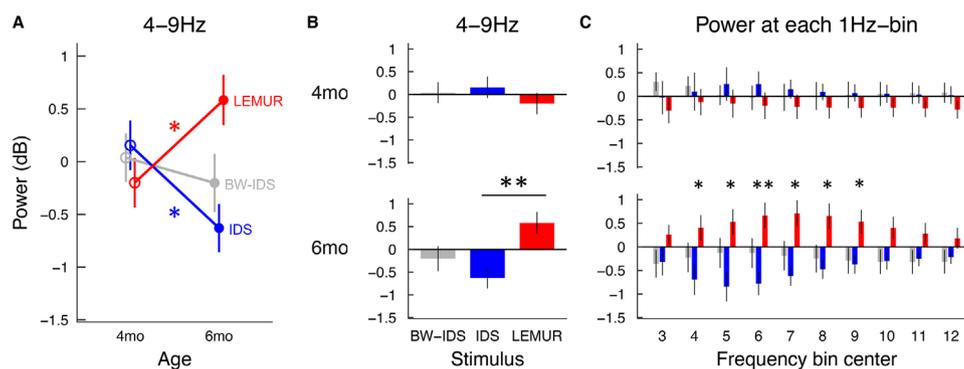


FIGURE 3 Changes in infant event-related neural power (as compared to pre-stimulus activity; in dB) was averaged across the entire response time region (0-2500 ms) and across centroparietal electrodes (Cz and Pz). (A) Between 4- and 6-months, infants' 4-9 Hz power decreased in response to infant-directed speech (IDS; blue) and increased in response to non-human primate vocalizations (LEMUR; red) but did not significantly change in response to backwards speech (BW-IDS; gray). (B) This developmental difference was largely due to differences in the 6-month-olds' responses to IDS and LEMUR (lower panel). (C) This effect was consistent across each individual frequency (1 Hz integer-centered bins) within the broader low-frequency region identified by the permutation tests in Figure 2D. ** = $p < 0.01$

2.8 | Statistical analysis

To define regions of interest, time- and time-frequency permutation-based t-tests were computed in MATLAB using *statcond* (see Figures S2 & S3). Within- and across-subject comparisons were computed using R v.3.5.2 (R Core Team, 2018) and SPSS v.26 (SPSS, Inc.). All dependent variables conformed to the expectations of linear models (Levene's test for the homogeneity of variances and Mauchly's test of sphericity). While the comparison of a number of electrodes and frequencies imposes a problem of correction for multiple comparisons, the application of such a correction would greatly inflate the probability of false negative results due to the strong correlation of activity in adjacent electrodes and frequency bins (i.e., multicollinear-

ity). Therefore, statistics are reported on averaged electrodes and frequency bins in the main results section with uncorrected p-values. Statistics for comparisons of individual, 1-Hz width, frequency bins can be found in Figure 3B and Table S3. Statistical analyses and figures from both individual electrodes of interest (Cz and Pz) are also presented in the supplemental materials (see Supplemental Results and Figures S4-S7).

2.9 | Stimuli and data sharing

All stimuli and data not under copyright restrictions are posted at <https://osf.io/pmt46/>.

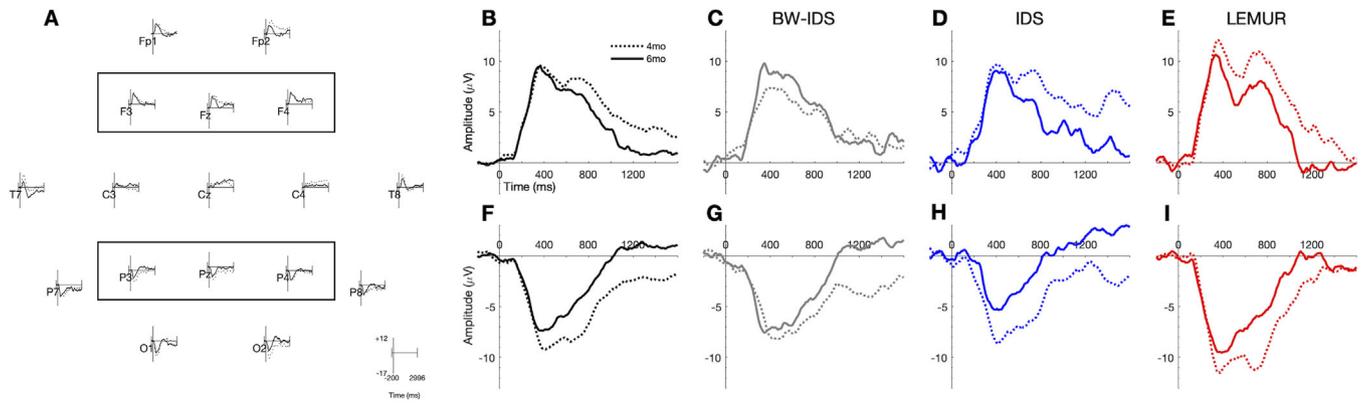


FIGURE 4 Event-related potentials (ERPs) were recorded in 4- (dotted line) and 6-month-olds (solid line) in response to the three acoustic stimuli: BW-IDS (gray), IDS (blue), and LEMUR (red). (A) Responses were averaged across stimuli (black lines), revealing a peak ~400 ms at (B-E) anterior (F3, Fz, F4) and (F-I) posterior (P3, Pz, P4) electrodes. The amplitude of this component decreased with age only in the posterior region of interest (F). Follow-up comparisons of the mean amplitude of this component (200–600 ms) for the responses to each stimulus type revealed that a maturational reduction in the amplitude of this ERP component occurred only in response to IDS (H)

3 | RESULTS

3.1 | Developmental changes in centroparietal 4–9 Hz EEG power in response to speech and non-human vocalizations

There was a significant interaction between age (4- vs. 6-month-olds) and stimulus type (BW-IDS; IDS; LEMUR) for event-related 4–9 Hz EEG power (repeated measures ANOVA: $F_{(2,64)} = 4.157$, $p = 0.020$, $\eta_p^2 = 0.115$; Figure 3A). This effect was specific to this frequency band; it did not occur at higher frequencies (see Figure 2). Although there were no main effects of either age ($F_{(1,32)} = 0.445$, $p = 0.510$) or stimulus type ($F_{(2,64)} = 1.241$, $p = 0.296$) on 4–9 Hz power, planned follow-up comparisons for each stimulus type did reveal significant differences in 4–9 Hz power between 4- and 6-month-olds in response to IDS and LEMUR, the two stimuli that have been shown to initially support infant categorization. As predicted, 4–9 Hz power *decreased* between 4- and 6-months in response to IDS (4mo: $M = 0.155$; $SD = 0.972$; 6mo: $M = -0.629$; $SD = 0.941$; ANOVA: $F_{(1,32)} = 5.709$, $p = 0.023$, $d = 0.820$). In contrast, 4–9 Hz power *increased* between 4- and 6-months in response to LEMUR (4mo: $M = -0.199$; $SD = 0.965$; 6mo: $M = 0.584$; $SD = 0.984$; ANOVA: $F_{(1,32)} = 5.485$, $p = 0.026$, $d = 0.803$). There was no change between 4- and 6-months in response to BW-IDS, a stimulus that fails to support object categorization at both ages (4mo: $M = 0.038$, $SD = 0.945$; 6mo: $M = -0.202$, $SD = 1.138$; ANOVA: $F_{(1,32)} = 0.447$, $p = 0.508$).

Planned follow-up comparisons of each age group's response to IDS and LEMUR revealed that, as expected, in 4-month-olds, there were no significant differences in 4–9 Hz power in response to IDS and LEMUR, the two types of stimuli that both support infant categorization at 4 months (paired t-test, IDS vs. LEMUR: $t(16) = 0.939$, $p = 0.362$). However, in 6-month-olds, for whom the cognitive advantage of IDS persists but the advantage of LEMUR has faded, 4–9 Hz power in response to these two stimuli differed significantly (paired t-test, IDS vs. LEMUR: $t(16) = 3.297$, $p = 0.005$, $d = 0.799$; Figure 3B). In

response to IDS, there was an event-related *decrease* in 4–9 Hz power relative to pre-stimulus baseline (one-sample t-test: $t(16) = 2.757$, $p = 0.014$, $d = 0.668$). In response to LEMUR, there was an event-related *increase* in power relative to pre-stimulus baseline (one-sample t-test: $t(16) = 2.445$, $p = 0.026$, $d = 0.594$). Examining the individual data revealed that 12 out of 17 participants (70.6%) showed this effect (Figure S8).

Permutation t-tests were performed for each time-frequency point to further visualize which time-frequency regions best captured differences in EEG power between the responses to IDS and LEMUR stimuli at each age (Figure 2D; masked to plot only $p < 0.05$). Follow-up comparisons of power at each 1-Hz, integer-centered, frequency bin also revealed that the effect was consistent across frequencies within this (literature-based) bin and was consistent for the duration of the stimulus (Figure 3C and Table S3; individual data presented in Figure S9).

3.2 | Developmental changes in the event-related potential (ERP) in response to speech

Previous studies of 3- to 9-month-old infants' ERPs in response to speech sounds describe a positive anterior peak (~200–300 ms) and a negative posterior (~240 ms) peak (Nacar Garcia et al., 2018; Peña et al., 2010). Here, we observed the large initial peak slightly later, around 400 ms, over both anterior (F3, Fz, F4; Figure 4B) and posterior (P3, Pz, P4; Figure 4F) regions. The mean amplitude of this peak (averaged over 200–600 ms) differed as a function of stimulus type only for the posterior ROI (repeated measures ANOVA: $F_{(2,64)} = 6.420$; $p = 0.003$, $\eta_p^2 = 0.167$), and follow-up comparisons of each stimulus type at each age revealed a reliable developmental change only for the response to IDS; the amplitude of this peak decreased in negativity from 4 to 6 months (4mo: $M = -6.456$, $SD = 3.267$; 6mo: $M = -3.744$, $SD = 4.178$; ANOVA: $F_{(1,32)} = 4.443$, $p = 0.043$, $d = 0.723$). This developmental decrease in amplitude in response to IDS was not observed for either LEMUR or BW-IDS and differentiated the response to IDS

from the other two stimuli in 6-month-olds (IDS vs. BWS: $t(16) = 1.940$, $p = 0.070$, $d = 0.470$; IDS vs. LEMUR: $t(16) = 3.361$, $p = 0.004$, $d = 0.815$).

4 | DISCUSSION

Infants' auditory-evoked neural responses offer novel insights into how infants' cognitive and attentional states may be modulated in response to acoustic stimuli that support early cognition. The current study reveals that human speech and non-human primate vocalizations modulate 6-month-olds' 4–9 Hz neural activity, suggesting these sounds engage early components of infants' attention, and do so in distinct ways. These results provide the first evidence of neural mechanisms that support infants' acquisition of an increasingly precise and uniquely human language-cognition link.

These results add to a growing body of evidence documenting the rapid organization of cortical networks in the infant brain for processing speech and language [for review, see (Perani et al., 2011)]. Recent neuroimaging studies have described distinct patterns of activity in the infant brain that are differentially activated by speech, compared to non-linguistic stimuli such as backwards speech, sine-wave contours, tones, scrambled speech, whistled surrogate language, or non-human primate calls (Dehaene-Lambertz et al., 2002; May et al., 2017; Minagawa-Kawai et al., 2011; Peña et al., 2003; Perani et al., 2011; Shultz et al., 2014). Although specialized networks for processing speech develop within the first few months, the infant brain remains sensitive to the vocalizations of other species. For example, in 4-month-olds, non-human primate vocalizations elicit greater neural activation than speech does in brain regions [i.e., bilateral anterior superior temporal gyrus (Minagawa-Kawai et al., 2011)] that will later be recruited to specialize in speech processing (Fecteau et al., 2004). This finding, which indicates that there may be a protracted period in human infancy during which the infant brain remains sensitive to non-human primate vocalizations—despite infants' preference for speech, which emerges by 3 months (Shultz & Vouloumanos, 2010; Vouloumanos et al., 2010; Vouloumanos & Werker, 2007)—converges well with behavioral evidence that non-human primate vocalizations support object categorization until 6 months (Ferry et al., 2013), and that non-human primate vocalizations can boost cognitive abilities even in 6-month-olds if infants were previously and systematically exposed to those sounds (Perszyk & Waxman, 2016).

In response to infant-directed speech, 6-month-olds' neural responses showed a decrease in 4–9 Hz power. This event-related desynchronization of infants' neural activity elicited by speech may reflect a transition to a neurocognitive state that is primed for learning, or at least a state that optimizes establishing links between external stimuli and internal representations. This interpretation is consistent with several previous demonstrations that when attention is manipulated directly, infants exhibit reduced neural activity or desynchronization within this 4–9 Hz frequency range. For example, reduced neural activity or desynchronization within this frequency range was observed for a range of attention-related tasks, including sustained

visual attention [7- to 12-month-olds, 5.6–9.6 Hz (Stroganova et al., 1999)], sustained audio-visual attention [6- to 12-month-olds, 6–9 Hz (Xie et al., 2018)], and periods of joint attention in infant-adult dyads [9-month-olds, 5–7 Hz: (Hoehl et al., 2014); 12-month-olds, 4–6 Hz (St. John et al., 2016)]. Because speech is a particularly salient stimulus for young infants, we suspect that it may similarly engage exogenous attentional processes that support downstream cognitive processes such as object categorization [for reviews of this idea, see (de Diego-Balaguer et al., 2016; Perszyk & Waxman, 2018)]. Together, these findings suggest that the reduction in infants' 4–9 Hz activity observed here might reflect a speech-driven cognitive state that is advantageous for learning.

In contrast, non-human primate vocalizations elicited an increase in 6-month-olds' 4–9 Hz power. This event-related synchronization of infants' neural activity may reflect disengagement from this once-privileged stimulus that supported categorization at 4 months, but whose role has since been neutralized or modified. An increase in neural activity within this range is thought to be a neural signature of inhibition and disengagement of cortical areas that process task-irrelevant information; increased activity in this frequency band has been identified in cortical regions processing irrelevant or distracting information during attention-related tasks in adults, and is thought to be a signature of suppression for stimuli or stimulus features that are to be ignored [for review, see (Foxy & Snyder, 2011; Klimesch, 1999, 2012)].

Further evidence that this neural response may reflect disengagement from a stimulus that once modulated neural activity comes from the observation that backwards speech—a stimulus that never supports infant categorization—did not modulate neural activity in either 4- or 6-month-olds. This discrepancy in the neural responses to acoustic stimuli that elicit identical behavioral responses at 6-months suggests that infants may preserve a trace of the connection between non-human primate vocalizations and cognition—perhaps in the event that this stimulus later proves meaningful. This might occur by means of exposure; non-human primate vocalizations continue to boost categorization in 6-month-olds who were systematically exposed to these vocalizations (Perszyk & Waxman, 2016). If so, the reduction in 4–9 Hz activity observed in response to speech may reflect how familiarity guides infants to specify which stimuli they link to cognition. Future work is necessary to test whether this neural signature may help to identify stimuli for which the link to cognition remains accessible via exposure, and if exposure to such stimuli decreases neural activity.

The reduction in amplitude of the negative posterior ERP peak in response to speech observed in 6-month-olds may reflect the influence of exposure on the neural processing of acoustic stimuli. This peak was selectively reduced in response to speech, both in relation to the response evoked by the same speech stimuli in 4-month-olds, and in relation to the responses evoked by the other two stimuli (backwards speech and non-human primate vocalizations). This may reflect the predictability of speech sounds, since reduced amplitude of ERP components are typically associated with inputs that elicit less surprise. Between 4- and 6-months, the infant brain gains familiarity with speech sounds and likely processes them more efficiently, recruiting fewer neural resources to process these more frequent stimuli.



Previous work has described changes in the latency of an early large component, similar in morphology, occurring around 240 ms (Nacar Garcia et al., 2018; Peña et al., 2010). Nacar Garcia et al. observed only a latency difference in this peak for native- compared to non-native languages for monolingual 4-month-olds, but for the latter it remains unknown whether there were stimulus-related differences in the latency or amplitude of this peak. Although the peak we observe here occurs later than in this prior work, the timing is consistent with other speech-evoked potential work in infants this age (for review, see Trainor, 2007). It is possible that the timing of this peak could be due to the acoustic variability in our stimulus set. In contrast to previous investigations that included only natural speech, here we interspersed natural speech with stimuli that cannot be produced by a vocal tract (backwards speech) and stimuli produced by non-humans (non-human primate vocalizations). Moreover, the reduction of the amplitude of this peak in response to speech is consistent with evidence that this component may be modulated by stimulus familiarity; in infants, the amplitude of the visual-evoked central negative component (Nc; 400–800 ms) is reduced when a visual stimulus is more familiar (Wahl et al., 2019). Together, these findings suggest that similar neural generators may be involved in the discriminative processes that lead infants to selectively favor speech over other acoustic stimuli, and to link speech to cognition.

These neurodevelopmental changes in infants' neural responses to each of these three acoustic stimuli align well with the established behavioral effect of each stimulus type on infant categorization. Nonetheless, several compelling questions remain. First, the mechanism by which human speech and non-human primate vocalizations support object categorization in 4-month-olds remains elusive. In 4-month-olds, there were no differences in the neural responses to acoustic stimuli that boost 4-month-olds' categorization (infant-directed speech and non-human primate vocalizations) compared to an acoustic stimulus that does not (backwards speech). There are several possible interpretations for this observation. For one, it is possible that the neural measures analyzed here are not sensitive enough to detect the differences in 4-month-olds' neural responses to these stimuli that are implicated in object categorization. Alternatively, it is also possible that at 4 months, hearing these stimuli does not yet draw upon the neural resources we have measured here to facilitate categorization performance, and some other, perhaps more basic, mechanism is responsible. For example, differences in the auditory-evoked response in 4-month-olds may emerge at more peripheral levels of auditory processing, which might be detectable, for example, in the brainstem's frequency-following response to these acoustic stimuli. Another possibility is that these neural responses are the product of distinct functions at 4, as compared to 6 months of age. For example, longitudinal assessment of silent ("resting") state EEG relative power in this frequency range (6–9 Hz) reveals that activity in this band undergoes distinct changes between 5 and 10 months (Marshall et al., 2002). This suggests that intrinsic neural activity under 10 Hz is also developing across these early months of life.

Second, although we demonstrate here that initially-privileged acoustic stimuli modulate 4–9 Hz activity in 6-month-olds, and sim-

ilar influences on neural activity have been described in relation to attentional engagement in infancy (Orehova et al., 2001; Stroganova et al., 1999; Xie et al., 2018), a direct link between the neural activity observed here and attentional processes involved in the object categorization task remains elusive. It also remains unknown which aspects of the broader domain of attention are implicated in infants' neural and behavioral tuning in response to acoustic stimuli and their effects on cognitive tasks. In very young infants, attentional processes are mostly exogenous: novel or salient stimuli elicit orienting responses, which shift infants' attention towards salient features of their environment and inhibit responses to distractors [for reviews, see (de Diego-Balaguer et al., 2016; Gomes et al., 2000; Oakes & Amso, 2018)]. It is not clear whether it would be possible to identify and directly manipulate these aspects of attention in the context of the categorization task. Furthermore, although the neural activity investigated here was in response to the acoustic stimuli alone (without the additional dimension of the visual stimuli as in the object categorization task), if these changes in neural activity can be taken to reflect a change in infants' cognitive state elicited by these acoustic stimuli, it is likely that these acoustic stimuli would similarly influence behavior in a multimodal task like object categorization. Future work is necessary to directly assess how auditory attention influences visual attention during infancy, and specifically how these processes interact to support cognitive processes including those that support categorization.

Finally, another open question is which acoustic features, singly or in combination, of these stimuli engage infants' attention in such a way as to boost categorization. In the current study, the non-human primate tokens were selected to approximate the mean duration of the human speech tokens (see Table S1). Acoustic differences do, however, remain. For example, the frequency composition differs between stimuli, as the speech tokens have more concentrated energy in the 3–4 Hz band than do the more broadband non-human primate tokens (see Figures 1 and S1 for further depiction of acoustic features). Additionally, the speech and non-human primate vocalizations also differ in their pitch range and contour, which may be an especially salient feature for young infants. Future work employing more fine-grained analyses to compare the spectro-temporal modulation spectra and pitch contours of these stimuli could shed light on these acoustic differences. That being said, previous work has failed to find any correlational relationships between stimulus features such as pitch, perceived loudness, acoustic complexity, or amplitude variation and infants' categorization behavior (Woodruff Carr et al., 2021) or infants' preferences to listen to speech over other sounds (Shultz & Vouloumanos, 2010).

Moreover, differences in infants' responses to these stimuli cannot be attributed *entirely* to low-level acoustic features such as acoustic complexity, spectral composition, or the spectro-temporal modulation structure. The backwards speech tokens used here and in the previous behavioral studies (Ferry et al., 2010, 2013) are identical in acoustic complexity and spectral composition to the speech tokens—but elicit strikingly different behavioral and neural responses. Of course, reversal of a speech recording modifies the rhythmic and prosodic cues that neonates use to discriminate speech (Goswami, 2019;



Mehler et al., 1988), and results in an amplitude envelope profile that is not compatible with a sound naturally produced by a vocal tract. Because the amplitude envelope of sounds is encoded early along the auditory pathway (Galbraith et al., 2004; E. C. Smith & Lewicki, 2006), the time-reversed envelope may be a cue that contributes to a distinction between the acoustic stimuli that will and will not carry communicative information at a very early age. Nevertheless, it remains an open question whether other features of acoustic stimuli, such as the origin of a sound (i.e., whether it is naturally produced or artificially engineered; whether it is produced by a human or non-human vocal tract) or its linguistic content may explain emerging developmental differences between backwards speech, speech, and non-human primate vocalizations for both infant categorization and the modulation of neural activity. Future work investigating other non-human vocalizations and non-speech human vocalizations could provide indications as to the relevance of these features.

These findings shed new light on earlier behavioral results and provide novel insights into how very young infants begin to link the sounds they hear to the objects they encounter in their environment. The data provide a more nuanced understanding of how an initially privileged acoustic stimulus (non-human primate vocalizations) and a stimulus that is never linked to cognition (backwards speech) may have different effects on infant cognition, despite equivocal outcomes that are limited by the challenges inherent in assessing infant behavior. Furthermore, the results suggest that the quintessentially human link between language and cognition may be guided by maturationally-sensitive systems of arousal and attention within the first 6 months of life.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

ACKNOWLEDGMENTS

We thank members of the Northwestern University Infant and Child Development Center for insightful discussions that supported this research. We thank Nonye Ogbuefi, Mary Okematti, Murielle Standley, and Ren Ashley Mondesir for their assistance with data collection. We also thank Ahren B. Fitzroy for analytic and theoretical advice, Xiangbin Teng and Ann R. Bradlow for analytic scripts and advice for analyzing the acoustic stimuli, and Chris Mercer for providing the recordings of the non-human primate vocalizations. This work was supported by the National Institutes of Health (R01 HD08831-03SI) and an Alumnae of Northwestern University Grant.

DATA AVAILABILITY STATEMENT

All stimuli and data not under copyright restrictions are posted at <https://osf.io/pmt46/>.

ORCID

Kali Woodruff Carr  <https://orcid.org/0000-0002-3142-0476>

Sandra R. Waxman  <https://orcid.org/0000-0002-5009-2068>

REFERENCES

- Althaus, N., & Mareschal, D. (2012). Using saliency maps to separate competing processes in infant visual cognition. *Child Development, 83*(4), 1122–1128. <https://doi.org/10.1111/j.1467-8624.2012.01766.x>
- Aslin, R. N. (2007). What's in a look? *Developmental Science, 10*(1), 48–53. <https://doi.org/10.1111/j.1467-7687.2007.00563.x>
- Bell, M. A. (2002). Power changes in infant EEG frequency bands during a spatial working memory task. *Psychophysiology, 39*(4), 450–458. <https://doi.org/10.1017/S0048577201393174>
- Bergelson, E., & Swingle, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences, 109*(9), 3253–3258. <https://doi.org/10.1073/pnas.1113380109>
- Berger, H. (1929). Über das Elektrenkephalogramm des Menschen. *Archiv Für Psychiatrie Und Nervenkrankheiten, 87*(1), 527–570. <https://doi.org/10.1007/BF01797193>
- Boersma, P., & Weenink, D. (2001). Praat, a system for doing phonetics by computer. *Glott International, Vol., 5*, pp. 341–345. Retrieved from <http://www.praat.org/>
- Bosseler, A. N., Taulu, S., Pihko, E., Mäkelä, J. P., Imada, T., Ahonen, A., & Kuhl, P. K. (2013). Theta brain rhythms index perceptual narrowing in infant speech perception. *Frontiers in Psychology, 4*, 690. <https://doi.org/10.3389/fpsyg.2013.00690>
- Carey, S. (2011). Précis of the origin of concepts. *Behavioral and Brain Sciences, 34*(3), 113–124. <https://doi.org/10.1017/S0140525X10000919>
- Cohen, M. X. (2014). Analyzing Neural Time Series Data. In *Analyzing Neural Time Series Data*, <https://doi.org/10.7551/mitpress/9609.001.0001>
- Colombo, J. (2002). Infant attention grows up: The emergence of a developmental cognitive neuroscience perspective. *Current Directions in Psychological Science, 11*(6), 196–200. <https://doi.org/10.1111/1467-8721.00199>
- Davidson, R., & Fox, N. (1982). Asymmetrical brain activity discriminates between positive and negative affective stimuli in human infants. *Science, 218*(4578), 1235–1237. <https://doi.org/10.1126/science.7146906>
- de Diego-Balaguer, R., Martinez-Alvarez, A., & Pons, F. (2016). Temporal attention as a Scaffold for language development. *Frontiers in Psychology, 7*, 44. <https://doi.org/10.3389/fpsyg.2016.00044>
- DeBoer, T., Scott, L. S., & Nelson, C. A. (2007). Methods for acquiring and analyzing infant event-related potentials. In M. De Haan (Ed.), *Infant EEG and Event-Related Potentials* (pp. 5–38). Psychology Press.
- Dehaene-Lambertz, G., & Dehaene, S. (1994). Speed and cerebral correlates of syllable discrimination in infants. *Nature, 370*(6487), 292–295. <https://doi.org/10.1038/370292a0>
- Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science, 298*(5600), 2013–2015. <https://doi.org/10.1126/science.1077066>
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. In *Journal of Neuroscience Methods*, (Vol., 134). Retrieved from <http://www.sccn.ucsd.edu/eeglab/>
- Ellis, D. P. W. (1998). Gammatone-like spectrograms. Retrieved from <http://cobweb.ecn.purdue.edu/~malcolm/interval/1998-010/>
- Fecteau, S., Armony, J. L., Joannette, Y., & Belin, P. (2004). Is voice processing species-specific in human auditory cortex? An fMRI study. *NeuroImage, 23*(3), 840–848. <https://doi.org/10.1016/j.neuroimage.2004.09.019>
- Ferry, A. L., Hespos, S. J., & Waxman, S. R. (2010). Categorization in 3- and 4-month-old infants: An advantage of words over tones. *Child Development, 81*(2), 472–479. <https://doi.org/10.1111/j.1467-8624.2009.01408.x>
- Ferry, A. L., Hespos, S. J., & Waxman, S. R. (2013). Nonhuman primate vocalizations support categorization in very young human infants. *Proceedings of the National Academy of Sciences, 110*(38), 15231–15235. <https://doi.org/10.1073/pnas.1221166110>
- Foxe, J. J., & Snyder, A. C. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Frontiers in Psychology, 2*, 154. <https://doi.org/10.3389/fpsyg.2011.00154>

- Friedrich, M., & Friederici, A. D. (2011). Word learning in 6-month-olds: Fast encoding-weak retention. *Journal of Cognitive Neuroscience*, 23(11), 3228–3240. https://doi.org/10.1162/jocn_a.00002
- Friedrich, M., & Friederici, A. D. (2017). The origins of word learning: Brain responses of 3-month-olds indicate their rapid association of objects and words. *Developmental Science*, 20(2), e12357. <https://doi.org/10.1111/desc.12357>
- Friedrich, M., Wilhelm, I., Born, J., & Friederici, A. D. (2015). Generalization of word meanings during infant sleep. *Nature Communications*, 6(1), 6004. <https://doi.org/10.1038/ncomms7004>
- Fujioka, T., & Ross, B. (2008). Auditory processing indexed by stimulus-induced alpha desynchronization in children. *International Journal of Psychophysiology*, 68(2), 130–140. <https://doi.org/10.1016/j.ijpsycho.2007.12.004>
- Fulkerson, A. L., & Waxman, S. R. (2007). Words (but not tones) facilitate object categorization: Evidence from 6- and 12-month-olds. *Cognition*, 105(1), 218–228. <https://doi.org/10.1016/j.cognition.2006.09.005>
- Galbraith, G. C., Amaya, E. M., Diaz De Rivera, J. M., Donan, N. M. M., Duong, M. T., Hsu, J. N., Tran, K., Tsang, L. P. (2004). Brain stem evoked response to forward and reversed speech in humans. *NeuroReport*, 15(13), 2057–2060. <https://doi.org/10.1097/00001756-200409150-00012>
- Gomes, H., Molholm, S., Christodoulou, C., Ritter, W., & Cowan, N. (2000). The development of auditory attention in children. *Frontiers in Bioscience*, 5(3), 108–120. <https://doi.org/10.2741/Gomes>
- Goswami, U. (2019). Speech rhythm and language acquisition: an amplitude modulation phase hierarchy perspective. *Annals of the New York Academy of Sciences*, 1453(1), 67–78. <https://doi.org/10.1111/nyas.14137>
- Hirsh-Pasek, K., Cauley, K. M., Golinkoff, R. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, 14(1), 23–45. <https://doi.org/10.1017/S030500090001271X>
- Hoehl, S., Michel, C., Reid, V. M., Parise, E., & Striano, T. (2014). Eye contact during live social interaction modulates infants' oscillatory brain activity. *Social Neuroscience*, 9(3), 300–308. <https://doi.org/10.1080/17470919.2014.884982>
- Jusczyk, P. W. (1999). How infants begin to extract words from speech. *Trends in Cognitive Sciences*, 3(9), 323–328. [https://doi.org/10.1016/S1364-6613\(99\)01363-7](https://doi.org/10.1016/S1364-6613(99)01363-7)
- Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Research Reviews*, 29(2–3), 169–195. [https://doi.org/10.1016/S0165-0173\(98\)00056-3](https://doi.org/10.1016/S0165-0173(98)00056-3)
- Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, 16(12), 606–617. <https://doi.org/10.1016/j.tics.2012.10.007>
- Krause, C. M., Pörn, B., Lang, A. H., & Laine, M. (1997). Relative alpha desynchronization and synchronization during speech perception. *Cognitive Brain Research*, 5(4), 295–299. [https://doi.org/10.1016/S0926-6410\(97\)00009-8](https://doi.org/10.1016/S0926-6410(97)00009-8)
- Kushnerenko, E., Eponiemi, R., Balan, P., Fellman, V., Huotilainen, M., & Näätänen, R. (2002). Maturation of the auditory event-related potentials during the first year of life. *NeuroReport*, 13(1), 47–51. <https://doi.org/10.1097/00001756-200201210-00014>
- Kwon, M.-K., Setoodehnia, M., Baek, J., Luck, S. J., & Oakes, L. M. (2016). The development of visual search in infancy: Attention to faces versus salience. *Developmental Psychology*, 52(4), 537–555. <https://doi.org/10.1037/dev0000080.supp>
- Libertus, M. E., Pruitt, L. B., Woldorff, M. G., & Brannon, E. M. (2009). Induced alpha-band oscillations reflect ratio-dependent number discrimination in the infant brain. *Journal of Cognitive Neuroscience*, 21(12), 2398–2406. <https://doi.org/10.1162/jocn.2008.21162>
- Lindsley, D. B. (1939). A longitudinal study of the occipital alpha rhythm in normal children: Frequency and amplitude standards. *Pedagogical Seminary and Journal of Genetic Psychology*, 55(1), 197–213. <https://doi.org/10.1080/08856559.1939.10533190>
- Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, 8, 213. <https://doi.org/10.3389/fnhum.2014.00213>
- Mai, X., Xu, L., Li, M., Shao, J., Zhao, Z., Lamm, C., Fox, N. A., Lozoff, B. (2014). Sounds elicit relative left frontal alpha activity in 2-month-old infants. *International Journal of Psychophysiology*, 94(3), 287–291. <https://doi.org/10.1016/j.ijpsycho.2014.09.008>
- Marshall, P. J., Bar-Haim, Y., & Fox, N. A. (2002). Development of the EEG from 5 months to 4 years of age. *Clinical and Experimental Rheumatology*, 113, 1199–1208. Retrieved from www.elsevier.com/locate/clinph
- May, L., Gervain, J., Carreiras, M., & Werker, J. F. (2017). The specificity of the neural response to speech at birth. *Developmental Science*, 21(3), 1–9. <https://doi.org/10.1111/desc.12564>
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29(2), 143–178. [https://doi.org/10.1016/0010-0277\(88\)90035-2](https://doi.org/10.1016/0010-0277(88)90035-2)
- Minagawa-Kawai, Y., van der Lely, H., Ramus, F., Sato, Y., Mazuka, R., & Dupoux, E. (2011). Optical brain imaging reveals general auditory and language-specific processing in early infant development. *Cerebral Cortex*, 21(2), 254–261. <https://doi.org/10.1093/cercor/bhq082>
- Murphy, G. L. (2004). The big book of concepts. In MIT Press. MIT Press.
- 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. *Scientific Reports*, 8(1), 2770. <https://doi.org/10.1038/s41598-018-20824-0>
- Oakes, L., & Amso, D. (2018). Development of Visual Attention. In S. Ghetti (Ed.), *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience IV* (pp. 3–35). John Wiley & Sons, Inc.
- Orehkova, E. V., Stroganova, T. A., & Posikera, I. N. (2001). Alpha activity as an index of cortical inhibition during sustained internally controlled attention in infants. *Clinical Neurophysiology*, 112(5), 740–749. [https://doi.org/10.1016/S1388-2457\(01\)00502-8](https://doi.org/10.1016/S1388-2457(01)00502-8)
- Peña, M., Maki, A., Kovačić, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., & Mehler, J. (2003). Sounds and silence: An optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences*, 100(20), 11702–11705. <https://doi.org/10.1073/pnas.1934290100>
- Peña, M., Pittaluga, E., & Mehler, J. (2010). Language acquisition in premature and full-term infants. *Proceedings of the National Academy of Sciences*, 107(8), 3823–3828. <https://doi.org/10.1073/pnas.0914326107>
- Perani, D., Saccuman, M. C., Scifo, P., Anwander, A., Spada, D., Baldoli, C., Poloniato, A., Lohmann, G., Friederici, A. D. (2011). Neural language networks at birth. *Proceedings of the National Academy of Sciences*, 108(38), 16056–16061. <https://doi.org/10.1073/pnas.1102991108>
- Perszyk, D. R., & Waxman, S. R. (2016). Listening to the calls of the wild: The role of experience in linking language and cognition in young infants. *Cognition*, 153, 175–181. <https://doi.org/10.1016/j.cognition.2016.05.004>
- Perszyk, D. R., & Waxman, S. R. (2018). Linking language and cognition in infancy. *Annual Review of Psychology*, 69, 231–250. <https://doi.org/10.1146/annurev-psych-122216>
- Perszyk, D. R., & Waxman, S. R. (2019). Infants' advances in speech perception shape their earliest links between language and cognition. *Scientific Reports*, 9(1), 3293. <https://doi.org/10.1038/s41598-019-39511-9>
- Purhonen, M., Kilpeläinen-Lees, R., Valkonen-Korhonen, M., Karhu, J., & Lehtonen, J. (2005). Four-month-old infants process own mother's voice faster than unfamiliar voices-Electrical signs of sensitization in infant brain. *Cognitive Brain Research*, 24(3), 627–633. <https://doi.org/10.1016/j.cogbrainres.2005.03.012>
- Quinn, P. C., Lee, K., & Pascalis, O. (January 4). (2019) Face processing in infancy and beyond: The case of social categories. *Annual Review of Psychology*, 70, 165–189. <https://doi.org/10.1146/annurev-psych-010418-102753>
- R Core Team. (2018). R: A Language and Environment for Statistical Computing.



- Saby, J. N., & Marshall, P. J. (2012). The utility of EEG band power analysis in the study of infancy and early childhood. *Developmental Neuropsychology*, 37(3), 253–273. <https://doi.org/10.1080/87565641.2011.614663>
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926–1928. <https://doi.org/10.1126/science.274.5294.1926>
- Shultz, S., & Vouloumanos, A. (2010). Three-month-olds prefer speech to other naturally occurring signals. *Language Learning and Development*, 6(4), 241–257. <https://doi.org/10.1080/15475440903507830>
- Shultz, S., Vouloumanos, A., Bennett, R. H., & Pelphrey, K. (2014). Neural specialization for speech in the first months of life. *Developmental Science*, 17(5), 766–774. <https://doi.org/10.1111/desc.12151>
- Smith, E. C., & Lewicki, M. S. (2006). Efficient auditory coding. *Nature*, 439(7079), 978–982. <https://doi.org/10.1038/nature04485>
- Smith, J. R. (1938). The electroencephalogram during normal infancy and childhood: iii. preliminary observations on the pattern sequence during sleep. *Pedagogical Seminary and Journal of Genetic Psychology*, 53(2), 471–482. <https://doi.org/10.1080/08856559.1938.10533822>
- Smith, J. R. (1939). The “occipital” and “pre-central” alpha rhythms during the first two years. *Journal of Psychology: Interdisciplinary and Applied*, 7(2), 223–226. <https://doi.org/10.1080/00223980.1939.9917630>
- Smith, J. R. (1941). The frequency growth of the human alpha rhythms during normal infancy and childhood. *Journal of Psychology: Interdisciplinary and Applied*, 11(1), 177–198. <https://doi.org/10.1080/00223980.1941.9917028>
- St. John, A. M., Kao, K., Choksi, M., Liederman, J., Grieve, P. G., & Tarullo, A. R. (2016). Variation in infant EEG power across social and nonsocial contexts. *Journal of Experimental Child Psychology*, 152, 106–122. <https://doi.org/10.1016/j.jecp.2016.04.007>
- Stroganova, T. A., Orekhova, E. V., & Posikera, I. N. (1999). EEG alpha rhythm in infants. *Clinical Neurophysiology*, 110(6), 997–1012. [https://doi.org/10.1016/S1388-2457\(98\)00009-1](https://doi.org/10.1016/S1388-2457(98)00009-1)
- Thut, G., Nietzel, A., Brandt, S. A., & Pascual-Leone, A. (2006). α -Band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *Journal of Neuroscience*, 26(37), 9494–9502. <https://doi.org/10.1523/JNEUROSCI.0875-06.2006>
- Tincoff, R., & Jusczyk, P. W. (2012). Six-month-olds comprehend words that refer to parts of the body. *Infancy*, 17(4), 432–444. <https://doi.org/10.1111/j.1532-7078.2011.00084.x>
- Trainor, L. J. (2007). Event-related potential (ERP) measures in auditory development research. In L. Schmidt & S. Segalowitz (Eds.), *Developmental Psychophysiology: Theory, Systems, and Methods* (pp. 69–102). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511499791.005>
- Tummeltshammer, K. S., Mareschal, D., & Kirkham, N. Z. (2014). Infants' selective attention to reliable visual cues in the presence of salient distractors. *Child Development*, 85(5), 1981–1994. <https://doi.org/10.1111/cdev.12239>
- Van Diepen, R. M., Foxe, J. J., & Mazaheri, A. (2019). The functional role of alpha-band activity in attentional processing: the current zeitgeist and future outlook. *Current Opinion in Psychology*, 29, 229–238. <https://doi.org/10.1016/j.copsyc.2019.03.015>
- Vouloumanos, A., Hauser, M. D., Werker, J. F., & Martin, A. (2010). The tuning of human neonates' preference for speech. *Child Development*, 81(2), 517–527. <https://doi.org/10.1111/j.1467-8624.2009.01412.x>
- Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: evidence for a bias for speech in neonates. *Developmental Science*, 10(2), 159–164. <https://doi.org/10.1111/j.1467-7687.2007.00549.x>
- Wahl, S., Marinović, V., & Träuble, B. (2019). Gaze cues of isolated eyes facilitate the encoding and further processing of objects in 4-month-old infants. *Developmental Cognitive Neuroscience*, <https://doi.org/10.1016/j.dcn.2019.100621>
- Waxman, S. R., & Gelman, S. A. (2009). Early word-learning entails reference, not merely associations. *Trends in Cognitive Sciences*, 13(6), 258–263. <https://doi.org/10.1016/j.tics.2009.03.006>
- Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology*, 29(3), 257–302. <https://doi.org/10.1006/cogp.1995.1016>
- Woodruff Carr, K., Perszyk, D. R., & Waxman, S. R. (2021). Birdsong fails to support object categorization in human infants. *PLOS ONE*, 16(3), e0247430. <https://doi.org/10.1371/journal.pone.0247430>
- Xie, W., Mallin, B. M., & Richards, J. E. (2018). Development of infant sustained attention and its relation to EEG oscillations: An EEG and cortical source analysis study. *Developmental Science*, 21(3), e12562. <https://doi.org/10.1111/desc.12562>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Woodruff Carr, K., Perszyk, D. R., Norton, E. S., Voss, J. L., Poeppel, D., & Waxman, S. R. (2021). Developmental changes in auditory-evoked neural activity underlie infants' links between language and cognition. *Developmental Science*, 1–12. <https://doi.org/10.1111/desc.13121>