Cortical Oscillations in Pre-verbal Infants Track Rhythmic Speech and Non-speech Stimuli

Áine Ní Choisdealbha, Adam Attaheri, Sinead Rocha, Natasha Mead, Helen Olawole-Scott, Perrine Brusini, Samuel Gibbon, Panagiotis Boutris, Declan Hines, Christina Grey, Sheila Flanagan, and Usha Goswami

1.   Introduction

The acoustic features of the sounds we hear are encoded in our neural responses to those sounds. For example, if we listen to a sound repeated at a particular rhythm, our neural oscillations align to that rhythm. Speech is a quasi-rhythmic stimulus, and our neural oscillations track the slow rhythms of its amplitude envelope (Ding et al., 2016). Frequencies in the amplitude envelope which correspond to the rates at which syllables and syllabic stress are produced (around two to four cycles per second) appear important both for neural tracking of the envelope and for the intelligibility of speech to the listener (Doelling et al., 2014). If these rhythms are important for speech processing in the developed brain, could they play a role in how infants process speech and acquire language? To answer this question, it is important to first find out how rhythmic auditory stimulation is processed by the developing brain.

The potential importance of neural oscillations in the delta (~0.5-4Hz) and theta (~4-8Hz) frequency bands is complemented by modeling of the structure of the amplitude envelopes of adult-directed speech (ADS) and infant-directed speech (IDS). The modeling showed that the amplitude envelope of speech is composed of nested modulations at rates corresponding to stress patterns (~2 Hz AM band), syllables (~5 Hz AM band), and phonemes (~20 Hz AM band; Leong & Goswami, 2015). Subsequent modelling of speech directed at infants of seven, nine and eleven months of age revealed more modulation power at the stressed syllable rate (~2 Hz, corresponding temporally to the neurophysiological delta band) than there is in ADS. Conversely, in ADS, there was more modulation power in the theta band than in IDS at seven- and nine-month-olds, but no difference between ADS and IDS at eleven months (Leong et al., 2017). This suggests that the manner in which we speak to infants provides them with additional low-frequency temporal information, particularly acoustic information which conveys syllable stress patterns.

Another source of evidence for the potential role of low frequency speech rhythms and neural oscillations in language development comes from research on dyslexia. Temporal Sampling (TS) Theory (Goswami, 2011) posits that phonological difficulties such as those seen in dyslexia arise in part because low-

frequency neural oscillations do not appropriately align to matching amplitude modulation information in the speech envelope. Relative to both chronological and reading age-matched controls, the neural responses of children with dyslexia show less accurate representation of 2Hz amplitude modulations in the speech envelope (Power et al., 2016). This suggests that the delta band may be particularly important in language processing, potentially because the corresponding rate of syllabic stress in speech creates an over-arching framework in which other components of language like syllables and phonemes are nested (Leong & Goswami, 2015). Tracking of even very simple acoustic rhythm at 2Hz is also affected in children with dyslexia. The phase of their 2Hz neural oscillations, relative to the stimulus rate, differs from children without dyslexia (Power et al., 2013). Furthermore, the difference in phase between stimulus and response is correlated with performance on standardized tests of phonological awareness and reading.

To investigate whether this link between neural responses to slow rhythmic stimuli and language development is present early in life, the longitudinal Cambridge UK "BabyRhythm" project recorded rhythmic neural and motor activity in infants in response to auditory stimulation. The project followed these infants into toddlerhood, taking behavioural measures of language encompassing vocabulary, phonology, grammar, and timing. In the paradigm described here, infants were presented with audiovisual stimuli encompassing speech and non-speech sounds played at 2Hz. In this paper, we aim to establish that there is a delta band oscillatory response to these stimuli in the infant brain. There is ongoing debate in the literature about whether the phenomenon of cortical tracking at the same frequency as an incoming stimulus is due to entrainment of endogenous neural rhythms to the stimulus, or whether it might arise from stimulus-evoked potentials (e.g. Zoefel, ten Oever, & Sack, 2018). The resolution of this debate is beyond the scope of this paper. Regardless of the precise neural explanation of the phenomenon, our aim is to investigate whether it occurs in the infant brain.

A growing number of studies demonstrate that the infant brain does track auditory input, including the speech amplitude envelope. Jessen and colleagues (2019) showed seven-month-old infants a cartoon movie, and found that their EEG tracked many auditory and visual features of the stimulus, including its soundtrack. Further research has shown that this stimulus tracking extends to the speech amplitude envelope. Seven-month-olds' low-frequency (< 8Hz) neural tracking of IDS is more accurate than tracking of ADS (Kalashnikova et al., 2018) and newborn infants track both native and non-native speech (Ortiz-Barajas, Guevara & Gervain, 2021). Attaheri and colleagues (2022) found that neural tracking of IDS was most accurate in the delta (0.5 to 4Hz) band, compared to the theta and alpha bands, supporting further investigation of the particular importance of 2Hz tracking during development.

In this study, we look at local peaks in EEG power and at relative increases in inter-trial coherence (ITC) while infants listened to rhythmic stimulus trains delivered audio-visually at a repetition rate of 2 Hz. We aim to find whether there is an increase in oscillatory power at 2Hz and to determine if these oscillations

are consistently in the same phase relative to the stimulus. This would be indicative of cortical tracking. We also examine whether these responses change between six and nine months of age, to see if there is improvement in tracking ability during this early stage in the development of receptive language. If neural activity in the infant delta band tracks 2Hz rhythms, this will support future investigation of individual differences in tracking, such as preferred phase alignment between oscillation and stimulus. In turn, these individual measures can be related to behavioural measures of language ability to uncover the role of delta band cortical tracking in the early acquisition of language.

2. Methods
2.1. Participants

Participants were 115 infants enrolled in the Cambridge UK "BabyRhythm" project. At six months, 113 infants took part in the paradigm reported here. Their mean age at the time of EEG recording was 6 months 2 days ± 7 days. One-hundred-and-eight infants are included in the analyses. The five excluded infants did not provide at least five clean EEG trials in any condition. At nine months, 108 infants took part, of whom 106 had taken part at six months of age. Mean age at the nine-month EEG recording was 8 months 27 days ± 5 days. Two infants are not included in the analyses. One rejected the EEG cap, and one recording was missing information about stimulus presentation.

2.2. Stimuli and apparatus

Infants were shown two types of audiovisual stimulus. One was a woman repeating the syllable "Ta" twice per second. Her head and shoulders were shown in the video against a blue background. The other was a video of a small ball bouncing on a drum creating a drumbeat sound twice per second. The videos were twelve seconds in duration and contained 24 beats. The gap between the final beat of one video and the first beat of the next varied in length between 1.5 and 2 seconds. Originally, shorter videos – eight seconds long with 16 beats – were proposed for use. The twelve second videos were chosen, to give infants a longer period over which to entrain. Due to experimenter error, some infants (38 at six months, 23 at nine months) nevertheless viewed the eight-second video of the bouncing ball.

These videos were displayed on a computer monitor. The sounds were played through Q acoustics 2020i speakers driven by a Cambridge Audio Topaz AM5 stereo amplifier. Stimuli were presented using Psychtoolbox (Kleiner, Brainard & Pelli, 2007) in Matlab 2017a (The MathWorks, Inc., Natick, MA). EEG data were recorded at 1000Hz through EGI 64-channel Geodesic Sensor Nets with the four facial channels removed, via an EGI GSN amplifier (Electrical Geodesics Inc., Eugene, OR, USA). Data were collected via EGI Netstation and event-related triggers were inserted using a StimTracker (Cedrus Corporation, San Pedro, CA) reacting to the auditory stimulus.

2.3. Procedure

During recording, infants were seated either in an infant chair or, in some cases, on their parent's lap. After setting up the EEG equipment, infants were shown repeated videos of one of the two types of audiovisual stimulus until the experimenters judged that they were no longer interested in it. An attention-grabber video was then played, followed by repeated videos of the other stimulus until the infant's interest waned, at which point the first video was played again. This process was repeated until fifty iterations of each type of video were played, or until the infant became fussy, upset, or otherwise unwilling to continue watching videos.

At this point, five minutes of EEG were recorded in silence without a video playing, generating a "silent" condition. For some infants (34 at six months, 25 at nine months) the silent condition lasted three minutes. It was later extended to obtain more EEG data. To maintain infant contentment during the recording, an experimenter blew bubbles or slowly flipped through a picture book.

2.4. EEG data processing and analysis

EEG were preprocessed in Matlab 2018b using the EEGlab 14 toolbox (Delorme & Makeig, 2004). First, data were filtered with a highpass of 0.2Hz and a lowpass of 45Hz. For each recording, persistently bad channels were identified by eye. This amounted to 4 ± 2.6 channels per infant. These were removed alongside the two channels in front of the infant's ears, which did not fit flush to many infants' faces, and the missing data were interpolated using the eeg_interp function.

Each recording was then visually inspected and periods of extreme noise were removed. Further noise reduction was performed using Artifact Subspace Reconstruction (Kothe & Makeig, 2013). Data were then epoched into segments of 10.5 seconds (6.5 seconds for the shorter drum videos) starting from the third "beat" in each video, with a baseline correction applied based on the preceding 500ms. The silent condition was similarly split into segments of 6.5 or 10.5 seconds. As the gap between audiovisual trials varied, similar pseudorandom gaps were introduced by varying the onset of each silent condition segment by between 10.5 (or 6.5) and 10.95 (or 6.8) seconds relative to the onset of the preceding segment, in steps of 10ms.

Individual segments were visually inspected and any segments with remaining artefacts were removed. On average, infants sat for 83 audiovisual trials at six months with a standard deviation of 20 trials. Forty-four (SD = 24) remained after cleaning. At nine months they sat for an average of 91 trials (SD = 13) and 54 (SD = 20) remained after cleaning. Data were re-referenced to the head average, excluding peripheral channels around the forehead, ears and neck.

A Fast Fourier Transform (FFT) was run on the data using the MTMFFT method in the Fieldtrip Toolbox (Oostenveld et al., 2011) with a Hanning taper.

This resulted in power values for 146 frequency bins between 0.5 and 15Hz (120 for the shorter 6.5-second trials). After the FFT was run, a normalisation procedure was applied to the data. This procedure involves subtracting the mean power of the two frequency bins either side of a particular frequency bin from that frequency bin (Nozaradan et al., 2011). In FFT EEG data, there is typically higher overall power in the lower frequency bands; this normalisation approach reveals where there are peaks exceeding the mean power of local frequencies.

For the inter-trial coherence analysis, a time-frequency analysis was run in Fieldtrip using five Morlet wavelets. This number was selected as a compromise between temporal and spectral precision. Values were obtained for frequency bins between 1 and 15Hz in steps of 0.1Hz and for time windows in steps of 100ms between stimulus onset and 10s. After the time-frequency analysis was run on the data epoched to the onset of the stimulus, the data were epoched again. This time, the start of each epoch was jittered to begin between 10 milliseconds and 490 milliseconds after the onset of the stimulus. This means that the start of each epoch would fall somewhere between the onsets of two syllables or two drumbeats. The time frequency analysis was run again on this data. This was done because high ITC values can be driven by a general increase in oscillatory power as well as by alignment in oscillatory phase across trials. If there is a general increase in power at 2Hz, we would also expect to see an increase in this jittered ITC at 2Hz. We subtracted the jittered ITC from the standard ITC to control for overall increases and power. The remaining values could then be used to determine if inter-trial coherence was time-locked to the stimulus.

2.5. Statistical analysis

Data were analysed via linear mixed effect regressions in R using the lme4 toolbox (Bates & Sarkar, 2007). For the FFT analysis, the dependent variable was relative power. For the ITC analysis, it was the ITC (r) value minus the jittered ITC value. In each case, factors of age (6 and 9 months), frequency bin (2, 4 and 7Hz) and condition (silent, drum, syllable) were included. The frequency bins were selected for the following reasons. The 2Hz bin corresponds to the stimulus frequency, where effects would be expected. The 4Hz bin is a theta band harmonic of 2Hz. Theta oscillations are involved in speech processing (Doelling et al., 2014) and may play a role in language development in conjunction with delta oscillations (Goswami, 2018). Thus we expect effects at this frequency as well, but to a lesser extent than at 2Hz. Finally, we examine 7Hz as a comparator frequency, as it is a relatively low frequency that is not a harmonic of 2Hz. It is also in the infant alpha band, not delta or theta. Each model included a random intercept on participant identity, and a random slope on age. The 24-beat and 16-beat stimuli are analysed separately because FFTs performed on shorter epochs of EEG data necessarily result in higher overall power than FFTs performed on longer epochs.

3. Results
3.1. Power

At six months, relative power at 2Hz was 0.36dB (standard deviation = 0.36) for the drumbeat, 0.35dB (SD = 0.31) for the syllable, and 0.03dB (0.24) for the silent condition. At nine months, these numbers were 0.33dB (0.28), 0.3dB (0.42) and 0dB (0.24) respectively. For the group viewing the shorter drumbeat video, these figures were 0.44dB (SD = 0.47) for the drumbeat and 0.05dB (SD = 0.43) for the silent condition at six months, and 0.35dB (SD = 0.39) and –0.16dB (SD = 0.39) respectively at nine months. Figure 1 depicts the relative power spectra, with peaks seen for the audiovisual stimuli in the 2Hz bin, alongside smaller peaks at the 4Hz bin. Henceforth, we refer to the frequency bins targeted for analysis as "peaks".

Results of the statistical model indicated significant effects of condition, $(F_{(2,1488)} = 66.66, p < 0.0001)$ and peak $(F_{(2,1418)} = 154.54, p < 0.0001)$ as well as an interaction between condition and peak $(F_{(2,1418)} = 33.39, p < 0.0001)$. There was a marginal effect of age $(F_{(1,1522)} = 3.44, p = 0.064)$, but no significant interactions with age. Investigation of fixed effects revealed that the 2Hz peak was higher in the drum $(\beta = 0.035, t = 6.663, p < 0.0001)$ and syllable $(\beta = 0.033, t = 7.325, p < 0.0001)$ conditions than in the silent condition. The 4Hz peak was also higher in these conditions (Drum: $\beta = 0.011, t = 2.401, p = 0.017$; Syllable: $\beta = 0.012, t = 2.467, p = 0.014$). The only age-related difference in the simple effects was a marginal drop in the nine month 4Hz peak to the drum, relative to silence $(\beta = -0.096, t = -1.909, p = 0.056)$.

For the subset of infants who viewed the shorter drumbeat video, a comparison to similarly short segments of silent condition EEG data also showed effects of condition $(F_{(1, 240)} = 13.32, p < 0.0001)$ and peak $(F_{(2, 227)} = 3.85, p = 0.023)$ as well as an interaction between condition and peak $(F_{(2, 227)} = 5.3, p = 0.006)$. There was a marginal effect of age $(F_{(1,165)} = 3.88, p = 0.051)$ and a marginal interaction between age and peak $(F_{(2,227)} = 2.92, p = 0.056)$. Fixed effects indicated that overall, the 4Hz peak was smaller than the 7Hz peak $(\beta = -0.577, t = -3.314, p = 0.001)$ but it was larger at six than nine months $(\beta = 0.609, t = 2.558, p = 0.011)$. As the silent condition is the base case, and the effects therefore apply when there is no stimulation, these results suggest that there was overall more oscillatory activity at 7Hz (in the alpha band) than at 4Hz (theta band) at six months, but that the 4Hz peak increased in size by nine months of age.
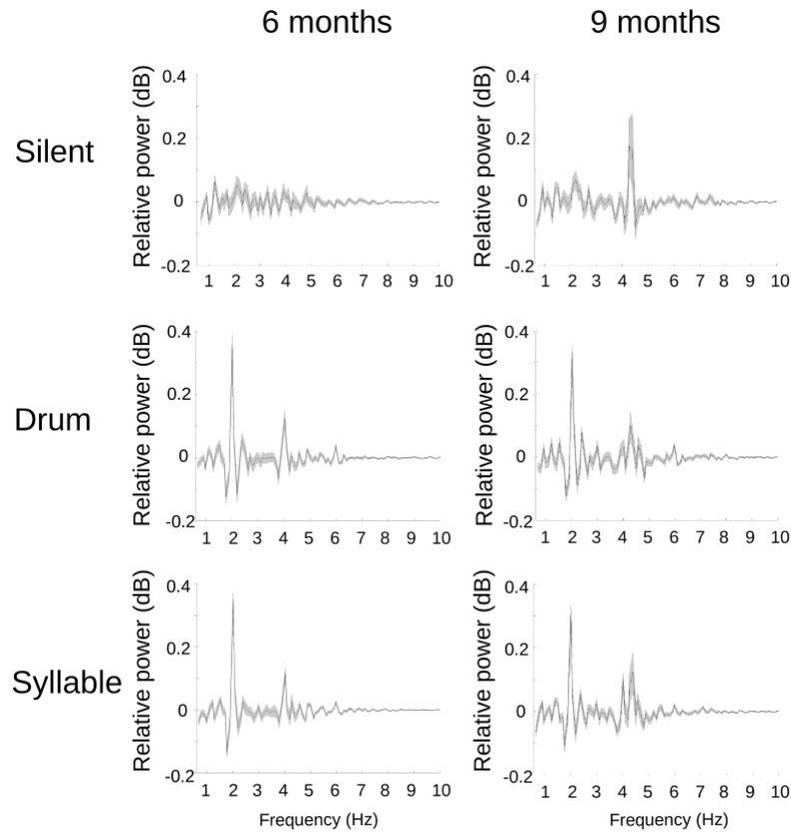
Figure 1: Relative EEG power by condition and age.

This difference is apparent in Figure 1, as a large 4Hz peak appears in the silent condition at nine months. This could point to a development shift in engagement of theta networks. The crucial finding, however, is that both the 2Hz and 4Hz peaks were greater in response to the drum (2Hz: $\beta = 0.612$, t = 2.57, $p = 0.011$; 4Hz: $\beta = 0.61$, t = 2.562, $p = 0.011$) than in the silent condition. This finding mirrors the effects found for the longer, 24-beat stimulus, suggesting that cortical tracking at 2Hz occurred for both audiovisual trial lengths.

3.2. Inter-trial coherence

At six months, the overall ITC r value at 2Hz was 0.24 (SD = 0.06), similar across all conditions. When using the jittered values to account for overall differences in power across conditions, the ITC r values were 0.02 (SD = 0.03) for the drumbeat, 0.03 (SD = 0.02) for the syllable, and 0 (SD = 0.02) for the silent condition. At

nine months, the overall r value at 2Hz was 0.22 (SD = 0.05). When corrected for overall differences in power, the ITC r values were 0.03 (SD = 0.03) for the drumbeat, 0.02 (SD = 0.02) for the syllable, and 0 (SD = 0.02) for the silent condition.
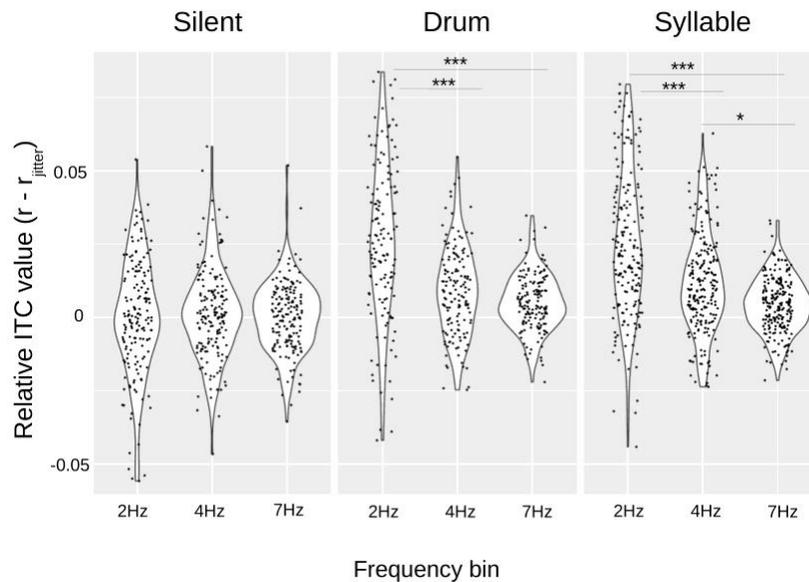

Figure 2: Distribution of relative ITC values by condition and frequency band, for the 24-beat stimuli. Dots depict individual data points.
*** p < 0.001, * p < 0.025

Results of the statistical model for the 24-beat stimuli indicated significant effects of condition, ($F_{(2,1432)}$ =102.19, $p < 0.0001$) and band ($F_{(2,1327)}$ = 92.09, $p < 0.0001$). There were interactions between condition and age ($F_{(2,1427)}$ = 4.87, $p = 0.008$), and condition and band ($F_{(4,1324)}$ = 24.9, $p < 0.0001$). Investigation of fixed effects revealed interactions between 2Hz and the drumbeat ($\beta = 0.023$, t = 5.745, p < 0.0001) and syllable ($\beta = 0.025$, t = 6.984, p < 0.0001) conditions, suggesting enhanced ITC to the AV stimuli at the stimulus rate. There was also an interaction between 4Hz and the syllable condition ($\beta = 0.009$, t = 2.435, p = 0.015). These results are depicted in Figure 2. There were no significant age-related effects or interactions.

For the shorter stimuli, the jittered 2Hz ITC r values at six months were 0 (SD = 0 .02) for the silent condition and 0.02 (SD = 0.03) for the drum. At nine months, these were –0.01 (SD = 0.03) and 0.01 (SD = 0.02) respectively. There were significant effects of condition ($F_{(1,99)}$ = 8.57, $p = 0.004$), band ($F_{(2,904)}$ = 4.27, $p = 0.014$), and age ($F_{(4,1324)}$ = 24.9, $p < 0.0001$). However, there was

no interaction between condition and frequency band , which would point to a relative difference between conditions at 2Hz. Likewise, there were no significant fixed effects. It was therefore not possible to identify phase consistency in the infants' neural responses to the shorter, 16-beat stimuli.

4. Discussion

The data presented here involved simple rhythmic stimuli and a beat-based analytic approach to the measured EEG, allowing us to focus on specific key frequencies of developmental interest regarding TS theory. For example, atypical processing of a rhythmic "ba" stream at 2Hz has been implicated in developmental language difficulties (Power et al., 2013). The data presented here indicate that infant neural oscillatory activity does respond to rhythmic 2Hz stimulation, showing a peak in EEG power at the same frequency. They also indicate that this increase in power is time-locked to the onset of the stimulus, whether speech ("ta") or nonspeech (drumbeat). This suggests that mechanisms for tracking auditory rhythm are present in infancy.

We did not find any differences in 2Hz power or phase consistency between the two types of audiovisual stimulus however. This allows for the possibility that either stimulus type could be used as an early marker of atypical processing. Future research incorporating language acquisition measures taken from the Cambridge UK "BabyRhythm" sample will investigate whether 2Hz processing is linked to language development. In related work, machine-learning algorithms accurately identified whether EEG data from these same infants measured at eight weeks of age was recorded as they heard the drumbeat or as they heard the repeated syllable (Gibbon et al., 2020). This suggests that although similar mechanisms may be involved in how these stimuli are processed in the infant brain, their neural representations differ.

Our findings are also consistent with other research showing neural tracking of audiovisual speech by infants (Attaheri et al., 2022; Jessen et al., 2019; Kalashnikova et al., 2018; Ortiz-Barajas, Guevara & Gervain, 2021). Prior research has however used IDS, and has utilized methods for examining the similarity of the neural response and the amplitude envelope of the stimulus.

At both six and nine months, infants showed an increase in normalised power in response to audiovisual speech and non-speech rhythms, relative to silence. We found that the expected increase in power occurred at 2Hz, which is the stimulus rate, and that it was accompanied by a relative increase in inter-trial coherence in the 2Hz band for the longer (24-beat) sequences. The latter finding suggests that the increase in 2Hz power is not a general increase in power but rather that the neural oscillation in response to the stimulus is in the same phase across repetitions of the same audiovisual trial. This is an important finding, as Power et al. (2013) found that atypical processing of a rhythmic "ba" syllable train was characterized by a difference in preferred phase between children with dyslexia and typically-developing control children. Having found evidence for phase consistency among infants in this sample, future work will examine individual

differences in these infants' preferred phase and how such differences may relate to language development.

We also found an increase in power in response to both stimuli at 4Hz. One reason for this increase is simply that it is a harmonic frequency of 2Hz. It would be unusual for a neural signal to take the form of a perfect sinusoid, and therefore the Fourier spectrum is expected to have power at the harmonics of the key frequency (see e.g. Zhou et al., 2016). Although this harmonic effect is seen for both audiovisual stimuli in the FFT analysis, in the ITC analysis we see a 4Hz effect for the syllable stimulus only. This indicates that although 4Hz power increases for both types of stimulus, there is phase consistency in the 4Hz oscillation in response to the speech stimulus only. This latter effect could be related to some acoustic feature of the syllable stimulus, or it could be driven by how speech is processed neurally. Attaheri and colleagues (2022) found significant neural tracking of infant-directed speech in the theta band as well as the delta band. It could be that a theta oscillation is drawn into phase with the stimulus-driven response in preparation for speech processing. Further work on the coupling of delta and theta oscillations in infant neural responses to speech is required to interrogate this possibility.

We did not find that cortical tracking improves across ages. There are multiple potential explanations. First, it may be that this auditory ability is so fundamental that it is present and developed early in life, and not subject to much change in later infancy. Indeed, Ortiz Barajas, Guevara and Gervain (2021) found better amplitude tracking in newborns than at six months of age. Attaheri and colleagues (2022) also found greater accuracy of neural tracking in younger infants (four-month-olds), with no differences emerging between seven and eleven months of age. Another explanation may be more physiological, because the amplitude of recorded infant neural activity measured at the scalp may decrease over time as the fontanelle closes. In this case, there may a difference in tracking ability between six and nine months that is masked by the physiological change. Another developmental explanation may be that as infants grow older, fundamental oscillatory responses to audiovisual stimuli occur alongside neural processes underlying more complex cognitive and perceptual processes. Consequently, the relative strength of the simple beat tracking response in the EEG recording may decrease.

Overall, the data presented here indicate that the mechanisms for cortical tracking of audiovisual rhythms are present in infancy. The infant brain tracks 2Hz audiovisual rhythms at the rate at which the stimulus is presented, and the data suggest that the 2Hz neural oscillation was drawn into a particular phase relative to the onset of the stimulus. There were some differences in theta band harmonic responses to the drum (non-speech) versus syllable (speech) stimuli, pointing to overall differences in how speech rhythm might be processed. This paves the way for a temporal sampling approach to understanding individual differences in language acquisition. We can examine how pre-verbal infants' individual differences in tracking of stimulation at the crucial 2Hz rate, implicated

in developmental language difficulties (Power et al., 2013), relate to the emergence of receptive and productive language.

References

Attaheri, Adam, Ní Choisdealbha, Áine, Di Liberto, Giovanni, Rocha, Sinead, Brusini, Perrine, Mead, Natasha, Olawole-Scott, Helen, Boutris, Panagiotis, Gibbon, Samuel, Williams, Isabel, Grey, Christina, Flanagan, Sheila, & Goswami, Usha (2022). Delta- and theta-band cortical tracking and phase-amplitude coupling to sung speech by infants. *NeuroImage*, *247*, 118698.

Bates, Douglas, & Sarkar, Deepayan (2007). The lme4 package. *R package version*, 2(1), 74.

Delorme, Arnaud, & Makeig, Scott (2004). EEGLAB: an open-source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods, 13*4, 9-21.

Doelling, Keith B., Arnal, Luc H., Ghitza, Oded, & Poeppel, David (2014). Acoustic landmarks drive delta–theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage, 85*, 761-768.

Ding, Nai, Melloni, Lucia, Zhang, Hang, Tian, Xing, & Poeppel, David (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience, 19*(1), 158-164.

Gibbon, Samuel, Attaheri, Adam, Ní Choisdealbha, Áine, Rocha, Sinead, Brusini, Perrine, Mead, Natasha, Boutris, Panagiotis, Olawole-Scott, Helen, Ahmed, Henna, Flanagan, Sheila, Mandke, Kanad, Keshavarzi, Mahmoud, & Goswami, Usha (2021). Machine learning accurately classifies neural responses to rhythmic speech vs. non-speech from 8-week-old infant EEG. *Brain and Language, 220*, 104968.

Goswami, Usha (2018). A neural basis for phonological awareness? An oscillatory temporal-sampling perspective. *Current Directions in Psychological Science, 27*(1), 56-63.

Goswami, Usha (2012). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences, 15(*1), 3-10.

Jessen, Sarah, Fiedler, Lorenz, Münte, Thomas F., & Obleser, Jonas (2019). Quantifying the individual auditory and visual brain response in 7-month-old infants watching a brief cartoon movie. *NeuroImage*, *202*, 116060.

Kalashnikova, Marina,, Peter, Varghese, Di Liberto, Giovanni M., Lalor, Edmund C., & Burnham, Dennis (2018). Infant-directed speech facilitates seven-month-old infants' cortical tracking of speech. *Scientific Reports*, *8*(1), 13745. doi: 10.1038/s41598-018-32150-6

Kleiner, Mario, Brainard, David, & Pelli, Denis (2007). What's new in Psychtoolbox-3? *Perception, 36*, ECVP Abstract Supplement.

Kothe, Christian Andreas, & Makeig, Scott (2013). BCILAB: a platform for brain–computer interface development. *Journal of Neural Engineering, 10*(5), 056014.

Leong, Victoria, & Goswami, Usha (2015). Acoustic-emergent phonology in the amplitude envelope of child-directed speech. *PloS one, 10*(12), e0144411.

Leong, Victoria, Kalashnikova, Marina, Burnham, Denis, & Goswami, Usha (2017). The temporal modulation structure of infant-directed speech. *Open Mind, 1*(2), 78-90.

Nozaradan, Sylvie, Peretz, Isabel, Missal, Marcus, & Mouraux, André (2011). Tagging the neuronal entrainment to beat and meter. *Journal of Neuroscience, 31*(28), 10234-10240.

Oostenveld, Robert, Fries, Pascal, Maris, Eric, Schoffelen, Jan-Mathijs (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience, 2011,* 156869.

Ortiz Barajas, Maria Clemencia, Guevara, Ramón, & Gervain, Judit (2021). The origins and development of speech envelope tracking during the first months of life. *Developmental Cognitive Neuroscience, 48,* 100915.

Power, Alan J., Colling, Lincoln J., Mead, Natasha, Barnes, Lisa, & Goswami, Usha (2016). Neural encoding of the speech envelope by children with developmental dyslexia. *Brain and Language, 160,* 1-10.

Power, Alan J., Mead, Natasha, Barnes, Lisa, & Goswami, Usha (2013). Neural entrainment to rhythmic speech in children with developmental dyslexia. *Frontiers in Human Neuroscience, 7*, 777.

Zhou, Hong, Melloni, Lucia, Poeppel, David, & Ding, Nai (2016). Interpretations of frequency domain analyses of neural entrainment: periodicity, fundamental frequency, and harmonics. Frontiers in human neuroscience, 10, 274.

Zoefel, Benedikt, ten Oever, Sanne, & Sack, Alexander T. (2018). The involvement of endogenous neural oscillations in the processing of rhythmic input: More than a regular repetition of evoked neural responses. *Frontiers in Neuroscience, 12*, 95.