Infant low-frequency EEG cortical power, cortical tracking and phase-amplitude coupling predicts language a year later

Adam Attaheri[1]*, Áine Ní Choisdealbha[1], Sinead Rocha[1,2], Perrine Brusini[1,3], Giovanni M. Di Liberto[4,5], Natasha Mead[1], Helen Olawole-Scott[1], Panagiotis Boutris[1], Samuel Gibbon[1], Isabel Williams[1], Christina Grey[1], Maria Alfaro e Oliveira[1], Carmel Brough[1], Sheila Flanagan[1] & Usha Goswami[1].

1 Centre for Neuroscience in Education, Department of Psychology, University of Cambridge, Cambridge, United Kingdom

2 Psychology and Sports Science, Anglia Ruskin University, Cambridge, United Kingdom

3 Institute of Population Health, University of Liverpool, Liverpool, United Kingdom

4 School of Computer Science and Statistics, Trinity College Dublin, Dublin, Ireland

5 Laboratoire des Systèmes Perceptifs, UMR 8248, CNRS, Ecole Normale Supérieure, PSL Research University, Paris, France

* Corresponding author, E-mail: aa974@cam.ac.uk

Abstract

Cortical signals have been shown to track acoustic and linguistic properties of continual speech. This phenomenon has been measured across the lifespan, reflecting speech understanding as well as cognitive functions such as attention and prediction. Furthermore, atypical low-frequency cortical tracking of speech is found in children with phonological difficulties (developmental dyslexia). Accordingly, low-frequency cortical signals, especially in the delta and theta ranges, may play a critical role in language acquisition. A recent investigation Attaheri et al., 2022 (1) probed cortical tracking mechanisms in infants aged 4, 7 and 11 months as they listened to sung speech. Results from temporal response functions (TRF), phase-amplitude coupling (PAC) and dynamic theta-delta power (PSD) analyses indicated speech

envelope tracking and stimulus related power (PSD) via the delta & theta neural signals. Furthermore, delta and theta driven PAC was found at all ages with gamma amplitudes displaying a stronger PAC to low frequency phases than beta. The present study tests whether those previous findings replicate in the second half of the same cohort (first half: N=61, (1); second half: N=61). In addition to demonstrating good replication, we investigate whether cortical tracking in the first year of life predicts later language acquisition for the full infant cohort (122 infants recruited, 113, retained). Increased delta cortical tracking and theta-gamma PAC were related to better language outcomes using both infant-led and parent-estimated measures. By contrast, increased ~4Hz PSD power and a greater theta/delta power ratio related to decreased parent-estimated language outcomes. The data are interpreted within a "Temporal Sampling" framework for developmental language trajectories.

1 Introduction

Infant speech perception develops rapidly during the first year of life, yet our understanding of the neural factors underpinning language acquisition is still incomplete. Whilst infant behavioral experiments have highlighted perceptual factors contributing to early phonological and morphological development, for example perceptual 'magnet' effects (2), the neural mechanisms supporting infant language acquisition are poorly understood. Here we contribute novel longitudinal neural data from a relatively large sample of 113 infants, utilising EEG recordings taken during the first year of life while infants listened to sung or chanted nursery rhymes. The neural measures are used to explore relations with subsequent language acquisition using a battery of standardised and experimental phonological, gestural and lexical measures taken from 12 to 24 months of age (3).

The auditory neuroscience of adult speech processing has revealed that accurate encoding of amplitude modulations (AMs) in the speech signal via neuronal oscillations is fundamental to many different aspects of language comprehension. Regarding development, and building on these neural insights, Temporal Sampling (TS) theory describes how low-frequency neural oscillations may be key to learning and extracting phonological information from the speech signal (4,5). According to TS theory, auditory cortical networks operating at delta, theta, beta and gamma frequencies 'sample' the envelope of the speech signal at matched frequencies, thereby underpinning the phonological (linguistic) encoding of speech by infants and children. Cortical signals have been shown to track various key properties of speech, using methodological advances such as Temporal Response Functions (TRFs), an analysis technique that explores how neural signals encode continuous sensory stimuli. Adult TRF studies have revealed how cortical tracking of the speech envelope reflects both bottom-up and top-down cortical processes in adult listeners, linking to selective attention, prediction, signal parsing and speech intelligibility (6–12). Recent infant EEG data has revealed robust cortical tracking of the speech envelope (13–15), particularly in the delta and theta frequency bands (1). Indeed, patterns of cortical tracking in delta and theta frequency bands make distinct contributions to speech encoding by adults (16), and there is evidence for preferential tracking in the delta band during the first year of life (~0.5Hz–4 Hz, see 1,17). Consequently, language acquisition by infants may depend in part on successful low-frequency cortical tracking of the speech signal.

Further, low-frequency cortical oscillations do not work in isolation from each other but show dynamic interactions during speech processing. There is a nested hierarchy of cortical oscillations that match the frequency of multiple components of the speech signal (18). Low-frequency (delta, theta) neural phase dynamics temporally organize the amplitude of high-

frequency signals in a process known as phase-amplitude coupling (PAC; (19–21)). Adult electrophysiological and MEG recordings have shown that PAC between delta-gamma, delta-beta and theta-gamma is induced in response to continuous speech (18,22,23), as well as to rhythmic speech stimuli (sung and chanted nursery rhymes, (17)). PAC is thought to group speech information into linguistic units (24,25). Like natural speech, rhythmic speech stimuli induce PAC in infants and adults (1). Interestingly, PAC emerges as early as 4 months of age and remains stable, while both delta- and theta-band cortical tracking show developmental changes over the first year of life (1). Accordingly, PAC may also relate to later language outcomes.

Studies relating infant cortical tracking to later language outcomes are rare (26). However, research on the role of cortical dynamics in language development has progressed by studying older children with language difficulties. For example, recent EEG modelling studies (27) have shown that delta-theta PAC is atypical in children with developmental language disorder (DLD) while the theta/delta power ratio during passive story listening is atypical in children with developmental dyslexia. Children with dyslexia have poor phonological processing skills while children with DLD have poor syntactic skills. Accordingly, both PAC and the theta/delta power ratio in infants may predict subsequent language development, in addition to low-frequency cortical tracking. The developmental theoretical framework encapsulated by TS theory informed Araujo et al.'s (27) modelling (Pre-print), and predicts that cortical tracking within specific low-frequency bands (i.e. delta, 1–4 Hz and theta, 4–8 Hz) and their interactive dynamics is central to language acquisition (4,28,29). In brief, if low-frequency cortical tracking and/or dynamic relations between frequency bands are atypical, then developmental trajectories for language may also be atypical. Previous electrophysiological studies testing TS theory with children have revealed a significant role for delta band encoding in both phonological and

93 vocabulary development (30–33). However, PAC (delta-beta PAC) was not significantly
94 different in children with dyslexia in a language processing (sentence repetition) task (32).
95 Infants and toddlers at risk for dyslexia differ from not-at-risk infants in measures of acoustic
96 processing that relate to cortical tracking, with implications for later language development
97 (14,38). Understanding how dynamic cortical processes in the infant brain affect language
98 development has yet to be studied in detail.

99 In the current study, we explore the potential role of individual differences in low-frequency
100 cortical tracking, PAC and theta-delta power relations in predicting individual differences in later
101 language acquisition. Participating infants were drawn from the Cambridge UK BabyRhythm
102 project, a longitudinal study of 122 infants (122 recruited, 113 retained) from two- to 42-months-
103 of-age, investigating early neural and motor tracking of auditory and audiovisual rhythms in
104 relation to language outcomes. Here we examine whether increases in low-frequency EEG power
105 spectral density (PSD) in the delta and theta range, low-frequency cortical tracking (delta, theta
106 and alpha) and PAC in response to rhythmic speech predict performance on multiple language
107 and communication measures administered in later infancy and toddlerhood. EEG data was
108 collected longitudinally at 4-, 7- and 11-months and used to predict subsequent performance in
109 phonological, gestural and lexical measures taken from 12 to 24 months of age. It was expected
110 that stronger cortical tracking would be predictive of better language outcomes (26). As PAC did
111 not show strong development with age, no strong predictions were made. However, given the
112 DLD modelling data noted earlier (27), poorer PAC may be linked to worse language outcomes.
113 Further, a higher theta/delta power ratio may be linked to worse language outcomes (27). In
114 Araújo et al., preprint, higher ratios in the dyslexic group were not driven by group differences in

either delta or theta power per se, but by their joint dynamics. Accordingly, we also examine the theta/delta ratio as a potential predictor of language outcomes.

2 Materials and Methods

2.1 Infant participants and ethics

All the data reported here is from the longitudinal Cambridge UK BabyRhythm project (122 infants recruited, 113 retained). The cohort was split into two halves based on recruitment date) to allow for confirmatory analyses of neural factors by comparing EEG data from the first and second halves of the cohort. The current report is from the second half of the sample. Results from the first half of the sample are reported in (1). Due to participant withdrawals, missed appointments and data exclusion the final number of data points for the second half of the sample was, ~4-months (N=56), ~7-months (N=51), ~11-months (N=52). Missing data was due to missed appointments or technical issues with the data files, ~4-months (N=4), ~7-months (N=10) and ~11-months (N=7). Other infants were excluded due to having fewer than 42 data epochs available after pre-processing, ~4-months (N=1), ~7-months (N=0) and ~11-months (N=2).

The two halves of the cohort were combined for the novel language analyses reported here (language outcome data was still being collected when Attaheri et al., 2022 (1) was being prepared). The full cohort provided EEG data for analysis as follows, ~4-months (N=110, aged 115.5 ± 5.0 days), ~7-months (N=104, aged 213.5 ± 6.1 days) and ~11-months (N=107, aged 333.6, ± 5.0 days) [mean ± standard deviation (*SD*)].

Infants were recruited from a medium sized city in the United Kingdom and surrounding areas via multiple means including flyers in hospitals, schools and antenatal classes, research presentations at maternity classes and online advertising. All infants were born full term (37-42

gestational weeks) and had no diagnosed developmental disorder. The study was reviewed by the Psychology Research Ethics Committee of the University of Cambridge. Parents gave written informed consent after a detailed explanation of the study and families were reminded that they could withdraw from the study at any point during the repeated appointments (8 EEG recordings at 2-, 4-, 5-, 6-, 7-, 8-, 9- and 11-months; 6 language follow-ups at 12-, 15-, 18-, 24-, 30- and 42-months).

2.2 Stimuli

A selection of 18 common English language nursery rhymes were chosen as the stimuli. Audio-visual stimuli of a singing female (head only) were recorded using a Canon XA20 video camera at 1080p, 50fps and with audio at 4800 Hz. A native female speaker of British English used infant directed speech to melodically sing (for example “Mary Mary Quite Contrary”) or rhythmically chant (for nursery rhymes like “There was an old woman who lived in a shoe”) the nursery rhymes whilst listening to a 120 bpm metronome. Although the nursery rhymes had a range of beat rates and indeed some utilized a 1 Hz rate, the metronome was used to keep the singer on time. The beat was not present on the stimulus videos, but it ensured that a consistent quasi-rhythmic production was maintained throughout the 18 nursery rhymes. To ensure natural vocalisations the nursery rhyme videos were recorded while being sung or rhythmically chanted to an alert infant.

2.3 EEG data collection

Infants were seated in a highchair, approximately one meter in front of their primary care giver, within a sound-proof acoustic chamber. EEG data was recorded at a sampling rate of 1000 Hz using a GES 300 amplifier connected to an appropriately sized 64 channel infant net (Geodesic Sensor Net, Electrical Geodesics Inc., Eugene, OR, USA). The infant was seated

~650mm away from the presentation screen and sounds were presented at 60dB (checked by sound level meter) from speakers (Q acoustics 2020i driven by a Cambridge Audio Topaz AM5 stereo amplifier) placed either side of the screen. Whilst the infant attended to the screen, the 18 nursery rhyme videos played sequentially, each repeated 3 times (54 videos, with a presentation time of 20’ 33’’ in total). If the infant lost attention to the screen an attention grabber video was played at the end of that nursery rhyme. If the infant became too fussy a short break was allowed before continuing, otherwise the session was ended. All infants included for analysis listened to at least 2 repetitions of each nursery rhyme (minimum of 36 nursery rhymes, lasting 13’ 42’’). This stimulus period was followed by 5 minutes of silent recording (hereafter silent state). To ensure compliance from the infant, and therefore a good EEG signal during the silent state, it was necessary for a researcher to sit alongside the infant. To ensure consistency across participants, the researcher performed the same action of silently blowing bubbles and showing the same picture book to the infant during this period.

2.4 EEG preprocessing

All analyses were conducted with custom-made scripts in Matlab 2017a (The MathWorks, Inc., Natick, MA) incorporating the EEGLab toolbox (36).

The four (channels 61, 62, 63 and 64) facial electrodes (present on the larger nets) were excluded from all analyses because they do not exist on the infant-sized EGI Geodesic sensor nets. The EEG data, from the remaining 60 channels, was filtered (*pop_eegfiltnew* function of EEGLab toolbox) into a broadband signal (0.5-45 Hz for PSD and PAC, methods for acquiring PAC bands of interest further detailed in (1)) or the frequency band of interest (delta, 0.5-4 Hz, theta, 4-8 Hz or alpha [control band] 8-12 Hz for the mTRF analyses) using zero-phase bandpass Hamming windowed FIR filters (transition band widths of 2 Hz with cutoff frequencies at -6 dB,

183 0-5 Hz, 3-9 Hz and 7-13 Hz respectively). The EEG data was down sampled to 100 Hz to reduce
184 the computational load. Next, the *clean_asr* EEGLab function (36,37) was used to clean noise
185 artifacts from the data by identifying and removing bad principal components via a modified
186 PCA procedure (see (1)). Further bad channels were identified via probability and kurtosis and
187 were interpolated (via spherical interpolation), if they were 3*SD* away from the mean (average
188 number of interpolated channels; 4mo = 6.9, 7mo = 6.5, 11mo = 6.6) before all channels were re-
189 referenced to a 60-channel average reference.

190 EEG responses were epoched into trials aligned to the start and ending at the completion
191 of a nursery rhyme phrase (e.g. "Mary had a little lamb"), producing EEG responses to 83
192 phrases (*M length* ± *SD*: 4.23sec ±0.88) which were repeated a maximum of 3 times in the
193 experiment (249 epochs in total). This epoching procedure was decided as infant EEG typically
194 contains short, irregular, movement artifacts and using shorter epochs increases the likelihood of
195 generating a whole artifact-free epoch whilst maintaining long enough epochs for optimising the
196 model fit with the mTRF toolbox (38). The infants occasionally rested their head or touched the
197 EEG cap in such a way that specific channels showed short bursts of noise whilst the rest of the
198 channels remained clean. To account for this, and retain a larger volume of data, epoch by epoch
199 channel rejection was conducted. Per epoch, probability and kurtosis were used to identify bad
200 channels and were interpolated (via spherical interpolation) if they were 3*SD* away from the
201 mean. Finally, bad epochs were rejected with the *pop_autorej* function (EEGLab), removing
202 epochs with fluctuations above 1000uV and values outside a 3*SD* of the probability threshold.

203 2.5 Multivariate Temporal Response Function (mTRF)

204 TRFs describe the linear relationship between an input and an output signal, taking into
205 account that such a relationship may not be instantaneous and could extend over a certain

window of time (38). TRFs can be estimated with system identification methodologies, such as the multivariate temporal response function (mTRF), which is a de-convolution method based on multiple lagged linear regression. Here, we applied the mTRF in the backward direction to reconstruct the stimulus envelope from the EEG signals. In turn, this informs us on the strength of the stimulus-EEG relationship, which we take as an index of cortical tracking of the stimulus envelope. The backwards 'stimulus reconstruction' mTRF model has important advantages compared to the forward linear encoding models which was previously used in infants study (13,14). Specifically, backward TRFs combines data from all the EEG channels simultaneously in a multivariate manner to reconstruct the univariate stimulus envelope, while forward TRF models only predict one EEG channel at a time. A Pearson's correlation value was calculated for each participant, representing how well the model reconstructed the stimulus envelope (a metric for cortical tracking), when trained with either the delta (0.5-4Hz), theta (4-8Hz) or alpha (8-12Hz) EEG band data. See (1) for full a detailed description of the mTRF method and analysis.

2.6 mTRF auditory stimuli preprocessing

The envelope of the auditory signal was extracted by taking the absolute value of the analytic signal generated by the Hilbert transform (Matlab). As the envelope of the lower frequencies is linearly relatable to the EEG signal (39–41) the envelope of the stimuli was filtered between 0.5 Hz and 15 Hz (lowpass; 6th order Butterworth filter. Highpass; 9th order Butterworth filter). The resultant envelopes were normalised using *nt_normcol* (Noisetools). Finally, the stimulus envelopes were down sampled to 100 Hz to match the EEG signal.

2.7 Phase amplitude coupling (PAC)

All remaining epochs after preprocessing were concatenated back into one continual EEG signal. Due to the increased susceptibility of PAC analysis to noise, manual rejection was conducted to further remove noisy periods of data from the signal.

For each infant's EEG data, low-frequency phase (LFP) and high-frequency amplitude (HFA) were extracted with a zero-phase FIR filter (Hamming window), separately for all 60 electrodes. LFP centre frequencies were filtered from 2 to 8 Hz, in 1 Hz steps with a 2 Hz bandwidth, and HFA centre frequencies were filtered from 17.5 Hz to 42.5 Hz, in 5 Hz steps with a 5 Hz bandwidth before a normalised modulation index value was computed for each channel (19,42). The EEG bands were defined as delta (2-4 Hz), theta (4-8 Hz) beta (15-30 Hz) and gamma (30-45 Hz). The channel exhibiting the strongest nMI, within predefined phase and amplitude band groupings (delta/beta, delta/gamma, theta/beta, theta/gamma), was taken forward for further analysis. See (1) for a detailed description of the PAC method and analysis.

2.8 Spectral analysis (periodogram PSD estimate)

The same concatenated data sets created for the PAC analysis were also used for the periodogram Power Spectral Density (PSD) estimates, in which all epochs after preprocessing were concatenated back into one continual signal with manual rejection used to remove noisy periods of data from the signal. A one-sided PSD estimate was conducted separately for each electrode channel using the periodogram function (Matlab). The average PSD across channels was next calculated with the maximum value, within a predefined 0.25Hz window centered on the peak of interest, taken per participant for further analysis. See (1) for a detailed description of the PSD method and analysis.

2.8.1 Theta / delta (PSD) ratio

The PSD peak per participant at 4.35 Hz and 1.95 Hz, taken from the spectral decomposition of the EEG signal described above, was used to create a theta / delta ratio calculation. Per participant, the 4.35 Hz PSD value was divided by the 1.95 Hz PSD value (4.35/1.95) to give a theta / delta ratio per participant.

2.9 Language measures

A full description of the language measures used here can be found in (3). The parent-estimated measures were taken from a carer-completed questionnaire, the UK-CDI. The UK-CDI estimates both receptive and productive vocabulary (43), which we refer to here as language comprehension or language production. In the current report, the 24-month CDI measures are used (see Rocha et al., preprint (3), for rationale). Only measures up to the age of 24 months were included, as the 30- and 42-month measures are still being processed. Five measures, taken from 3 of our experimental tasks, were included as the infant-led language measures. These were 1, gesture (adapted from (44)), in which pointing towards a present and absent target was used a pre-verbal measure of communicative intent at 12 months; 2, the computerized comprehension task (CCT), a touch screen-based measure of infant vocabulary in which the infant is asked to select a picture of a named item (45), and 3, a custom designed non-word repetition (NWR) task, which measured the child's productive phonology. Three measures were taken from the NWR task administered at 24 months, (i) accurate reproduction of consonants, (ii) number of syllables reproduced, and (iii) stress patterns (canonical or non-canonical, see Rocha et al., preprint (3)). This NWR task measured phonological production of both nonsense words (e.g. "punky") and real words (e.g. "puppy") to name a series of toys.

2.10 Language outcomes: statistics

Multivariate linear models (using Rstudio *lm* function) were run separately for the PSD, mTRF, theta-delta ratio and PAC measures. Due to the inherent differences between the infant-led and parent-estimated measures, separate multivariate linear models were conducted for each, providing a general measure of language abilities taken from either infant-led or parent-estimated language measures. This approach was considered most appropriate as additional variance can be extracted from an infant's overall language performance which may not be visible by looking at individual measures separately. However, we also include Bonferroni-corrected individual univariate linear models to highlight the individual contribution of each language measure to the overall infant-led or parent-estimated multivariate linear model results.

3 Results

3.1 PSD, mTRF, PAC and theta/delta ratio

First, we present the power spectral density (using a periodogram; PSD), cortical tracking (using multivariate temporal response functions; mTRFs), phase-amplitude coupling (normalised modulation index, nMI; PAC) and theta/delta PSD ratio analyses using EEG recorded from the second half of our cohort of infants. We compare these data to the analyses from the first half of the cohort (1), and also merge the neural data to explore these neural factors in the combined full sample of 113 infants. Due to missed recording sessions, technical issues with data files and not enough trials after preprocessing, less than 113 data points are included in each analysis. Furthermore, outlier analysis was also conducted to remove extreme data points that would compromise each of the LMEMs. The number of data points included are given separately for each analysis.

3.1.1 Power spectral density (PSD) analyses

Figure 1, Spectral decomposition of the EEG signal in the a) first half, b) second half and c) full infant cohort (1–12 Hz plotted from 0.5 to 45 Hz calculated), in response to nursery rhyme stimulation. A periodogram was used to obtain a power spectral density (PSD) estimate separately for 4- (red), 7- (green) and 11- (blue) months data. Bold lines indicate the mean values and pale shading plots the standard deviation of the data. Outlier analysis was also conducted to remove extreme data points that would compromise the LMEM. Panel A is reproduced from Attaheri et al. 2022 (1), with permission.

The distribution of low-frequency oscillations in the EEG recorded from the second half of the sample and the full cohort were obtained using the power spectral density (PSD) estimate (Fig. 1). The results from the first half of the cohort (1) are reproduced for comparison. Given the differences in the location of the PSD peaks for the first and second halves of the cohort, the distribution of PSD peaks in low-frequency power for the full sample were first established by creating a grand average PSD estimate across channels, recording type and ages. This revealed three frequency peaks centred around 1.92 Hz, 4.05 Hz and 4.35 Hz (Fig. S1). The distribution of the PSD peaks were similar to those observed in the first half of the sample (see (1)) and in the second half of the sample (Fig. S2). It is important to note that when the full cohort was analysed, the ~4Hz peak split into two locations, peaking at 4.05 Hz and 4.35 Hz. To further investigate whether the PSD was significantly higher in response to the nursery rhymes compared to the silent state, and whether this relationship was affected by the age of the participants, a linear mixed effects model (LMEM) was conducted separately for the 1.92 Hz, 4.05 Hz and 4.35 Hz peaks. A random intercept (AR1) per participant was included to account for individual differences across the 3 recording sessions (4-, 7- and 11-months). The full sample submitted to the LMEM was as follows, 4-months (N=96), 7-months (N=88) and 11-months (N=92). This process was conducted separately for the first half of the sample, the second half of

the sample and the full cohort (Table 1). To allow for individual variation across participants, a maximum peak value per participant was taken from the participants' 60 channel averaged data, within 0.25 Hz windows centred around the peak of interest (1.92 Hz, 4.05 Hz and 4.35 Hz). Note that a smaller window was used around the peaks of interest than used in Attaheri et al., 2022 (1), this was necessary to avoid overlap between the 4.05 Hz and 4.35 Hz windows. Fixed effects of recording type (nursery rhymes, silent state), age (4-, 7- or 11-months) and a recording type by age interaction were included in the model. The ~4.35 Hz peak is reported first, as this was the most robust.

~4.35 Hz

	1st half of sample		*2nd half of sample*	
	F	*p*	*F*	*p*
Recording type	F(1, 124.16 = 12.51)	***0.001***	F(1, 107.77 = 7.67)	***0.007***
Age	F(2, 153.41 = 10.01)	***$7.7x10^{-5}$***	F(2, 114.77 = 4.86)	***0.009***
Recording type * Age	F(2, 123.93 = 3.30)	***0.04***	F(2, 107.90 = 3.54)	***0.032***

	Full sample	
	F	*p*
Recording type	F(1, 236.60 = 19.25)	***$1.7x10^{-5}$***
Age	F(2, 281.12 = 14.13)	***$1.0x10^{-6}$***
Recording type * Age	F(2, 236.48 = 6.95)	***0.001***

Table 1. Tests of fixed effects for LMEM conducted on PSD power peak at 4.35Hz. Results reported for separate LMEMs run on the first half, second half and full sample of data. Significant results (p<0.05) denoted by ***bold italic text***. The Satterthwaite approximation was applied to approximate the degrees of freedom, due to the missing cases.

The 4.35 Hz peak showed the largest stimulus-related increase in PSD of the three peaks analysed. The LMEM's revealed a significant effect of recording type, with the nursery rhyme (NR) condition producing significantly greater PSD power than the silent state (SS; Table 1). A significant effect of age was observed, with ~4.35 Hz power increasing parametrically with age from 4- months to 7- months to 11- months (estimates of fixed effects provided in Table S1). The observed recording type by age interaction was due to a developmental increase in NR induced

power above the SS. Bonferroni-corrected pairwise comparisons showed that PSD increased significantly between 4- and 11-months (M increase = 18.41, SE ± 7.43, $p = 0.043$) as well as between 4- and 7- months (M increase = 30.56, SE ± 5.77, $p = 6.7\text{x}10^{-7}$). Hence there was a developmental increase in 4.35 Hz PSD for both the nursery rhymes and the silent state. This may indicate age-related maturation of theta activity in general. Importantly, we also observed a stimulus-driven interaction with age for the 4.35 Hz PSD peaks, with the largest increase observed at 11-months (4-months, NR, $M = 38.9$ $SE \pm 4.9$; SS, $M = 37.9$ $SE \pm 5.5$. 7-months, NR, $M = 63.0$ $SE \pm 6.9$; SS, $M = 50.6$ $SE \pm 7.2$. 11-months, NR, $M = 84.8$ $SE \pm 5.8$; SS, $M = 53.1$ $SE \pm 6.0$). This showed that the interaction was driven by a greater PSD in response to the nursery rhymes at 11-months. This pattern replicated the theta PSD findings observed in the first half of the sample (see (1)). In summary, the peak at 4.35 Hz showed robust replication across the first and second half of the sample, with consistent increases in NR related power displaying a developmental increase in PSD magnitudes, above SS.

~4.05 Hz

	1st half of sample		*2nd half of sample*	
	F	*p*	*F*	*p*
Recording type	F(1, 223.47 = 2.72)	0.100	F(1, 101.13 = 5.65)	***0.019***
Age	F(2, 160.05 = 4.74)	***0.010***	F(2, 111.74 = 2.39)	0.097
Recording type * Age	F(2, 223.50= 1.07)	0.345	F(2, 101.07 = 1.50)	0.228

	Full sample	
	F	*p*
Recording type	F(1,424.40 = 7.33)	***0.007***
Age	F(2,298.70 = 3.54)	***0.030***
Recording type * Age	F(2, 424.16 = 1.77)	0.172

Table 2. Tests of fixed effects for LMEM on PSD power peak at 4.05Hz. Results reported for separate LMEMs run on the first half, second half and full sample of data. Significant results (p<0.05) denoted by ***bold italic text***. The Satterthwaite approximation was applied to approximate the degrees of freedom, due to the missing cases.

For the 4.05 HZ peak, the LMEM results replicated less well across the first half of the sample, the second half of the sample and the full cohort. Significant effects of both recording type and age were only observed when analysing the full cohort. The significant effect of age was driven by the 7-month (*M* = 63.1 *SE* ± 6.0) PSD power being larger than the 11-month (*M* = 52.5 *SE* ± 3.2), however this result is in part due to the 0.25 Hz window not fully encapsulating the 11-month peak for all participants (Fig. 1). A significant effect of recording type was observed in the second half of the sample, with a trend observed in the first half of the sample (Table 2). A significant effect of age was observed in the first half of the sample with a trend towards significance observed in the second half of the sample. As the data were more robust for the 4.35 HZ peak, and as the 0.25 Hz window did not fully capture the 4.05 Hz peak for all participants, the 4.35 Hz PSD peak was taken forward as the 'theta' peak for analyses with the language data.

~1.92 Hz

	1st half of sample		*2nd half of sample*	
	F	*p*	*F*	*p*
Recording type	F(1, 96.59 = 9.71)	***0.002***	F(1, 87.49 = 0.25)	0.616
Age	F(2, 127.50 = 1.00)	0.371	F(2, 109.00 = 3.34	***0.039***
Recording type * Age	F(2, 96.74 = 0.61)	0.544	F(2, 87.50 = 7.39)	***0.001***

	Full sample	
	F	*p*
Recording type	F(1,188.40 = 8.47)	***0.004***
Age	F(2,232.10 = 0.36)	0.698
Recording type * Age	F(2,96.74 = 0.61)	***0.021***

Table 3. Tests of fixed effects for LMEM on PSD power peak at 1.92Hz. Results reported for separate LMEMs run on the first half, second half and full sample of data. Significant results (p<0.05) denoted by ***bold italic text***. The Satterthwaite approximation was applied to approximate the degrees of freedom, due to the missing cases.

For the 1.92 Hz peak, the LMEM conducted on the full cohort revealed a significant effect of recording type (Table 3). However, there were differences in this effect across the first

and second halves of the sample. Upon further inspection, the non-significant level of recording type in the second half of the sample was due to the large amount of PSD power in the silent state (SS) data at 4-months (4-month SS, M = 116.3 $SE \pm 10.6$, NR, M= 97.3 $SE \pm 10.2$). Posthoc paired sample t-tests on the 7- and 11-month data both showed a consistent stimulus-induced increase in PSD power, matching the first half of the sample (7-month SS, M= 116.3 $SE \pm 13.5$, NR, M= 124.2 $SE \pm 13.1$, p = 0.007; 11-month SS, M= 94.0 $SE \pm 11.6$, NR, M=76.7 $SE \pm 11.9$, p = 1.90×10^{-27}). A significant recording type by age interaction was observed in both the second half of the sample and the full cohort. In conclusion, the 1.92Hz stimulus-related increase in PSD power is broadly replicated across the first and second halves of the sample once the anomalous SS power, present in the 4-month data in the second half of the sample, was taken into account.

3.1.2 Cortical tracking (mTRF) analyses

Figure 2, mTRF replication. Grand averaged stimulus reconstruction values (Pearson's r) per frequency band in the a) first half, b) second half and c) full infant cohort. Each bar represents the grand average r value across the participants at the 3 age points: 4-months (red), 7-months (green) and 11-months (blue). Age responses are grouped along the X axis by EEG frequency band, delta (0.5–4 Hz) theta (4–8 Hz) and alpha (8–12 Hz). Light grey boxes signify the average random permutation value across subjects, calculated separately for each band. Outlier analysis was conducted to remove extreme values which would compromise the LMEMs. Panel a is reproduced from Attaheri et al. (2022), with permission.

A linear mixed effects model (LMEM) was constructed using the same parameters as described in Attaheri et al., 2022 (1) to examine potential developmental changes in cortical tracking in the delta, theta and alpha (control) bands. The full sample submitted to the mTRF LMEM was as follows, 4-months (N=108), 7-months (N=104) and 11-months (N=106). Main effects of data type (real r values, random permutation r values), frequency band (delta, theta or alpha) and age (4-, 7- or 11-months) were investigated along with interactions between data type

397 by frequency band, data type by age and age by frequency band. A random intercept (AR1) per
398 participant was included to account for individual differences across the 3 recording sessions (4-,
399 7- and 11-months).

400 The main pattern of results replicated between the first and second halves of the sample,
401 with significant fixed effects of data type and frequency band, as well as significant interactions
402 between data type by frequency band and age by frequency band (Table 4). There were no
403 significant effects of age nor data type by age interactions when analysing the second half of the
404 sample in isolation, although these effects were present in the first half of the sample (Table 4).
405 When using the full cohort, tests of fixed effects showed significant effects of data type,
406 frequency band and age. These results indicate that the real *r* values (cortical tracking) in
407 response to nursery rhymes were significantly above randomly permuted *r* values and that the
408 cortical tracking values differed by frequency band and age. A significant interaction between
409 data type and age indicated that the difference between random and real *r* values varied by age.
410 Examination of the estimates of fixed effects (Table S2) revealed that the large, above chance,
411 increase in cortical tracking at four months (average across all bands, random permuted *r* values,
412 $M = 0.12$, $SE \pm 0.001$; real *r* values, $M=0.19$, $SE \pm 0.001$) was driving the interaction. The
413 significant interaction of frequency band by data type indicated that cortical tracking in the delta
414 and theta bands was driven by their values being significantly above the randomly permuted *r*
415 values (as compared to the alpha base case; Table S2). The marginal means across the three ages,
416 revealed that cortical tracking estimates (*r* values) were significantly higher in both delta
417 (random permuted *r* values, $M = 0.18$, $SE \pm 0.001$; real *r* values, $M = 0.031$, $SE \pm 0.001$), and
418 theta bands (random permuted *r* values, $M = 0.12$, $SE \pm 0.001$; real *r* values, $M = 0.016$, $SE \pm$
419 0.001), compared to the alpha base case (random permuted *r* values, $M = 0.006$, $SE \pm 0.001$; real

r values, *M* = 0.007, *SE* ± 0.001). Furthermore, alpha tracking was again not significantly above chance (Table S2) when using the full sample, mimicking the results from the first half of the sample (1). The significant interaction between age and frequency band showed that the relative difference in *r* values between the three age groups were different in the three frequency bands. This was predominantly driven by 4-month delta (*M* = 0.27, *SE* ± 0.001) producing larger *r* values than the 11-month delta (*M* = 0.22, *SE* ± 0.001) and the 7- month delta *(M* = 0.24, *SE* ± 0.001).

In summary, all the main effects reported by Attaheri et al. 2022 (1) were replicated in the second half of the sample, with only age effects varying. This replication strengthens the conclusion that delta and theta EEG responses track the acoustic envelope of sung and chanted speech from the age of 4-months and that this infant neural response is robust and reliable.

Cortical tracking

	1st half of sample		*2nd half of sample*	
	F	*p*	*F*	*p*
Data type	F(1, 901.80) = 118.66	***4.81× 10^{-26}***	F(1, 793.07 = 80.29)	***2.28x10^{-18}***
Age	F(2, 755.51) = 5.07	***0.006***	F(2, 253.92 = 1.36)	0.258
Frequency band	F(2, 903.22) = 401.94	***1.40 × 10^{-125}***	F(2, 794.49 =277.10)	***5.05x10^{-92}***
Data type * age	F(2, 901.80) = 3.41	***0.033***	F(2, 793.07 = 1.04)	0.354
Data type * freq. band	F(2, 901.80) = 47.66	***2.11 × 10^{-20}***	F(2, 793.06 =28.16)	***1.53x10^{-12}***
Age *freq. band	F(4, 755.51) = 5.08	***0.006***	F(4, 794.54 = 2.55)	***0.038***

	Full sample	
	F	*p*
Data type	F(1, 1636.52 = 192.77)	***1.62x10^{-41}***
Age	F(2, 691.86 = 5.22)	***0.006***
Frequency band	F(2, 1646.47, = 658.70)	***6.65x10^{-211}***
Data type * age	F(2, 1636.52 = 3.85)	***0.021***
Data type * freq. band	F(2, 1636.52 = 72.18)	***9.13x10^{-31}***
Age *freq. band	F(4, 1645.40 = 5.43)	***2.44x10^{-4}***

Table 4. Tests of fixed effects for LMEM of cortical tracking across data type (real r values, random permutation r values), frequency band (delta, theta or alpha) and age (4-, 7- or 11-

433 months). Results reported for separate LMEMs run on the first half, second half and full sample
434 of data. Significant results ($p<0.05$) denoted by ***bold italic text***. The Satterthwaite approximation
435 was applied to approximate the degrees of freedom, due to the missing cases.

436 **3.1.2 Phase amplitude coupling (PAC) analysis**

437

438 Figure 3, PAC replication. Violin plots of phase amplitude coupling (normalised modulation
439 index, nMI) at 4-months (red), 7-months (green) and 11-months (blue) for the a) first half, b)
440 second half and c) full infant cohort. The nMIs, from all significant analysis windows, were
441 averaged for each individual infant. The PAC pairing with the maximum nMI per infant, from
442 within the pre-defined frequency bands of interest (delta/beta, delta/gamma, theta/beta and
443 theta/gamma; beta 2–4 Hz, theta 4–8 Hz, beta 15–30 Hz and gamma 30–45 Hz), were included
444 in the grand average violin plot. Horizontal black and red lines represent the group mean and
445 median respectively. Outlier analysis was conducted to remove extreme data points that would
446 compromise the LMEM. Panel A is reproduced from Attaheri et al. 2022 (1), with permission.

447 The PAC analyses (Fig. 3) were conducted for the second half of the sample and the full
448 cohort, replicating the LMEM conducted with the first half of the sample (1). The full sample
449 submitted to the PAC LMEM was as follows, 4-months (N=108), 7-months (N=103) and 11-
450 months (N=107). This LMEM included fixed effect factors of low-frequency phase (delta or
451 theta), high-frequency amplitude (beta or gamma) and age (4-, 7- or 11-months). Interactions of
452 low-frequency phase by age, high-frequency amplitude by age and low-frequency phase by high-
453 frequency amplitude were also included in the model, with a random intercept (AR1) per
454 participant (see methods section and Attaheri et al., 2022 (1) for detailed methods). Tests of fixed
455 effects are reported in Table 5 and estimates of fixed effects are provided in Table S3. It is
456 important to note that all the normalised modulation index (nMI) values included in the LMEM
457 displayed significant coupling above a randomly permuted surrogate distribution, thus the
458 LMEM intended to explore whether the nMI values were significantly different when either delta
459 or theta was the low-frequency carrier phase or when either beta or gamma were the high-
460 frequency amplitudes, and if this relationship was affected by age.

Analysis of the full cohort revealed a significant effect of high frequency amplitude (HFA), demonstrating that the nMI values where larger when gamma (M = 2.929, SE ± 0.31) was the HFA, rather than beta (M = 2.820, SE ± 0.31), when coupling with low-frequency phases (Table 5). The non-significant effect of low-frequency phase (LFP) demonstrated that the gamma amplitudes coupled equally as strongly with delta and theta phases (Table 5). The main effect of high frequency amplitude (HFA) observed in the first half of the sample and the full cohort did not replicate between the first and second half of the sample. However, there was a significant LFP by HFA interaction in the second half of the sample. This arose because the relationship between the LFPs and the HFAs varied by frequency band. When delta was the LFP, delta-beta (M = 2.884, SE ± 0.063) PAC was stronger than delta-gamma PAC (M = 2.827, SE ± 0.063). This pattern reversers when theta was the LFP with theta-gamma (M = 2.961, SE ± 0.063) demonstrating stronger coupling than theta-beta (M = 2.737, SE ± 0.063). Posthoc analysis revealed that the LFP by HFA interaction was driven most strongly by the strength of the theta-gamma nMI values. A posthoc paired sample t-test revealed theta-gamma was significantly greater than theta-beta ($p = 0.025$), whilst there was no significant difference between delta-beta and delta-gamma ($p = 0.324$). Although the main effect of HFA did not replicate between the first and second halves of the sample, theta-gamma was consistently greater than theta-beta across the first and second halves of the sample. All other fixed effects were not significant in either half of the sample or together in the full sample (Table 5). In summary, significant PAC was present at all ages. When delta was the carrier phase, equal PAC was found for both beta and gamma bands. However, when theta was the carrier phase, there was stronger PAC with the gamma band amplitudes.

PAC

	1st half of sample		*2nd half of sample*	
	F	*p*	*F*	*p*
Age	F(2, 311.00) = 0.29	0.749	F(2,175.074) = 2.70	0.070
Low freq. phase (LFP)	F(1,506.35) = 0.81	0.369	F(1,412.165) = 0.01	0.908
High freq. amp (HFA)	F(1, 506.35) = 6.35	***0.012***	F(1,411.251) =2.19	0.140
Age * LFP	F(2, 506.34) = 0.26	0.768	F(2,412.122) = 0.75	0.473
Age * HFA	F(2, 506.34) = 2.62	0.074	F(2,411.151) = 0.32	0.728
LFP * HFA	F(2, 506.35) = 0.01	0.913	F(1,411.251) = 6.25	***0.013***
Age * LFP * HFA	F(2, 506.34) = 0.50	0.606	F(2,411.151) = 0.30	0.744

	Full sample	
	F	*p*
Age	F(2,438.48) = 2.16	0.117
Low freq. phase (LFP)	F(1,912.95) = 0.37	0.541
High freq. amp (HFA)	F(1,912.02) = 8.02	***0.005***
Age * LFP	F(2,912.75) = 0.42	0.658
Age * HFA	F(2,911.75) = 2.45	0.087
LFP * HFA	F(1,912.02) = 3.33	0.068
Age * LFP * HFA	F(2,911.75) = 0.025	0.975

Table 5, Tests of fixed effects, from LMEM, assess how PAC varied by low frequency phase (LFP), high frequency amplitude (HFA) and age (4-, 7- or 11-months). Results reported for separate LMEMs run on the first half, second half and full sample of data. Significant results ($p<0.05$) denoted by ***bold italic text***. The Satterthwaite approximation was applied to approximate the degrees of freedom, due to the missing cases.

3.1.4 Theta/delta PSD power ratio

A significant difference between the theta/delta PSD power ratio across the three ages was also found (Fig. 4), indicated by a one-way ANOVA ($F(2,273) = 9.54$, $p = 9.389\text{x}10^{-5}$). The full sample submitted to the ANOVA was as follows, 4-months (N=96), 7-months (N=88) and 11-months (N=92). The amount of theta PSD power increased relative to the amount of delta PSD power as our sample aged from 4- to 11- months. For comparison, the child modelling using EEG recorded from children with and without dyslexia aged around 9 years showed that a higher theta/delta ratio was associated with poorer language outcomes (27).

Figure 4, Theta / delta PSD ratio. The PSD peak per individual at 4.35 Hz and 1.95 Hz, taken from the spectral decomposition of the EEG signal, was used to create a theta / delta ratio calculation (4.35/1.95). Data provided at 4-months (red), 7-months (green) and 11-months (blue) with standard error bars in black.

3.2 Neural measures as predictors of language acquisition

Multivariate linear models were used to test whether the neural findings reported in 3.1 (for the full sample) were related to the different language measures outlined in the methods above. These language measures were selected as most reliable in previous analyses of the Cambridge UK BabyRhythm cohort, as they had the largest number of infants contributing data and scores were neither at floor nor ceiling (3). As noted, the parent-estimated measures were receptive and productive vocabulary scores from the UK-CDI (43) and the infant-led measures were (1) The ability to point at 12-months, (2) Word recognition at 18-months on a computerised comprehension test (CCT), and (3) Phonological processing at 24-months, as measured by non-word repetition (NWR). Due to the inherent difference in parent-estimated and infant-led measures of language, separate multivariate linear models were run for each set of scores. The results of these multivariate analyses for global parent-estimated or infant-led outcomes are reported in Tables 6-11 along with posthoc univariate linear models. The univariate analyses enable estimation of the individual contribution of each language measure to the global parent-estimated or infant-led language performance. Due to the multiple comparisons involved, the post hoc univariate tests utilised Bonferroni correction, leading to modified alpha significance levels for infant-led ($\alpha < 0.01$) and parent-estimated ($\alpha < 0.025$) measures. The univariate analyses often included more cases, as the global analyses only included infants who contributed data points to all language measures and the neural measure being analysed.

521 **3.2.1. PSD and language outcomes**

522 The 11-months EEG data were selected for comparison to both the parent-estimated and
523 infant-led language measures, due to the observed developmental maturation of the ~4 Hz peak.
524 The analyses for the 4-month and 7-month EEG are provided as Tables S4 to S10. As described
525 above, the 4.35 Hz peak was the more robust and consistent of the two theta peaks (see Tables 1
526 & 2) and therefore was used for predicting the language outcomes (4.05 Hz peak results are also
527 reported in Tables S6-S8). The analyses for the ~1.92 Hz peak are also reported here. 77 infants
528 provided complete data for all the parent-estimated measures and PSD power results, whilst 70
529 infants provided data for all the infant-led language measures and PSD power.

530 Regarding global language outcomes, the multivariate models indicated that increased
531 ~1.92 Hz PSD power in response to nursery rhymes did not lead to significant prediction of
532 either infant-led ($F(5,65) = 1.257$, $p = 0.285$) nor parent-estimated ($F(2,75) = 0.560$, $p = 0.574$)
533 outcomes (Table 6). However, the posthoc univariate linear models revealed a trend for pointing
534 at 12-months to be correlated with increases in ~1.92 Hz power (Table 1a; $p = 0.014$, Bonferroni
535 corrected trend α values, $p = <0.02$ for infant-led univariate models).

536 The multivariate analyses for ~4.35 Hz PSD power in response to nursery rhymes
537 showed that increased power led to a significant decrease ($F(2,75) = 3.18$, $p = 0.047$) in the
538 global parent-estimated language outcomes (Table 7). It is important to note that whilst the
539 posthoc univariate linear model analyses indicated that comprehension and production did not
540 individually make significant contributions to this prediction (Table 7b), the direction of the beta
541 estimates suggest that higher power was associated with fewer known words later in
542 development. Taken together, these data suggest that there is a global effect of theta power on
543 parent-estimated (UK-CDI) language outcomes, but that this is not driven by either single parent-

544 estimated measure. By contrast, infant-led language outcome measures were not significantly
545 predicted by the theta (~4.35 Hz) peak for either the global multivariate linear model (F(5,65) =
546 1.758, p= 0.134) nor the individual posthoc univariate models (Table 7a)

~1.92 Hz PSD power at 11 months

a, Infant-led language measures

Multivariate linear model	Df_{num}	Df_{denom}	F	p
Infant-led language measures	5	65	1.257	0.29
Univariate linear models (post hoc)	Df	$\beta_{estimate}$	SE	p
NWR Consonants, 24 months	1,79	5.23×10^{-5}	8.50×10^{-5}	0.95
NWR Syllables, 24 months	1,79	5.54×10^{-4}	8.56×10^{-4}	0.53
NWR Stress, 24 months	1,79	5.51×10^{-4}	7.83×10^{-4}	0.48
CCT, 18 months	1,79	7.16×10^{-4}	6.32×10^{-4}	0.26
Pointing, 12 months	1,90	0.003	0.001	**0.014**

b, Parent-estimated language measures

Multivariate linear model	Df_{num}	Df_{denom}	F	p
Parent-estimated language measures	2	*75*	0.560	0.57
Univariate linear models (post hoc)	Df	$\beta_{estimate}$	SE	p
CDI comprehension, 24 months	1,76	0.373	0.445	0.40
CDI production, 24 months	1,76	0.179	0.465	0.70

547 Table 6, Multivariate and univariate linear models, investigating whether ~1.92 Hz PSD power at
548 11 months predicted either infant-led or parent-estimated language outcomes. The table details
549 the multivariate linear models (reporting the global effects) and the post-hoc univariate linear
550 models (ran after the main multivariate model) describing whether ~1.92 Hz PSD power at 11
551 months predicted, a) infant-led language measures or b) the parent-estimated language measures.
552 Bonferroni correction for multiple comparisons in the univariate models led to modified
553 significant alpha levels of, p = <0.01 for infant-led and p = <0.025 for parent-estimated (denoted
554 by ***bold italic text***). Modified alpha level for trends are p = <0.02 for infant-led and p = <0.05 for
555 parent-estimated (denoted by **bold text**).

~4.35 Hz PSD power at 11 months

a, Infant-led language measures

Multivariate linear model	Df_{num}	Df_{denom}	F	p
Infant-led language measures	5	65	1.76	0.13
Univariate linear models (post hoc)	Df	$\beta_{estimate}$	SE	p
NWR Consonants, 24 months	1,79	-4.93×10^{-4}	3.69×10^{-4}	0.19

NWR Syllables, 24 months	1,79	-8.36×10^{-4}	3.65×10^{-4}	0.025
NWR Stress, 24 months	1,79	-5.31×10^{-4}	3.40×10^{-4}	0.12
CCT, 18 months	1,79	-1.62×10^{-4}	2.76×10^{-4}	0.56
Pointing, 12 months	1,90	-2.95×10^{-4}	-5.98×10^{-4}	0.62

b, Parent-estimated language measures

Multivariate linear model	Df_{num}	Df_{denom}	*F*	*p*
Parent-estimated language measures	2	75	3.18	***0.047***
Univariate linear models (post hoc)	*Df*	$\beta_{estimate}$	*SE*	*p*
CDI comprehension, 24 months	1,76	-0.086	0.209	0.68
CDI production, 24 months	1,76	-0.346	0.215	0.11

Table 7, Multivariate and univariate linear models, investigating whether ~4.35 Hz PSD power at 11 months predicted either infant-led or parent-estimated language outcomes. The table details the multivariate linear models (reporting the global effects) and the post-hoc univariate linear models (ran after the main multivariate model) describing whether ~4.35 Hz PSD power at 11 months predicted a) infant-led language measures or b) the parent-estimated language measures. Bonferroni correction for multiple comparisons in the **univariate models** led to modified significant alpha levels of, p = <0.01 for infant-led and p = <0.025 for parent-estimated (denoted by ***bold italic text***). Modified alpha level for trends are p = <0.02 for infant-led and p = <0.05 for parent-estimated (denoted by **bold text**).

A further pair of multivariate linear model analyses showed that a larger theta/delta PSD power ratio (4.35 Hz / 1.92 Hz) led to a decrease in performance for both infant-led ($F(5,65) = 2.684$, $p = 0.029$) and parent-estimated ($F(2,75) = 3.65$, $p = 0.031$) global outcomes. The more theta power that was present relative to delta power, the better the infants' performance on either type of outcome. Again, the posthoc univariate linear models suggested that the global effects were not driven by any single individual measure (Table 8).

Theta / delta ratio of PSD power

a, Infant-led language measures

Multivariate linear model	Df_{num}	Df_{denom}	*F*	*p*
Infant-led language measures	5	65	2.68	***0.029***
Univariate linear models (post hoc)	*Df*	$\beta_{estimate}$	*SE*	*p*
NWR Consonants, 24 months	1,79	-0.019	0.032	0.56
NWR Syllables, 24 months	1,79	-0.058	0.032	0.07

NWR Stress, 24 months	1,79	-0.044	0.029	0.14
CCT, 18 months	1,79	-0.023	0.023	0.32
Pointing, 12 months	1,90	-0.065	0.049	0.19

b, Parent-estimated language measures

Multivariate linear model	Df_{num}	Df_{denom}	*F*	*p*
Parent-estimated language measures	2	75	3.65	***0.031***
Univariate linear models (post hoc)	*Df*	$\beta_{estimate}$	*SE*	*P*
CDI comprehension, 24 months	1,76	-6.760	16.44	0.68
CDI production, 24 months	1,76	-28.55	16.82	0.094

Table 8, Multivariate and univariate linear models, investigating whether delta/theta ratio of PSD power (~1.92 Hz/4.35 Hz) at 11 months predicted either infant-led or parent-estimated language outcomes. The table details the multivariate linear models (reporting the global effects) and the post-hoc univariate linear models (ran after the main multivariate model) describing whether delta/theta ratio of PSD power (~1.92 Hz/4.35 Hz) at 11 months predicted a) infant-led language measures or b) the parent-estimated language measures. Bonferroni correction for multiple comparisons in the **univariate models** led to modified significant alpha levels of, p = <0.01 for infant-led and p = <0.025 for parent-estimated (denoted by ***bold italic text***). Modified alpha level for trends are p = <0.02 for infant-led and p = <0.05 for parent-estimated (denoted by **bold text).**

3.2.3 mTRF and language outcomes

The 11-months EEG data was selected for analyses with both the parent-estimated and infant-led language measures, due to it being the closest in time to the language measures (please see supplementary information for the earlier age groupings; these did not reach significance). 89 infants provided complete data for all the parent-estimated measures and a cortical tracking value, whilst 82 infants provided data for all the infant-led language measures and a cortical tracking value.

The multivariate linear models (Table 9) showed that increased delta cortical tracking in response to nursery rhymes displayed a trend towards better language outcomes regarding the parent-estimated language measures ($F(2,87) = 3.00$, $p = 0.055$). This effect was not present for the infant-led measures, ($F(5,77) = 0.83$, $p = 0.350$). Furthermore, the posthoc univariate linear model analyses showed that both CDI comprehension and production at 24 months individually

were significantly predicted by delta band cortical tracking (*p*'s= 0.02, Table 9b). Theta band cortical tracking did not predict either infant-led (F(5,77) = 0.886, *p*= 0.495) or parent-estimated (F(2,87) = 0.250, *p*= 0.780) language outcomes (Table 10). As alpha cortical oscillations did not show above chance cortical tracking, comparison to language measures were not considered (however the data are provided in Tables S14-S16 for completeness)

Delta cortical tracking

a, Infant-led language measures

Multivariate linear model	Df_{num}	Df_{denom}	*F*	*p*
Infant-led language measures	5	77	0.83	0.35
Univariate linear models (post hoc)	*Df*	$\beta_{estimate}$	*SE*	*p*
NWR Consonants, 24 months	1,96	1.65	1.85	0.37
NWR Syllables, 24 months	1,96	1.43	1.85	0.44
NWR Stress, 24 months	1,96	2.89	1.67	0.09
CCT, 18 months	1,93	1.04	1.46	0.48
Pointing, 12 months	1,105	2.91	2.95	0.33

b, Parent-estimated language measures

Multivariate linear model	Df_{num}	Df_{demom}	*F*	*p*
Parent-estimated language measures	2	87	3.00	**0.055**
Univariate linear models (post hoc)	*Df*	$\beta_{estimate}$	*SE*	*p*
CDI comprehension, 24 months	1,90	2043.00	877.43	***0.02***
CDI production, 24 months	1,90	2235.55	948.65	***0.02***

Table 9, Multivariate and univariate linear models, investigating whether delta cortical tracking at 11 months predicted either infant-led or parent-estimated language outcomes. The table details the multivariate linear models (reporting the global effects) and the post-hoc univariate linear models (ran after the main multivariate model) describing whether delta cortical tracking at 11 months predicted a) infant-led language measures or b) the parent-estimated language measures. Bonferroni correction for multiple comparisons in the **univariate models** led to modified significant alpha levels of, p = <0.01 for infant-led and p = <0.025 for parent-estimated (denoted by ***bold italic text***). Modified alpha level for trends are p = <0.02 for infant-led and p = <0.05 for parent-estimated (denoted by **bold text**).

Theta cortical tracking

a, Infant-led language measures

Multivariate linear model	*Df num*	*Df demom*	*F*	*p*
Infant-led language measures	5	77	0.89	0.50
Univariate linear models (post hoc)	*Df*	*β estimate*	*SE*	*p*
NWR Consonants, 24 months	1,96	-4.32	2.91	0.14
NWR Syllables, 24 months	1,96	-3.51	2.92	0.23
NWR Stress, 24 months	1,96	-2.24	2.67	0.40
CCT, 18 months	1,93	-2.72	2.28	0.24
Pointing, 12 months	1,105	5.72	4.61	0.22

b, Parent-estimated language measures

Multivariate linear model	*Df num*	*Df demom*	*F*	*p*
Parent-estimated language measures	2	87	0.25	0.78
Univariate linear models (post hoc)	*Df*	*β estimate*	*SE*	*p*
CDI comprehension, 24 months	1,90	-828.31	1452.9	0.57
CDI production, 24 months	1,90	-1114.12	1570.3	0.48

Table 10, Multivariate and univariate linear models, investigating whether theta cortical tracking at 11 months predicted either infant-led or parent-estimated language outcomes. The table details the multivariate linear models (reporting the global effects) and the post-hoc univariate linear models (ran after the main multivariate model) describing whether theta cortical tracking at 11 months predicted a) infant-led language measures or b) the parent-estimated language measures. Bonferroni correction for multiple comparisons in the **univariate models** led to modified significant alpha levels of, p = <0.01 for infant-led and p = <0.025 for parent-estimated (denoted by ***bold italic text***). Modified alpha level for trends are p = <0.02 for infant-led and p = <0.05 for parent-estimated (denoted by **bold text**).

3.2.4 PAC and language outcomes

Phase amplitude coupling (PAC) at 4-months was selected for the language outcome analyses as PAC did not show any developmental improvement, indicating that the early presence of stronger PAC may be informative. 87 infants provided complete data for all the parent-estimated measures and a PAC value, whilst 81 infants provided data for all the infant-led language measures and a PAC value. Multivariate linear models were run separately to

623 investigate whether delta-beta, delta-gamma, theta-beta or theta-gamma PAC respectively
624 predicted the language outcome measures.

625 The multivariate linear models showed that increased theta-gamma PAC predicted a
626 global increase in both the infant-led ($F(5,76) = 3.541$, $p = 0.006$) and the parent-estimated
627 ($F(2,85) = 4.86$, $p = 0.010$) language measures (Table 11). Posthoc Bonferroni-corrected
628 univariate linear models highlighted that better non-word repetition of syllables at 24-months
629 ($p = .006$) and CDI production at 24-months ($p = .03$) were the strongest contributing factors.
630 Increased theta-beta PAC also predicted a global increase in the parent-estimated ($F(2,85) =$
631 3.34, $p = 0.040$) language measures. Posthoc univariate linear models however indicated that no
632 individual parent-estimated language measure drove the overall effect (Table 11b). Increased
633 theta-beta PAC did not predict the infant-led language outcome measures in the multivariate
634 analysis ($F(5,76) = 0.480$, $p = 0.790$) nor in the univariate analyses (Table 11a).

635 Delta-beta ($F(5,76) = 1.602$, $p = 0.170$) and delta-gamma ($F(5,76) = 0.544$, $p = 0.742$)
636 PAC did not predict the infant-led language outcome measures (Table 11a). Furthermore, delta-
637 beta ($F(2,85) = 1.62$, $p = 0.204$) and delta-gamma ($F(2,85) = 1.54$, $p = 0.241$) PAC did not predict
638 the parent-estimated language outcome measures either (Table 11b).

639 In summary, theta-based PAC was associated with an overall increase in both the infant-
640 led and parent-estimated global language outcomes. Theta-gamma PAC robustly affected both
641 groups of language measures, while theta-beta PAC affected parent-estimated measures alone.
642 Delta-based PAC did not make significant contributions to later language outcomes. When
643 gamma was the high frequency amplitude, both infant-led and parent-estimated language
644 outcomes showed significant predictive relations.

PAC

a, Infant-led language

	Delta-beta		*Delta-gamma*		*Theta-beta*		*Theta-gamma*	
Multivariate linear model	*F*	*p*	*F*	*p*	*F*	*p*	*F*	*p*
Infant-led language measures	1.60	0.17	0.54	0.74	0.48	0.79	3.54	***0.006***
	Delta-beta		*Delta-gamma*		*Theta-beta*		*Theta-gamma*	
Univariate linear models	$\beta_{est.}$	*p*	$\beta_{est.}$	*p*	$\beta_{est.}$	*p*	$\beta_{est.}$	*p*
NWR Consonants, 24 months	-0.04	0.46	-0.02	0.70	0.01	0.78	0.03	0.21
NWR Syllables, 24 months	-0.01	0.89	-0.05	0.34	0.01	0.07	0.07	***0.006***
NWR Stress, 24 months	-0.03	0.64	-0.04	0.42	-0.02	0.55	0.02	0.37
CCT, 18 months	-0.09	0.03	0.01	0.83	0.04	0.13	0.03	0.14
Pointing, 12 months	-0.12	0.19	0.05	0.56	-0.02	0.74	0.03	0.45

b, Parent-estimated language

	Delta-beta		*Delta-gamma*		*Theta-beta*		*Theta-gamma*	
Multivariate linear model	*F*	*p*	*F*	*p*	*F*	*p*	*F*	*p*
Parent-estimated lang measures	1.62	0.20	1.54	0.24	3.34	***0.04***	4.86	***0.01***
	Delta-beta		*Delta-gamma*		*Theta-beta*		*Theta-gamma*	
Univariate linear models	$\beta_{est.}$	*p*	$\beta_{est.}$	*p*	$\beta_{est.}$	*p*	$\beta_{est.}$	*p*
CDI comprehension, 24 mo	-47.5	0.12	43.7	0.10	-4.64	0.78	8.33	0.56
CDI production, 24 mo	-58.2	0.08	44.6	0.12	21.2	0.23	33.4	**0.03**

Table 11, Multivariate and univariate linear models, investigating whether delta-beta, delta-gamma, theta-beta or theta-gamma PAC at 4 months predicted either infant-led or parent-estimated language outcomes. The table details the multivariate linear models (reporting the global effects) and the post-hoc univariate linear models (ran after the main multivariate model) describing whether PAC at 4 months predicted a) infant-led language measures or b) the parent-estimated language measures. Bonferroni correction for multiple comparisons in the **univariate models** led to modified significant alpha levels of, p = <0.01 for infant-led and p = <0.025 for parent-estimated (denoted by ***bold italic text***). Modified alpha level for trends are p = <0.02 for infant-led and p = <0.05 for parent-estimated (denoted by **bold text**).

4 Discussion

According to recent extensions of Temporal Sampling theory (4,28) cortical tracking of infant directed speech (IDS) by delta- and theta-band neural signals and delta-theta neural dynamics may be key factors in explaining differences in the developmental trajectories for language acquisition. In the Cambridge UK BabyRhythm infant sample assessed here (N=113),

659 individual differences in delta-band cortical tracking, theta power, theta/delta power ratio, and
660 theta-led PAC were all significant predictors of language outcomes in infants during the first
661 year of life. These infant data are consistent with neural oscillatory factors that have been found
662 to explain individual differences in language development in children, for example the fidelity
663 and synchronicity of delta-band encoding of speech envelope information for phonological
664 development (32,46). They are also consistent with a recent EEG modelling preprint based on
665 child recordings during passive speech listening, which showed that the theta/delta power ratio
666 was related to phonological difficulties in dyslexia and that atypical theta-gamma PAC was
667 related to impaired language outcomes in DLD (27). Further, these infant neural oscillatory
668 factors were broadly consistent in terms of replicating our previous findings with the first half of
669 the infant cohort (1). The presence and maturation of low-frequency (<12 Hz) cortical
670 oscillations (PSD), cortical speech tracking (mTRF) and phase amplitude coupling (PAC) in
671 response to sung IDS was largely similar across both halves of the longitudinal Cambridge UK
672 BabyRhythm cohort.

673 The replication of stronger delta band tracking than theta band tracking in the youngest
674 (4-month-old) infants suggests that delta band cortical tracking may be a key early building
675 block for language acquisition. Efficient delta-band cortical tracking would enable infants to lock
676 on to the information-rich strong (or stressed) syllables in the speech signal, acoustic landmarks
677 that occur on average every 500ms across languages (47) . This mechanism would enable the
678 infant brain to begin to represent the acoustic structure of human speech (48,49). Dynamic
679 interactions with theta-band cortical tracking may then support further linguistic learning (see
680 below). The importance of delta-band cortical tracking for language outcomes is consistent with
681 child MEG studies, which show that typically-developing children exhibit efficient delta-band

682 tracking of speech-in-noise, but not efficient theta-band tracking (50). The finding that delta
683 band cortical tracking predicted early language comprehension and production is also congruent
684 with research in the adult brain, where delta-band tracking is associated with key language-
685 related skills such as perceptual grouping (51,52), discourse-level parsing related to phrasing
686 (53,54) and comprehending speech-in-noise (55).

687 In many adult studies, the theta band is considered the key linguistic oscillator (24). In
688 our infant cohort, better theta-gamma PAC at 4-months was associated with better outcomes in
689 both the infant-led and parent-estimated global language outcome measures, and theta-beta PAC
690 at 4-months predicted the parent-estimated global measure. Posthoc univariate tests revealed that
691 nonword repetition, a phonological output measure (specifically syllable production at 24-
692 months) and CDI production (parental estimation of vocabulary output) were strong drivers of
693 these effects. By contrast, delta-based PAC (with beta or gamma) did not predict any of our
694 language outcome measures. However, not all the effects found for theta-band cortical tracking
695 were positive in nature. Theta-band PSD increased significantly during the first year of life, and
696 greater theta-band PSD was associated with poorer global language outcomes as measured by
697 parental estimates of comprehension and production at 24-months. A larger theta/delta PSD
698 power ratio was also associated with poorer language outcomes, as measured by both parent-
699 estimated and infant-led measures. This latter finding underscores the importance of considering
700 neural dynamics when assessing developmental trajectories. In the previous EEG modelling
701 work with 9-year-old children with and without dyslexia that investigated theta/delta ratios, a
702 dyslexia-specific pattern of delta and theta power dynamics during natural speech processing was
703 seen to become progressively more atypical the worse the child's phonological awareness, and
704 this was not related to age (27).

705 The infant literature is sparse regarding PAC dynamics. For adults, theta-gamma PAC
706 has been suggested as a neural mechanism for processing phonetic information in the speech
707 signal (24,56,57). When the same rhythmic speech stimuli used in the current study were
708 presented to adult participants, theta/gamma coupling was stronger than delta/gamma coupling
709 (17). For our infant participants, presented with the same nursery rhyme stimuli, theta and delta
710 were equally strong carriers of higher-frequency amplitudes. Nevertheless, only theta-based PAC
711 predicted language outcomes. Indeed, it is the phase of theta band rather than delta band
712 oscillations that adjust to speech rate variations for adults, coupling with gamma amplitudes (22).
713 Whilst both delta and theta phases couple with higher-frequency amplitudes in the infant brain, it
714 seems that theta-based PAC is more strongly predictive of early language outcomes. The post-
715 hoc univariate analyses suggested that measures involving speech production, namely nonword
716 repetition and productive vocabulary, were important regarding these relations. Whilst the adult
717 PAC literature shows theta-gamma coupling is directly linked to speech processing (18), delta-
718 beta PAC is particularly associated with measures of speech production via audio-motor
719 dynamics (23,58). As the infants in our cohort continue to develop and produce more speech, it
720 may be that delta-beta PAC will also begin to play a role in predicting language outcomes.

721 The current study has several limitations. Recording EEG with infants is an inherently
722 noisy process, and this affected our data, as different infants had to be excluded from different
723 analyses, reducing sample size for the longitudinal language outcome assessments. A key source
724 of noisy data is infant fussing and movement, and noise was particularly apparent in the 7-
725 month-old and silent state recordings for our sample. A second limitation is that even though our
726 study followed a relatively large cohort of infants (122 recruited and 113 retained), a study with
727 a larger participant pool may have had increased chances of uncovering developmentally

important time-lagged associations. Notwithstanding the limitations of the study, here we present a large longitudinal sample demonstrating that neural dynamics to speech stimuli predict early infant language acquisition.

5 Conclusion

Estimates of the efficiency of low-frequency cortical tracking, phase amplitude coupling and theta-delta power dynamics in an infant's first year of life predict their early language outcomes. Importantly, the neural mechanisms measured here are automatic processes that are inherent to the human brain, and as such represent physiological priors for language learning in a Bayesian sense (59). Accordingly, individual differences in the efficiency of these mechanisms in the pre-verbal stage of language learning assessed here are unlikely to reflect top-down linguistic processes such as language comprehension (12). Delta-band cortical tracking at 11 months and theta−gamma PAC at 4 months of age significantly predicted language comprehension and production at 24 months. Conversely, increased theta PSD power and a larger theta/delta ratio were associated with decreased performance on the language outcome measures. While other studies have also indicated a predictive relationship between cortical tracking efficiency in infancy and language development (26), the current study is the first to show that individual differences in PAC and theta/delta power dynamics also predict infant language outcomes.

A Declaration of Competing Interest

The authors do not report any conflicts of interest.

Author Contributions

Adam Attaheri: EEG Paradigm Development, EEG Preprocessing, Investigation - Data Curation, Formal Analysis - design, creation and implementation of analysis, Writing - Original Draft.

Áine Ní Choisdealbha: EEG Preprocessing - Investigation - Data Curation - Writing - Review & Editing

Sinead Rocha: Investigation - Data Curation, Writing - Review & Editing

Perrine Brusini: EEG Paradigm Development, Investigation - Data Curation, Writing - Review & Editing

Giovanni Di Liberto: Formal analysis, Writing - Review & Editing

Natasha Mead: Investigation - Data Curation

Helen Olawole-Scott: Investigation - Data Curation

Panagiotis Boutris: Investigation - Data Curation

Samuel Gibbon: Investigation - Data Curation

Isabel Williams: Investigation - Data Curation

Christina Grey: Investigation - Data Curation

Maria Alfaro e Oliveira - Data Curation

Carmel Brough - Data Curation

Shelia Flanagan – Data curation

Usha Goswami: Conceptualization - Methodology, Funding Acquisition, Supervision, Project Administration, Writing - Original Draft

Data Sharing agreement

The analyses were conducted by using publicly available MATLAB toolboxes that can be downloaded at http://audition.ens.fr/adc/NoiseTools/, https://sccn.ucsd.edu/eeglab/download.php, https://github.com/mickcrosse/mTRF-Toolbox.

The custom code used for the mTRF, PSD, PAC and language analysis can be downloaded from https://osf.io/ea9bd/. The final data included in the manuscript figures and statistics can also be downloaded from https://osf.io/ea9bd/.

This data is currently being analysed with other techniques as part of the ongoing longitudinal Cambridge UK BabyRhythm project. Once accepted for publication, the raw EEG files will be uploaded to the same OSF server as above.

Funding sources

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 694786).

Acknowledgements

We would like to thank all the families and infants who kindly donated their time to this project. We would also like to thank Henna Ahmed for their time spent as an RA on the Cambridge UK

BabyRhythm project. Finally, we would like to acknowledge Mike X Cohen's statistical textbooks and online materials for guiding our statistical analysis.

References

1. Attaheri A, Choisdealbha ÁN, Di Liberto GM, Rocha S, Brusini P, Mead N, et al. Delta- and theta-band cortical tracking and phase-amplitude coupling to sung speech by infants. NeuroImage. 2022 Feb 15;118698.

2. Kuhl PK. Early language acquisition: cracking the speech code. Nat Rev Neurosci. 2004 Nov;5(11):831–43.

3. Rocha S, Ní Choisdealbha Á, Attaheri A, Mead N, Olawole-Scott H, Grey C, et al. Language acquisition in the longitudinal BabyRhythm cohort. PsyArXiv; Preprint.

4. Goswami U. Speech rhythm and language acquisition: an amplitude modulation phase hierarchy perspective. Ann N Y Acad Sci. 2019;1453(1):67–78.

5. Goswami U. A Neural Basis for Phonological Awareness? An Oscillatory Temporal-Sampling Perspective. Curr Dir Psychol Sci. 2018 Feb 1;27(1):56–63.

6. Baltzell LS, Srinivasan R, Richards VM. The effect of prior knowledge and intelligibility on the cortical entrainment response to speech. J Neurophysiol. 2017 Sep 6;118(6):3144–51.

7. Di Liberto GM, Crosse MJ, Lalor EC. Cortical Measures of Phoneme-Level Speech Encoding Correlate with the Perceived Clarity of Natural Speech. eNeuro [Internet]. 2018 Apr 16 [cited 2020 Jul 3];5(2). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5900464/

8. Di Liberto GM, O'Sullivan JA, Lalor EC. Low-Frequency Cortical Entrainment to Speech Reflects Phoneme-Level Processing. Curr Biol CB. 2015 Oct 5;25(19):2457–65.

9. Ding N, Simon JZ. Cortical entrainment to continuous speech: functional roles and interpretations. Front Hum Neurosci [Internet]. 2014 [cited 2020 Jul 3];8. Available from: https://www.frontiersin.org/articles/10.3389/fnhum.2014.00311/full

10. Luo H, Poeppel D. Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. Neuron. 2007 Jun 21;54(6):1001–10.

11. Millman RE, Johnson SR, Prendergast G. The Role of Phase-locking to the Temporal Envelope of Speech in Auditory Perception and Speech Intelligibility. J Cogn Neurosci. 2015 Mar 1;27(3):533–45.

12. Peelle JE, Gross J, Davis MH. Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. Cereb Cortex N Y N 1991. 2013 Jun;23(6):1378–87.

13. Jessen S, Fiedler L, Münte TF, Obleser J. Quantifying the individual auditory and visual brain response in 7- month-old infants watching a brief cartoon movie. bioRxiv. 2019 Apr 17;610709.

14. Kalashnikova M, Peter V, Di Liberto GM, Lalor EC, Burnham D. Infant-directed speech facilitates seven-month-old infants' cortical tracking of speech. Sci Rep [Internet]. 2018 Sep 13 [cited 2020 Jul 3];8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6137049/

15. Ortiz Barajas MC, Guevara R, Gervain J. The origins and development of speech envelope tracking during the first months of life. Dev Cogn Neurosci. 2021 Jan 20;48:100915–100915.

16. Molinaro N, Lizarazu M. Delta(but not theta)-band cortical entrainment involves speech-specific processing. Eur J Neurosci. 2018 Oct 1;48(7):2642–50.

17. Attaheri A, Panayiotou D, Phillips A, Ní Choisdealbha Á, Di Liberto GM, Rocha S, et al. Cortical Tracking of Sung Speech in Adults vs Infants: A Developmental Analysis. Front Neurosci [Internet]. 2022;16. Available from: https://www.frontiersin.org/article/10.3389/fnins.2022.842447

18. Gross J, Hoogenboom N, Thut G, Schyns P, Panzeri S, Belin P, et al. Speech Rhythms and Multiplexed Oscillatory Sensory Coding in the Human Brain. PLOS Biol. 2013 Dec 31;11(12):e1001752.

19. Canolty RT, Edwards E, Dalal SS, Soltani M, Nagarajan SS, Kirsch HE, et al. High Gamma Power Is Phase-Locked to Theta Oscillations in Human Neocortex. Science. 2006 Sep 15;313(5793):1626–8.

20. Canolty RT, Knight RT. The functional role of cross-frequency coupling. Trends Cogn Sci. 2010 Nov;14(11):506–15.

21. Lakatos P, Shah AS, Knuth KH, Ulbert I, Karmos G, Schroeder CE. An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. J Neurophysiol. 2005 Sep;94(3):1904–11.

22. Lizarazu M, Lallier M, Molinaro N. Phase−amplitude coupling between theta and gamma oscillations adapts to speech rate. Ann N Y Acad Sci. 2019 Oct;1453(1):140–52.

23. Arnal LH, Doelling KB, Poeppel D. Delta-Beta Coupled Oscillations Underlie Temporal Prediction Accuracy. Cereb Cortex N Y N 1991. 2014/05/20 ed. 2015 Sep;25(9):3077–85.

24. Giraud AL, Poeppel D. Cortical oscillations and speech processing: emerging computational principles and operations. Nat Neurosci. 2012 Mar 18;15(4):511–7.

25. Hyafil A, Fontolan L, Kabdebon C, Gutkin B, Giraud AL. Speech encoding by coupled cortical theta and gamma oscillations. Brownell H, editor. eLife. 2015 May 29;4:e06213.

26. Menn KH, Ward EK, Braukmann R, van den Boomen C, Buitelaar J, Hunnius S, et al. Neural Tracking in Infancy Predicts Language Development in Children With and Without Family History of Autism. Neurobiol Lang. 2022 Aug 17;3(3):495–514.

27. Araújo J, Simons BD, Varghese P, Mandke K, Kalashnikova M, Macfarlane A, et al. Atypical encoding of speech identifies children with Dyslexia versus Developmental Language Disorder. BioRxiv; Preprint.

28. Goswami U. Language acquisition and speech rhythm patterns: an auditory neuroscience perspective. R Soc Open Sci. 2022 Jul 27;9.

29. Goswami U. A temporal sampling framework for developmental dyslexia. Trends Cogn Sci. 2011 Jan 1;15(1):3–10.

30. Keshavarzi M, Di Liberto G, Gabrielczyk F, Wilson A, Macfarlane A, Goswami U. Atypical speech production of multisyllabic words by children with developmental dyslexia. 2022.

31. Mandke K, Flanagan S, Macfarlane A, Gabrielczyk F, Wilson A, Gross J, et al. Neural sampling of the speech signal at different timescales by children with dyslexia. NeuroImage. 2022 Jun 1;253:119077.

32. Power AJ, Colling LJ, Mead N, Barnes L, Goswami U. Neural encoding of the speech envelope by children with developmental dyslexia. Brain Lang. 2016 Sep;160:1–10.

33. Power AJ, Mead N, Barnes L, Goswami U. Neural entrainment to rhythmic speech in children with developmental dyslexia. Front Hum Neurosci [Internet]. 2013 [cited 2020 Jul 3];7. Available from: https://www.frontiersin.org/articles/10.3389/fnhum.2013.00777/full

34. Kalashnikova M, Goswami U, Burnham D. Sensitivity to amplitude envelope rise time in infancy and vocabulary development at 3 years: A significant relationship. Dev Sci. 2019 Nov 1;22(6):e12836.

35. Kalashnikova M, Goswami U, Burnham D. Delayed development of phonological constancy in toddlers at family risk for dyslexia. Infant Behav Dev. 2019 Nov 1;57:101327.

36. Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods. 2004 Mar 15;134(1):9–21.

37. Kothe CA, Makeig S. BCILAB: a platform for brain–computer interface development. J Neural Eng. 2013 Aug 28;10(5):056014.

38. Crosse MJ, Di Liberto GM, Bednar A, Lalor EC. The Multivariate Temporal Response Function (mTRF) Toolbox: A MATLAB Toolbox for Relating Neural Signals to Continuous Stimuli. Front Hum Neurosci [Internet]. 2016 [cited 2020 Jul 3];10. Available from: https://www.frontiersin.org/articles/10.3389/fnhum.2016.00604/full

39. Pasley BN, David SV, Mesgarani N, Flinker A, Shamma SA, Crone NE, et al. Reconstructing Speech from Human Auditory Cortex. PLOS Biol. 2012 Jan 31;10(1):e1001251.

40. Zion Golumbic EM, Ding N, Bickel S, Lakatos P, Schevon CA, McKhann GM, et al. Mechanisms Underlying Selective Neuronal Tracking of Attended Speech at a 'Cocktail Party.' Neuron. 2013 Mar 6;77(5):980–91.

41. O'Sullivan JA, Power AJ, Mesgarani N, Rajaram S, Foxe JJ, Shinn-Cunningham BG, et al. Attentional Selection in a Cocktail Party Environment Can Be Decoded from Single-Trial EEG. Cereb Cortex N Y NY. 2015 Jul;25(7):1697–706.

42. Özkurt TE, Schnitzler A. A critical note on the definition of phase–amplitude cross-frequency coupling. J Neurosci Methods. 2011 Oct 15;201(2):438–43.

43. Alcock K, Meints K, Rowland CF, Brelsford V, Christopher A, Just J. The UK Communicative Development Inventories: Words and Gestures - Manual and Norms. Emsworth, UK: J & R Press; 2020.

44. Liszkowski U, Carpenter M, Tomasello M. Reference and attitude in infant pointing. J Child Lang. 2007/01/25 ed. 2007;34(1):1–20.

45. Friend M, Keplinger M. An Infant-Based Assessment of Early Lexicon Acquisition. Behav Res Methods Instrum Comput J Psychon Soc Inc. 2003 Jun 1;35:302–9.

46. Molinaro N, Lizarazu M, Lallier M, Bourguignon M, Carreiras M. Out-of-synchrony speech entrainment in developmental dyslexia. Hum Brain Mapp. 2016;37(8):2767–83.

47. Dauer RM. Stress-timing and syllable-timing reanalyzed**A preliminary version of this paper was read at the Annual Meeting of the Linguistic Society of America, San Antonio, Texas, December 28–30, 1980. The experimental work for this study was carried out at the Phonetics Laboratory, Edinburgh University. J Phon. 1983 Jan 1;11(1):51–62.

48. Leong V, Kalashnikova M, Burnham D, Goswami U. The Temporal Modulation Structure of Infant-Directed Speech. 2017 Mar 27 [cited 2020 Jul 3]; Available from: https://www.repository.cam.ac.uk/handle/1810/263724

49. Leong V, Goswami U. Acoustic-Emergent Phonology in the Amplitude Envelope of Child-Directed Speech. PLOS ONE. 2015 Dec 7;10(12):e0144411.

50. Ghinst MV, Bourguignon M, Niesen M, Wens V, Hassid S, Choufani G, et al. Cortical Tracking of Speech-in-Noise Develops from Childhood to Adulthood. J Neurosci. 2019 Apr 10;39(15):2938–50.

51. Ding N, Melloni L, Zhang H, Tian X, Poeppel D. Cortical tracking of hierarchical linguistic structures in connected speech. Nat Neurosci. 2016 Jan;19(1):158–64.

52. Kösem A, van Wassenhove V. Distinct contributions of low- and high-frequency neural oscillations to speech comprehension. Lang Cogn Neurosci. 2017 May 28;32(5):536–44.

53. Boucher VJ, Gilbert AC, Jemel B. The role of low-frequency neural oscillations in speech processing: Revisiting delta entrainment. J Cogn Neurosci. 2019;31(8):1205–15.

54. Doelling KB, Arnal LH, Ghitza O, Poeppel D. Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. NeuroImage. 2014 Jan 15;85 Pt 2:761–8.

55. Vanthornhout J, Decruy L, Wouters J, Simon JZ, Francart T. Speech Intelligibility Predicted from Neural Entrainment of the Speech Envelope. J Assoc Res Otolaryngol. 2018 Apr 1;19(2):181–91.

56. Fries P. The Model- and the Data-Gamma. Neuron. 2009 Dec 10;64(5):601–2.

57. Schroeder CE, Lakatos P. Low-frequency neuronal oscillations as instruments of sensory selection. Trends Neurosci. 2009 Jan 1;32(1):9–18.

58. Poeppel D, Assaneo MF. Speech rhythms and their neural foundations. Nat Rev Neurosci. 2020 Jun 1;21(6):322–34.

59. Doelling KB, Arnal LH, Assaneo MF. Adaptive oscillators provide a hard-coded Bayesian mechanism for rhythmic inference. bioRxiv. 2022 Jan 1;2022.06.18.496664.

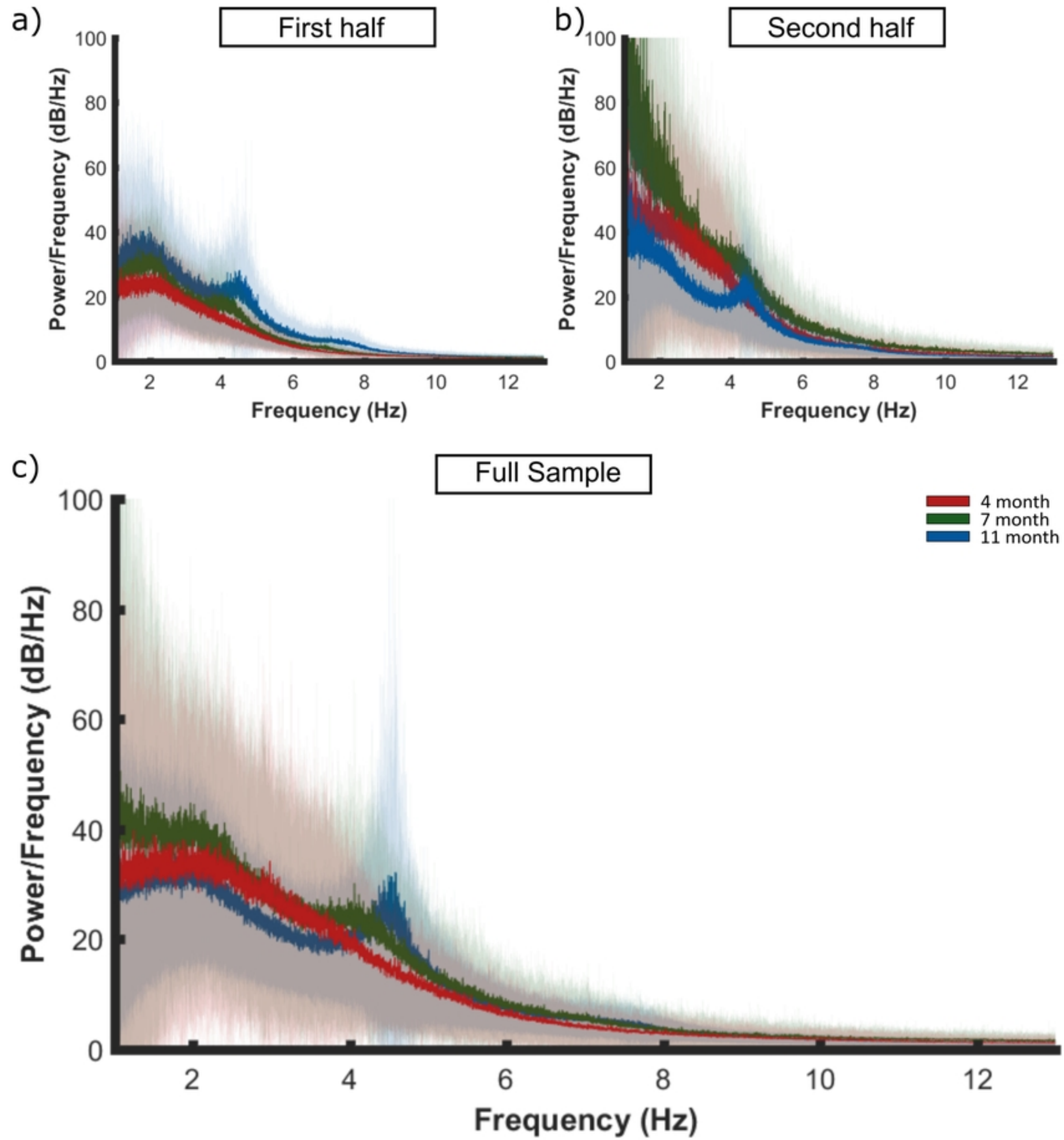
a)
First half
b)
Second half
c)
Full Sample
Power/Frequency (dB/Hz)
Frequency (Hz)
4 month
7 month
11 month

Figure 1

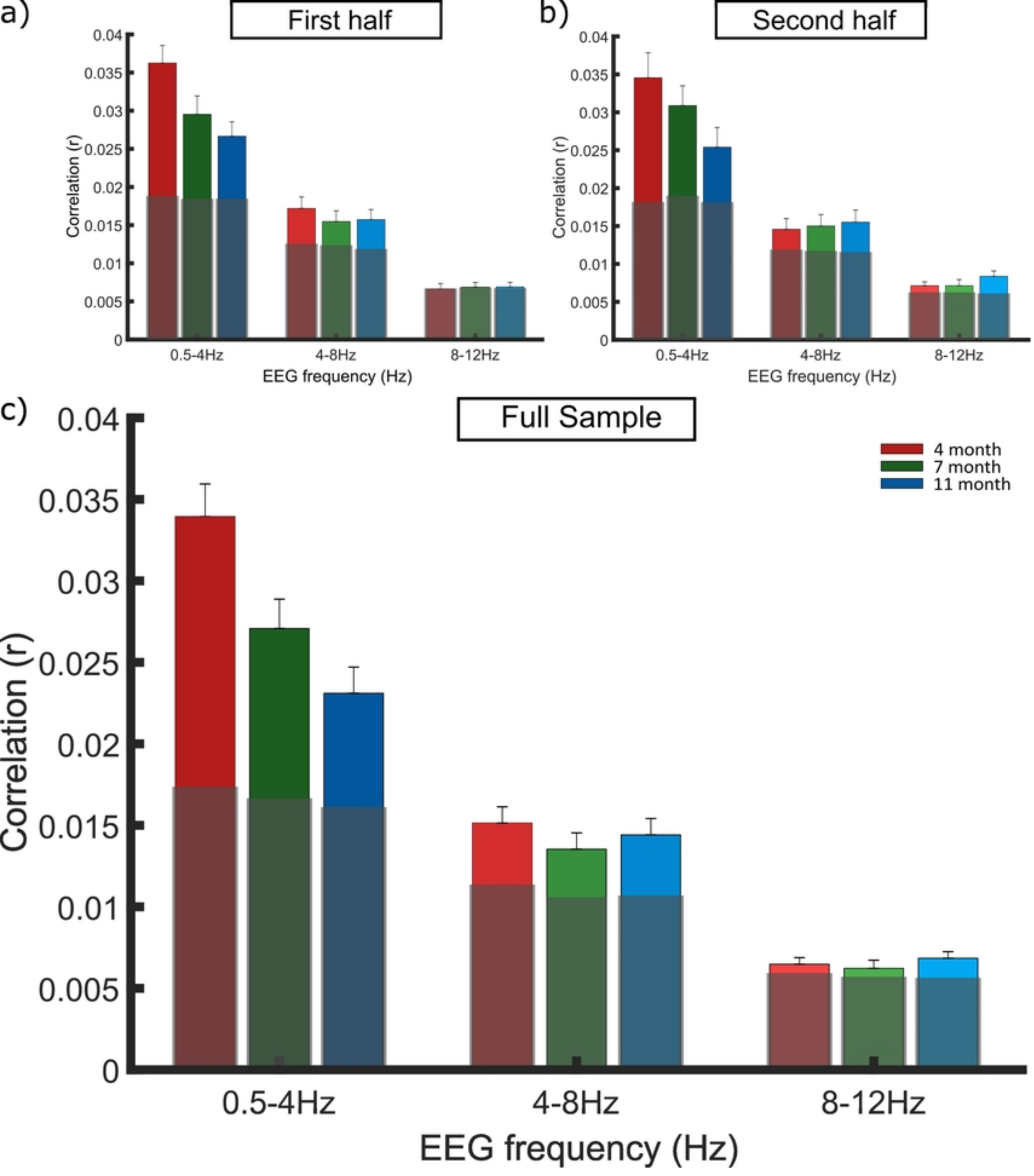
a)
First half
b)
Second half
c)
Full Sample
Correlation (r)
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
0.5-4Hz
4-8Hz
8-12Hz
EEG frequency (Hz)
4 month
7 month
11 month

Figure 2

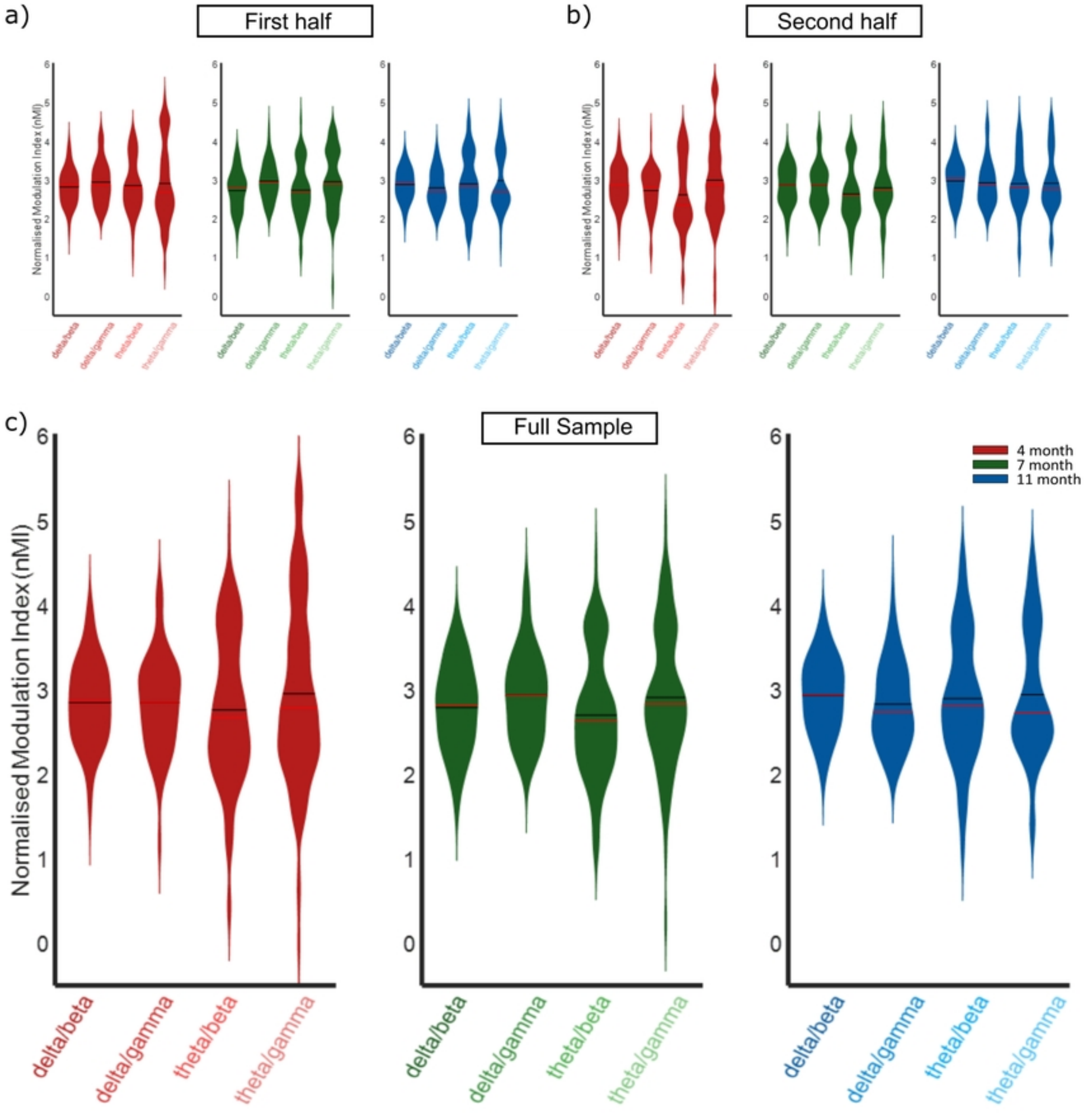
a)
First half
b)
Second half
Normalised Modulation Index (nMI)
delta/beta
delta/gamma
theta/beta
theta/gamma
c)
Full Sample
4 month
7 month
11 month

Figure 3

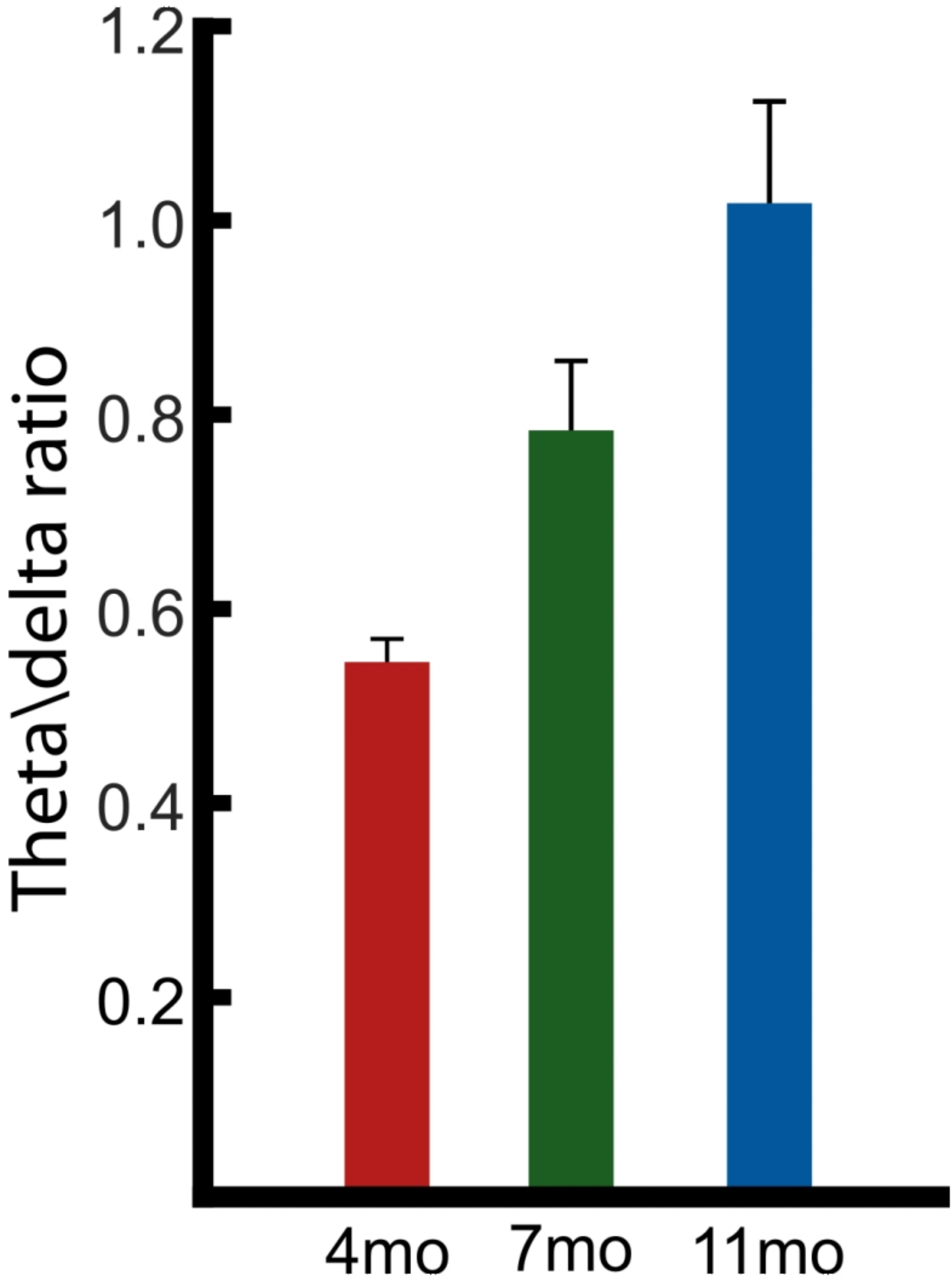

Figure 4