Mechanisms, Capabilities and Applications of Infant Neural and Physiological Oscillatory Entrainment

James Ernest White

A thesis submitted in partial fulfilment of the requirement of the University of East London for the degree of Doctor of Philosophy

September 2023

ABSTRACT

Infants entrain to a wide variety of simple rhythmic stimuli from birth and more complex stimuli during development (section 2.4). However, there is still limited understanding of infant neural and physiological entrainment mechanisms and how these compare with adult entrainment techniques. For example, there is growing evidence of internally generated and driven neural oscillatory entrainment in adults, but similar evidence is lacking in infant populations. There is also limited research that has investigated the effects of rhythm in naturalistic settings, such as the impact of rhythm in speech brain tracking, and disruption of infant rhythms caused by other external rhythmic stimuli.

This thesis assesses sub-second cognitive and physiological entrainment responses to provide new evidence in the above areas. In Chapter 3, neural entrainment data are presented discussing infant and adult entrainment preferences across stimulus frequency, along with spatial and temporal differences. In Chapter 4, neural entrainment to controlled temporally regular vs irregular stimuli and implications for mechanisms of infant and adult entrainment are explored. In Chapter 5, strength of speech-brain tracking to infant directed speech and song rhythms using continuous neural mapping investigated impacts of rhythm in infant directed language. Finally in Chapter 6, interactions between competing rhythmic infant crying and caregiver rocking in a home setting are investigated.

In brief, the results demonstrated for the first time that infant and adult entrainment preferences are different across stimulation frequencies; mechanisms of infant and adult entrainment are not the same; entrainment to infant directed speech vs song were differentiated across frequency bands and age ranges; and dynamic responses to infant crying can disturb infant rhythms to shorten cry length.

Discussion focusses on contribution to our understanding of infant entrainment capabilities and mechanisms, application of infant rhythmic entrainment to language acquisition, and dysregulation of infant rhythms with competing rhythms. Future research directions, methodological considerations and the need for more naturalistic studies are discussed.

Word count: 311

DECLARATION

This thesis focusses on data collected in one large study, a European Research Council (ERC) Horizon 2020 grant (award number 853251; PI Prof Sam Wass). The ethics approval letter and ethics application form for the project are in Appendix A.

Data collection was primarily performed by myself and seven other PhD students working on the project: Marta Perapoch Amadó, Emily Greenwood, Narain Viswanathan, Pierre Labendzki, Tom Northrop. Gio Esposito, and lab manager Katie Lancaster. The specific paradigms that are presented in this thesis were primarily designed by myself. All seven PhD students contributed to the design of the paradigms used for data collection, as well as data management and participant recruitment. I was responsible for writing the ethics application for this study.

All formulation and testing of hypotheses, data preprocessing, data analysis and writing of the text of this thesis are entirely my own work.

Open science statement

Public accessibility of data, protocols, code, and other materials utilized in the empirical chapters, as well as within the broader ONACSA project, is crucial. Upon completion and acceptance of publications, all relevant data, materials, stimuli, and custom code will be made publicly available. All data with personally identifiable information first be anonymised, and where this is not possible the data will be shared by request only in line with the ONACSA data management plan. An Open Science Framework project has been established at https://osf.io/fdsx8/, awaiting public release of data.

In Chapters 3 and 4, detailed procedures for stimuli creation, data recording, and analysis pipelines are provided (sections 3.2.2 to 3.2.6 and 4.2.2 to 4.2.6). Post-publication, sound stimuli and raw/preprocessed EEG data will be released publicly, along with code for stimuli production and analysis.

Chapter 5 outlines procedures for experimental setup, data recording, and analysis (sections 5.2.2 to 5.2.7). Audio data containing personally identifiable information will be available upon request with appropriate sharing agreements, while raw/preprocessed EEG data and analysis code will be released publicly.

Chapter 6 details procedures for experimental setup, wearable device specifications, data recording, and analysis (Sections 6.2.2 to 6.2.5). Audio and GPS data containing personally identifiable information will be shared upon request with appropriate sharing agreements, while proximity and actigraphy data will be made publicly available post-publication, alongside analysis code.

First-author publications arising from this thesis

Please note that my married name is James Ives, but for visa reasons related to my wife's nationality, visas and citizenship, I have not legally changed my name from James White to James Ives. Publications have been put forward with my married name.

Chapter 3 represents the work of the following publication under review at the European Journal of Neuroscience:

Ives, J., Labendzki, P., Perapoch Amadó, Greenwood, E., Viswanathan, N., Northrop, T., & Wass, S. (submitted). At Which Low Amplitude Modulated Frequency Do Infants Best Entrain? A Frequency Tagging Study, *bioRxiv 2022.12.08.519576;* doi: https://doi.org/10.1101/2022.12.08.519576

During my PhD I also contributed to the following publications

Wass, S. V., Perapoch Amadó, M., & **Ives, J**. (2022). Oscillatory entrainment to our early social or physical environment and the emergence of volitional control. Developmental Cognitive Neuroscience, 54. https://doi.org/10.1016/j.dcn.2022.101102

Perapoch Amado, M., Greenwood, E., **Ives, J**., Labendzki, P., Marriott-Haresign, I, Northrop, T., Phillips, E., Viswanathan, N., Whitehorn, M., Jones, E., & Wass, S. (preprint). The neural and physiological substrates of real-world attention change across development. 10.31219/osf.io/6z27a.

Perapoch Amado, M., Esposito, G., Greenwood, E., **Ives, J**., Labendzki, P., Lancaster, K., Northrop, T., Vaaras, E., Viswanathan, N., Tamarin, M., Crawley, M., Jones, E., & Wass, S. (preprint). Do shorter infant distress episodes with age result from improvements in caregiver-infant co- regulation or infant self-regulation? DOI: 10.13140/RG.2.2.10414.28484

TABLE OF CONTENTS

ABBREVIATIONS

AM	Amplitude Modulated
ANOVA	Analysis of Variance
ANOVAN	N-Ways ANOVA
ASSR	Auditory Steady State Responses
df	Degrees of Freedom
DLD	Developmental Language Disorder
ECG	Electrocardiogram
EEG	Electroencephalography
ERP	Event Related Potentials
FFT	Fast Fourier Transform
FM	Frequency Modulated
Hz	Hertz
ICA	Independent Components Analysis
ID	Infant Directed
ISI	Inter Stimulus Interval
ITC	Inter Trial Coherence
LMEM	Linear Mixed Effects Models
MMN	Mismatched Negativity
mTRF	Multivariate Temporal Response Function
PLV	Phase Locking Value
SBT	Speech Brain Tracking
SM	Supplementary Materials
SNR	Signal To Noise Ratio
SS-EPs	Steady State Evoked Potentials
SSVEP	Steady State Visual Evoked Potentials
WD	Wearable Device

TABLE OF FIGURES

Main text figures

Appendix figures

TABLE OF TABLES

Main text tables

Appendix tables

ACKNOWLEDGEMENTS

I'd like to thank a number of people for their role in supporting this piece of work and me throughout my PhD. First, I would like to thank Prof Sam Wass for giving me the opportunity, for supporting me academically, for his boundless enthusiasm and drive for developmental research, and for inviting me to be part of the ONACSA project and Baby Development Lab in UEL. The ONACSA project fascinated me from first reading, and I feel privileged to have had the chance to work on it. I hope to continue to work with the UEL Baby Dev Lab for as long as they'll have me.

Second, I'd like to thank the PhD students and staff that have worked on the ONACSA project: Marta Perapoch Amadó, Emily Greenwood, Narain Viswanathan, Pierre Labendzki, Tom Northrop, Gio Esposito and Katie Lancaster, who have been brilliant colleagues, sharing the highs and lows of testing very young infants, and have always been willing to listen and provide insightful ideas and feedback. More broadly I'd like to thank all members of the UEL Baby Dev Lab and the Psychology Office Technicians. The vibrant lab is filled with such knowledgeable, skilled, and expert people from all of whom I have learnt so much.

On a personal note, I would like to express my deepest gratitude to my family and my partner for their unwavering support and love throughout my PhD journey. Your encouragement and belief in me have been the cornerstone of my success, and I could not have accomplished this without you by my side.

To my family, thank you for always being there, for providing the foundation from which I could flourish and for your unflinching support at every stage of my life, to my mum Karen and her partner Paul, my sister Anni and my scientist-in-training nephew Levi, and my grandparents Shirley and Jim.

I’d also like to thank the incredible and emotionally overwhelming generosity of all those who got me back on my feet after the robbery, especially my parents-in-law Glenda and Terry, who were so kind to buy me a new laptop, without which I wouldn’t have been able to finish my analysis and write-up.

Finally, I’d like to thank my wife Allie who has been there for me at every step of this PhD as my loudest and most enthusiastic cheerleader. Thank you for all the brilliant support, your endless patience and understanding during the countless late nights and weekends I spent working on my research. Your sacrifices did not go unnoticed, and I am forever grateful for your constant encouragement and pride in my achievements.

Chapter 1 – Thesis Overview

This thesis aims to examine neural and physiological capabilities and mechanisms that drive infant entrainment to external rhythmic stimuli. This is particularly important to understanding the development and application of infant entrainment to rhythmic social structures such as language. This thesis incorporates 4 separate empirical analyses conducted on data collected either in a lab setting with EEG recording neural responses to controlled rhythmic audio across a range of frequencies and jittered rhythms; EEG recorded in an interactive context during free-flowing play or singing sessions between caregivers and infants; or finally in a home setting with custom made wearable devices to record infant and adult audio, physiology, and behaviour. Beginning with a focus on infant and adult capabilities to entrain to different rhythmic stimuli in Chapters 3 and 4, Chapter 5 goes on to examine speech-brain tracking to infant directed speech and song. Chapter 6 examines natural crying and rocking rhythms in a home environment and explores a rare setting where caregivers attempt to disrupt infant rhythms with their own competing rhythm.

In Chapter 2 the thesis starts by outlining why rhythms are important in the development of infant entrainment, followed by an overview of current evidence, benefits and capabilities of infant rhythmic entrainment throughout infant development. Next, a review of research that has studied infant entrainment to anisochronous, live and naturalistic stimuli is considered as well as speech brain tracking and potential avenues for research examining dysregulation of infant rhythmic entrainment. The Chapter ends with a discussion on the mechanisms of infant entrainment, how they might contrast with adult entrainment and future directions.

Chapters 3 and 4 investigate infant and adult neural entrainment using a frequency tagging paradigm to isochronous rhythmic stimuli at a range of frequencies (Chapter 3) and a group of jittered rhythms, defined as those with a consciously perceived isochronous rhythm but with individual beats presented ahead or behind that rhythm (Chapter 4). Chapter 3 examines whether infant and adult participants entrain more strongly to certain frequencies when all else is held constant, differences in power vs phase domains, spatial regions of interest and stimulus durations needed to engage rhythmic entrainment. Chapter 4 builds on work from Chapter 3 to examine how infant and adult entrainment is affected by temporal irregularity and offers a robust method to test entrainment mechanisms. Findings are discussed relative to previous mechanistic and developmental entrainment research including methodological considerations for future studies.

Chapter 5 expands on these findings by investigating speech-brain tracking in naturalistic infant directed speech and song, leveraging the similarities between these infant directed vocalisations and differences in tempo across infants at 5, 10 and 15 months old. Analyses are run to assess the amount of rhythmic vocalisation and dominant frequencies of rhythms embedded within the infant directed language conditions and how this correlates with speech-brain tracking. Implications of using naturalistic rather than recorded or manipulated audio are discussed.

Further natural rhythmic structures and dysregulation of infant rhythms are investigated in Chapter 6, which used day-long home recordings with custom made wearable devices that measure infant and caregiver continuous audio, GPS, proximity, actigraphy and heart rate. This offers a detailed look into the sub-second interplay of rhythmic infant negative arousal episodes and rhythmic caregiver soothing techniques. Findings are discussed relative to the lack of naturalistic studies that have investigated crying and rocking rhythms and in terms of how dysregulation of infant crying rhythms as measured by a decrease in crying duration is impacted by a variety of naturally occurring rhythmic rocking techniques.

In the General Discussion of Chapter 7, this thesis reviews the implications of the findings in Chapters 3-6 on our understanding of infant entrainment to rhythmic stimuli across a range of controlled and naturalistic settings. The thesis ends with a discussion on the development of further stimuli, methods and analyses to examine capabilities, mechanisms, applications and development of infant rhythmic entrainment and how this might inform not only infant studies but also adult entrainment studies.

Chapter 2 – General Introduction

2.1 Introduction

Rhythmic stimuli are increasingly present in modern society with rising technology use but are also shown in naturalistic systems such as heart rhythms, weather patterns, day-night cycles etc (Koukkari & Sothern, 2007). A rhythm is the repeated oscillation of strong and weak activity of any continuous stimulus. Often rhythms are thought of as isochronous, i.e. with equal temporal spacing between peaks or troughs in the activity fluctuation, but rhythms come in a variety of types including: anisochronous rhythms, those that are not equally temporally spaced; syncopated rhythms, with missing periods of activity; or complex rhythms that may be repeating patterns of tones or the convolution of multiple rhythms laid on top of one another (Fraisse, 1982).

An oscillator in the context of this thesis represents the signal of a specific rhythm within a system. Entrainment represents the similarity of the change in activity over time (Hoehl, Fairhurst & Schirmer, 2021; Wass, Perapoch Amadó & Ives, 2022), and in this thesis will largely be referred to in the context of the change in activity between two oscillators. This change over time may concurrently align the two oscillators or put the two oscillators in a lagged relationship (for review see Wass et al., 2020). The important characteristic is that this remains a stable relationship over time, i.e., a transient aligning of two oscillators with subsequent decaying alignment would not be considered entrainment. Similarly, the coincidental alignment of two oscillators independent of one another can also not be considered entrainment. For example, the swinging of two pendula on opposite sides of the earth could hold the same oscillation frequency and be perfectly aligned, but this would not be because of entrainment. Conversely, two pendula linked with a spring will entrain to be concurrently in phase as a result of energy transfer via the spring.

Concurrent entrainment of two oscillators, also known as “in phase” entrainment, represents activity when one oscillator peaks the other also peaks, when each troughs the other also troughs. Sequential entrainment occurs when one oscillator follows and overlaps the other, also knowns as “out of phase” entrainment, here when one oscillator peaks the other will peak before the first troughs. Oscillators that are entrained in an opposing manner are a special form of sequential entrainment, also known as “anti-phase” entrainment, which occurs when one oscillator peaks, the other troughs and the two oscillators alternate between the two states. These types of entrainment between two oscillators can be thought of either as driven by one oscillatory generator and the other following, or by the mutual changing of oscillatory rhythms to each attempt to synchronise with the other (Wass et al., 2020; Wass, Perapoch Amadó & Ives, 2022).

In this thesis, infant entrainment refers to the similarity of oscillatory change between internally generated behavioural, physiological or neural oscillators and an external oscillator. The external oscillator can be generated in the environment in a directly perceivable way, for example a repeated flashing light, or could be indirectly perceivable, for example neural entrainment between two participants driven by neural phase resetting as a result of social behaviours (for review see Wass, Perapoch Amadó & Ives, 2022).

Infants have been shown to entrain to a range of stimuli behaviourally, physiologically and neurally (for review see Haegens & Golumbic, 2018; Hoehl, Fairhurst & Schirmer, 2021; Wass, Perapoch Amadó and Ives, 2022). However, entrainment research has largely focused on adult studies, which have predominantly driven the discussion around the underlying mechanisms and cognitive benefits of neural entrainment (for a review see Zoefel, ten Oever & Sack, 2018). We know for example that adults can use top-down attentional processes to promote one stream of information while inhibiting others, often known as the cocktail effect (for recent review see Bronkhorst, 2018; see also Golumbic et al., 2013).

Theories surrounding adult attentional entrainment have discussed low-frequency neural oscillations as dynamic support for attention through either Dynamic Attending Theory (DAT; e.g. Jones, 1976), or Active Sensing (e.g. Lakatos et al., 2005; Schroeder & Lakatos, 2009) as an attention based framework for entrainment. DAT posits that attention can be drawn on cyclically into peaks and troughs of attention synchronised to an external oscillator through an endogenously driven attentional energy oscillator, which is self-sustaining and can be coupled with external sensory stimuli. While active sensing on the other hand, suggests that slow neural oscillations can support attentional selection of sensory information. In this view, in phase alignment of endogenous oscillators with external stimuli facilitates stronger processing of the external stimulus, while out of phase alignment leads to a dulling inhibitory response.

While attention to external oscillatory stimuli has been shown to increase the neural response to that stimulus, attention to the stimulus is not required to generate neural entrainment (e.g. Robertson et al., 2012). In this thesis, entrainment will be largely discussed in relation to the view of active sensing rather than DAT. This is primarily because the active sensing approach takes fundamentalist approach to sensory entrainment, suggesting that the alignment of internal neural oscillations impacts responses rather than attentional energy. As much of the work that follows occurs in a passive manner, either while infants are asleep, or as one feature in a wider more naturalistic context.

Despite works investigating many aspects of adult entrainment, we currently know very little even about the limits of infant entrainment capabilities throughout development, the mechanisms of infant entrainment, whether they differ from adult entrainment, and the effects of intentional disruption of infant entrainment (e.g. Bánki et al., 2022; Frischen, Degé & Schwarzer, 2022; Hoehl, Fairhurst & Schirmer, 2021).

In the context of infant entrainment to external rhythms, the sections below explore the current research on infant entrainment capabilities, mechanisms of entrainment, naturalistic infant speech-brain tracking and research into dysregulation of infant rhythmic entrainment. First, components of rhythms are discussed and how they may impact entrainment. Followed second by the current evidence for benefits of infant entrainment. Third, infant rhythmic entrainment capabilities to isochronous and anisochronous stimuli are explored in connection with naturalistic environments. Fourth, speech brain tracking and language acquisition are used as a case study for the application of infant entrainment in their typical development, before discussion of the limited evidence of dysregulation of infant rhythmic entrainment and the case for further study are presented and discussed. Finally, mechanisms of entrainment, supported by research in adult participants is examined and how methods can be adapted further for infant study.

2.2 Why are rhythms important in the development of infant entrainment?

The infant world is an overwhelming tapestry of interleaved rhythms that cross into all sensory domains. Rhythms are important not just in natural and manmade environments but also play a key role in social development (Darnley, 1981; Kotz, Ravignani & Fitch, 2018; Mogan, Fischer & Bulbulia, 2017; Trainor & Cirelli, 2015). Infants need to learn to perceive, separate and understand these rhythms to become fully participant in modern life (e.g. Frischen, Degé & Schwarzer, 2022). For example, computer models of 4-month-olds processing human and non-human languages have shown that rhythmic components of the temporal envelope and acoustic stimuli may be needed to support cognition (Lau, Fyshe & Waxman, 2022).

2.2.1 What are the features of rhythmic stimuli?

Rhythm can be shown across sensory domains as isochronous or quasi-isochronous oscillations within stimuli, e.g. drumming, flashing lights, beeping sounds. The features of oscillatory stimuli can be separated into components that can help infants to understand underlying rhythms. This is most comprehensively explored in the auditory literature, but many of these features can be applied to other sensory stimuli.

Rhythms are built of individual stimulus beats. Rhythms are patterns of onsets and offsets, most often isochronous or quasi-isochronous but can also be constructed of repeating multitone sequences. The beat is classified as the underlying pulse of a rhythm, while meter represents the hierarchical organisation of strong and weak stimulus beats within a rhythm (Repp and Su, 2013). The combined perception of these components helps infants to entrain to external stimuli from both bottom-up and top-down driven processes that are mediated by experience (Repp and Su, 2013).

Rhythms can be created using either amplitude modulation or frequency modulation (AM or FM; see figure 2.1) of the underlying signal. Amplitude modulation is the result of rhythmic shifts to the power of a signal to create a distinguishable change. For example, in the visual domain, oscillating brightness and dimness of a light source; or in the auditory domain modulating the volume of a stimulus would create an amplitude modulated rhythm. Frequency modulated rhythms are the result of oscillatory changes in frequency from a base frequency to either a lower or higher frequency and back. For example, cyclically increasing and decreasing the pitch of an auditory stimulus.

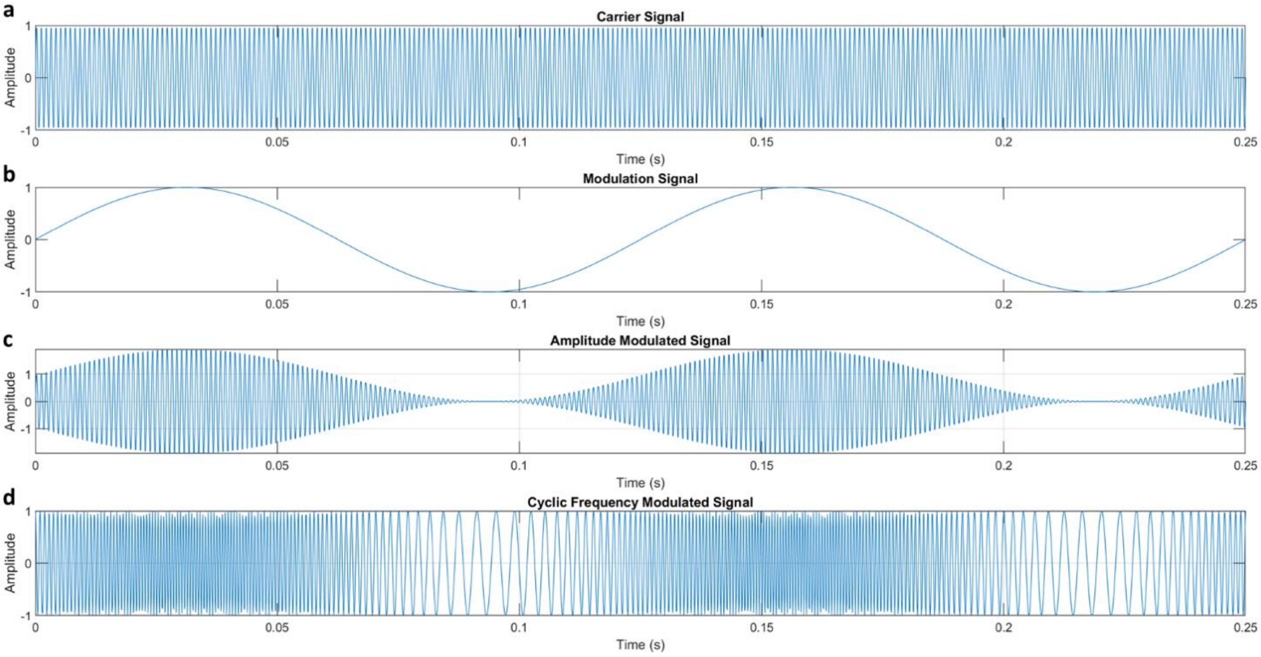


Figure 2.1 Examples of carrier, modulation, amplitude modulated and frequency modulated signals. A) showing 0.25 seconds of a 1000Hz sinusoidal signal without any modulations, b) shows an 8Hz modulation signal, c) carrier signal has been amplitude modulated by the modulation signal, d) the carrier signal has been frequency modulated ising the same modulation signal

2.2.2 Perception of rhythmic stimuli

Frequencies can be further subdivided based on how they are perceived. For example, the auditory frequency spectrum can be split into two broad frequency classes. First, frequencies with individually differentiated beats to the human ear, which will be referred to here as low amplitude modulated frequencies (generally below 50Hz) and those where the beats are perceived as a merged, continuous stimulus, often referred to as pitch, which for humans is often referred to as between 20-20,000Hz. There is some overlap between the two frequency classes depending on hearing ability.

Low amplitude modulated frequencies are where we would consider rhythm to be, which is an important range of frequencies. Slow amplitude modulated rhythms are prevalent in natural sounds, which may carry critical information (Singh et al., 2003). For example, Greenberg et al., (2003) demonstrated that properties of speech under 5Hz generally represent a lower branch of modulated speech, used for heavily stressed syllables, while speech properties with higher frequencies between 6-20Hz generally reflect unstressed syllables.

2.2.3 Phase of rhythmic stimuli

One further aspect of rhythmic stimuli is its phase, or the temporal position within an oscillation. Peaks of rhythmic stimuli generally represent high activity in a rhythmic stimulus, while troughs typically represent low activity. Signals can be concurrently in phase, i.e. when one signal has high activity the other does too; out of phase, when one signal lags behind the other by up to half an oscillation; or anti-phase, when one signal is high the other is low, see figure 2.2, reproduced from section 6.2.5.3, which shows signals in phase, out of phase and anti-phase as well as the phase difference, shown in polar plots and with phase locking value vectors shown in red. Phase alignment can be important for communication between two oscillators, for example this has been shown to be important for the effective communication of neuron populations within the brain (Fries, 2005). Whereas phase misalignment has been suggested to be similar to an inhibitory process of communication between oscillators (Fries, 2005).

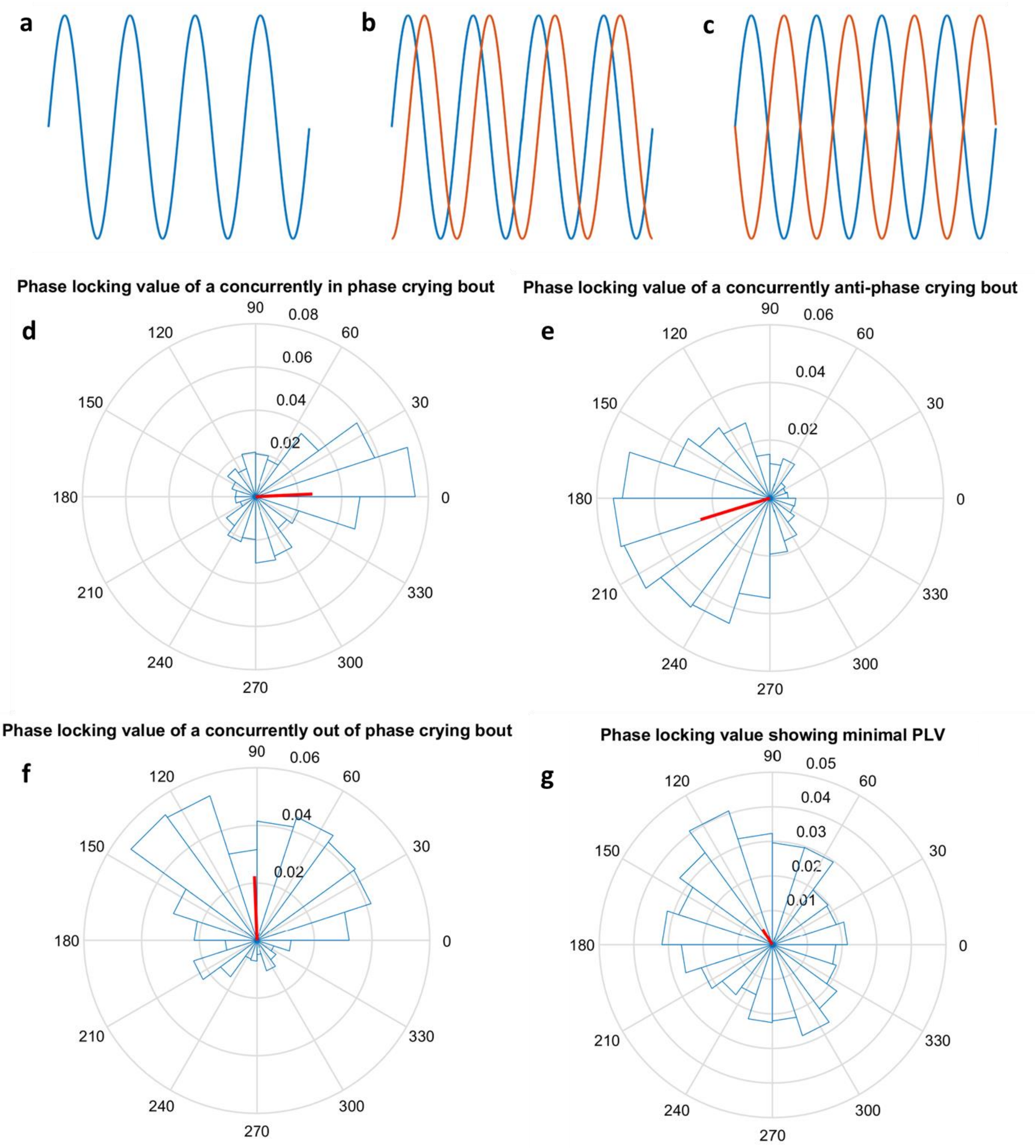


Figure 2.2, reproduced from section 6.2.5.3, showing example diagrams of when signals are a) in phase, when one signal is high the other is high, when one is low the other is low; b) out of phase, where there is a phase shift along the time axis so that one signal is ahead of another; c) showing signals in antiphase, i.e. when one is high the other is low and vice versa. Example circular histograms of real data with mean PLV vector lengths in red are shown in d-g. A longer mean vector length shows a stronger average phase angle and lower phase angle deviation, while spread out values cause a low

vector length and a high phase angle deviation. Mean phase angles are d) concurrently in phase between signals, e) antiphase between signals, f) out of phase between signals and g) spread relatively evenly around the polar plot.

2.2.4 Convergence of stimulus features driving entrainment

To further complicate matters, entrainment between two oscillators can be characterised as a temporal alignment in the frequency, power, phase domains, or a combination of the three. For example, much research has investigated whether there is increased power or phase alignment within frequency bands over particular time periods (e.g. Saby & Marshall, 2010), however this does not necessarily show the temporal alignment of individual peaks of power or phase activity aligned at specific frequencies, simply that there is a general increase in power or phase alignment within a range of frequencies (e.g. delta, theta or alpha frequency bands).

Conversely, amplitude peaks could be shown to be entrained to frequencies outside of the stimulation frequency, for example in work with harmonics of multiple frequencies there may be a response at the stimulation frequency but also at 2, 3, 4 … *n* multipliers of that stimulation frequency (e.g. Zhou et al., 2016). Entrainment between phase and amplitude have also been shown across multiple frequencies with "Phase Amplitude Coupling", which has been useful in differentiating diverse uses of a variety of frequency bands (e.g. Tort et al., 2010). Finally, there are many examples of phase entrainment without corresponding power entrainment, for example in phase resetting research, which has shown phase changes to align with external stimuli without increases in power (e.g. Voloh & Womelsdorf, 2016).

Given that there are many different types of entrainment possible, along with infant development and a wide range of rhythm in early infant environments this presents a potentially much larger challenge for developmental researchers hoping to develop a unified theory of infant entrainment development that progresses through to adult entrainment research.

2.2.5 Rhythm in the early infant environment

Foetuses are exposed to multiple largely isochronous rhythms *in utero* including the mother's heartbeat, respiration rate, the mother's voice, reverberant sounds from walking rhythms and muted external rhythmic sounds (Webb et al., 2015) and these rhythms may in fact play a role in early cortical development (Haslbeck & Bassler, 2018). Therefore, it is as expected that newborn infants have been shown to have remarkable auditory capabilities, many of which show the perception and discrimination of rhythmic stimuli (e.g. Edalati et al., 2023; Háden et al., 2015; Nazzi, Bertoncini & Mehler, 1998; Niepel et al., 2020; Winkler et al., 2009; also see section 2.3).

Behaviourally, abundant evidence suggests that our environment and more so infant environments are filled with rhythmic stimuli (Ciccone, 2013; de Barbaro and Fausey, 2021; Spagnola and Fiese, 2007). Rhythmic experiences have been argued to be foundations for care and a sense of wellbeing, security and learning (Ciccone, 2013). Infant-caregiver interactions have been described as "dances or choreographies" with caregiver adjustments to ensure infant wellbeing (Stern, 1977; 1985). Families usually settle into rhythmic sleeping, feeding, bathing, nappy/clothes changing and play schedules from very early on, which have been suggested to improve infant health and behaviour (e.g. Allen et al., 2016; Blume-Peytavi, et al., 2011; Fallon et al., 2016; Kumar et al., 2022; Spagnola and Fiese, 2007).

This is driven both by infant's internal oscillations (see section 2.3.2; Wass, Perapoch-Amadó & Ives, 2022) as bids for attention, hunger and tiredness, but also caregiver schedules and oscillations such as caregiver hunger, energy levels, hormone levels and work schedules (Feldman, 2003; 2006; 2007; 2012; 2016; Nishihara et al., 2002; Rosenbaum & Morett, 2009). Within larger activity units there are also smaller repetitions representing nested rhythmic structures, for example within the daily cycle of feeding, an infant will suck on a breast or bottle repeatedly and when they are older spoonfuls of food will be delivered to their mouths in a repeated and rhythmic way.

Strictly socially, infant-caregiver face-to-face interactions are also often naturally coordinated and rhythmic (Ciccone, 2013; Tronick & Cohn, 1989), with caregivers able to manage experiences of discontinuity, and support and manufacture behavioural rhythms within a social interaction (Ciccone, 2013). Ciccone (2013) suggested that synchronicity of rhythmic experience between internal and external rhythms is important for infant social development. While longitudinally Tronick & Cohn (1989) showed that caregiver-infant pairs demonstrated both concurrent and sequential synchrony that increased within the dyads between 3 to 9 months old.

Infant interactions have also been shown to have a higher use of rhythmic singing and nursery rhymes (e.g. Ilari 2005; Markova et al., 2019; Mendoza and Fausey, 2021). Caregivers reported that singing was the primary musical activity with their 7–9-month-old infants. Ninety two percent of caregivers surveyed reported that they sang to their children for a range of reasons including: to amuse, induce sleep, as background, to alert, quieten or calm fussy babies and more (Ilari, 2005). Research by Mendoza and Fausey (2021) used day long home recordings to demonstrate that music and caregiver singing are habitual parts of infants' auditory environments rather than specific musical events used to teach infants. This passive learning of infants' auditory environment has been shown to have an impact on infant knowledge and how they are able to perceive their surroundings, as infants have been shown to learn their culture's style of music and at older ages are less able to detect metrical violations of music outside of the infant's culture (Hannon and Trehub, 2005).

Further infant directed language has been repeatedly shown to be adapted to increase rhythmicity with an increased emphasis on key phonological units at the ~2Hz stress rate, ~5Hz syllable rate and ~20Hz phoneme rate suggesting a relationship between infant amplitude envelope sensitivity and the increased periodicity of infant directed language, which has been suggested to grab infant's attention and emphasise important aspects of the vocalisation (Leong and Goswami, 2015, Leong et al., 2017).

2.3 Evidence of infant entrainment to isochronous rhythms

2.3.1 Entrainment to simple isochronous rhythms in laboratory settings

Beginning with very simple and controlled stimuli, much entrainment research has focussed on infant neural entrainment to simple auditory or visual isochronous stimuli. Often referred to as auditory steady state responses (ASSRs) and steady state visual evoked potentials (SSVEPs) these stimuli are reliably produced by computer and can be administered cheaply and in the case of auditory stimuli while the infants sleep (for review see Kabdebon et al., 2022). Neural responses can be measured as rhythmic changes in power (e.g. with an FFT) or changes in phase (e.g. with PLV), which when combined with a signal to noise ratio calculation to remove the 1/f component of brain spectral decompositions can provide a distinct response. This technique is called frequency tagging because the frequency of interest is "tagged" as a response at the same frequency as the original stimulus in the brain. Infants are able to detect rhythm and rhythmic components from birth, for example previous research by Háden et al., (2015) demonstrated neurally with ERPs that newborn infants can detect onsets, offsets and the tempo of rhythms with changing presentation rates shortly after birth. As this response is present very early on it is ideal for testing a variety of oscillatory entrainment features.

It is well established using frequency tagging methods that if an isochronous stimulus is presented to infants even as early as the third trimester of birth, they will show a steady state cortical response (e.g. Daneshvarfard et al., 2019; Edalati et al., 2020; Niepel et al., 2020), this continues after birth through to adulthood (for review see Kabdebon et al., 2022; and chapter 3.1.2). Niepel, et al., (2020) used foetal magnetoencephalography to investigate infant auditory steady state responses to amplitude modulated tones *in utero* at 30 gestational weeks and found rhythmic auditory event-related responses to stimuli at 27 and 42Hz, consistent with entrainment of foetal neural activity with the auditory stimulus. Daneshvarfard et al., (2019) demonstrated entrainment of cortical auditory evoked potentials to rhythmically presented syllable stimuli at 1.6Hz with infants at 31 gestational weeks. However, as will be discussed in section 2.7 on mechanisms of infant entrainment, it is difficult to discern whether a rhythmic response to a rhythmic stimulus is a repeated stimulus response (here known as contingent responding) or whether the repeated stimulus induces oscillatory entrainment, aligning internal neural oscillators with the external stimulus.

One method to differentiate oscillatory entrainment from contingent responding is to produce a rhythm with an ambiguous meter and then track the neural response. Edalati et al., (2023) showed that even premature infants at 32 gestational weeks neurally tracked and entrained to both the rhythm and meter of a metrically ambiguous duple/triple (1.5/1Hz) and metrically unambiguous quadruple (0.75Hz) rhythms consisting of drumbeats. This demonstrates that even without clear markers delineating the metrical structure infants do not simply respond to the beat but impose a top-down meter on that rhythm. Similar findings have been shown at older ages, for example at 6-7 months (Cirelli et al., 2016; Flaten et al., 2022).

Further, studies have investigated the effects of missing beats by producing isochronous stimulus streams and either randomly or methodically removing beats. Mismatch negativity (MMN) responses to the missed beats are recorded in infant ERPs and comparisons are made between infant neural responses to beat and no beat trials. For example, Winkler (2009) studied newborn infant responses to missing metrically important down beats within an isochronous auditory rhythm, despite these beats not marked with any stress or other spectral features. The results demonstrated a neural violation response similar to an adult MMN response. However, when He, Hotson & Trainor, (2009) investigated the development of infant MMN to auditory pattern changes using a similar paradigm, they did not show an MMN response at 2 months, but instead showed a slow positive wave in response to changes in pitch or duration. At 4 months, the infants did show a sharp negative response which similar to adult responses. Taken together results show that while there is evidence of a MMN response to isochronous stimulus streams at various stages of early infancy, it has not been shown consistently. This may be due to the stimulus characteristics which can impact neural steady state responses (for review see section 3.1.5). Also, while it may be argued that a response to a beat that is not present may be considered concrete evidence of oscillatory entrainment, there are neural characteristics that may produce this response without oscillatory entrainment control, such as a reverberating linear dampened response (for further review see entrainment mechanisms section 2.7).

Another method of investigating entrainment to simple isochronous rhythms is to examine behavioural or neural responses to metrical violations in rhythms. In this paradigm, rhythms with different meters are presented to participants and responses to metrical violations are recorded. If participants entrain to the meters by contingently responding to the beats, then it would be expected that responses to all metrical violations would be similar as participants would be responding on a beat-by-beat basis. However, if participants employ oscillatory entrainment, then it would be expected this would be easier with culturally familiar over unfamiliar meters. Behaviourally, Hannon & Trehub, (2005) demonstrated with a familiarization preference looking task that 6-month-old infants were able to detect metrical violations in culturally familiar (duple, triple meter) and unfamiliar (quintuple, septuple meter) rhythmic patterns, but that 12-month-old infants and adults could not. In a follow up Hannon & Trehub, (2005) showed that the 12-month-old infants but not the adults when exposed to culturally unfamiliar meters would allow them to detect the same metrical violations as the 6-month-olds. A further study by Hannon, Vanden Bosch der Nederlanden & Tichko, (2012) showed that after exposing 5-year-olds to culturally unfamiliar rhythms they too were able to detect metrical violations in culturally unfamiliar rhythms, but that older infants up to 11 years old were not able to do the same.

Neurally and using a range of infant directed linguistic stimuli, Suppanen, Huotilainen & Ylinen (2019) investigating how rhythmic structures facilitate learning. Infant directed song, speech and nursery rhymes were recorded with metrical violations and infant responses to these violations was recorded using EEG. The results showed when presented with the most rhythmic stimulus, the nursey rhyme condition, infants showed the best detection of metrical violations and facilitated learning. The authors argued that rhythm may help scaffold future events onto a prediction template aiding metrical violation identification and language learning.

Finally, evidence from clinical studies has demonstrated that not all populations show the same steady state responses to the same stimuli. For example, Peter et al., (2023) showed in 7–12-year-old English speaking children that children with dyslexia or developmental language disorder (DLD) showed similar ASSRs when compared to their neurotypical counterparts at 5Hz and 8Hz but a reduced response at 2Hz. Slower 2Hz frequencies are associated with language and the authors argue that atypical synchronisation likely impairs syntactic and phonological processing which may impair speech. In this example though, the authors do not suggest that there are fundamental differences in the neural architecture between the groups, which may be needed to cause a change in contingent responding to the 2Hz stimuli. Thus, it suggests that at least in the 2Hz condition that the neurotypical infants induce oscillatory entrainment at this frequency in order to better process this information.

2.3.2 Vocalisations, conversations, and language acquisition

Language is arguably one of the most important and complex naturalistic rhythmic structures that young infants encounter on a regular basis. Language is largely rhythmic with multiple hierarchical levels of rhythmic structure, which are thought to help infants with language acquisition (for review see Langus, Mehler & Nespor, 2017; Leong et al., 2017, Leong and Goswami, 2015; Zettersten, 2019). Languages are not uniform however, and can broadly be split based on their rhythmic structure: rhythm based on phenomes, known as mora timed, such as Japanese; syllable timed, such as French and Spanish; or stress timed, such as English and Dutch (Kokelaar 2002). Research into linguistic rhythm largely focusses on prosody, the study of intonation, stress and rhythm that form suprasegmentsals (e.g. Cole, 2012).

Many studies have investigated rhythm perception, production and discrimination, which is not direct evidence for entrainment but does provide context for infant capabilities which may be driven by rhythmic entrainment. Research has shown behaviourally that newborn infants can discriminate between languages based on their rhythmic structure compared to their native language (e.g. Nazzi, Bertoncini & Mehler, 1998). Rhythmic distance between languages has been shown to affect development of speech perception in simultaneously bilingual learners by helping learners to discriminate between languages. Comparing rhythmically dissimilar Spanish-English bilingual learners to an earlier cohort of rhythmically similar Spanish-Catalan bilinguals showed that Spanish-English bilinguals were better able to distinguish acoustically similar vowel contrasts that were phonemic in one but not the other language. The authors argue that this suggests a mechanism of language tagging and sorting that helps bilinguals with statistical learning and word segmentation for the two languages (Sundara & Scutellaro 2011; for further review of the role of rhythm in infant language discrimination see Gasparini et al., 2021).

Rhythm discrimination has been shown to have a long-term positive association with better language processing or an infant's native language. Höhle, et al., (2014) studied German 4-month-olds who completed a rhythm discrimination task between trochaic and iambic disyllabic sequences. The infants were also given SETK3-5 language performance and SON-R nonverbal intelligence tests. The results showed that early rhythmic discrimination correlated with later SETK3-5 outcomes on linguistic but not memory or nonverbal intelligence scores at 5 years old. The authors argue that early rhythm discrimination at 4 months old can correlate with later language skills.

Rhythm perception has been linked to better grammar skills in 5–7 year old children (Gordon et al., 2015). This was further demonstrated with short term priming of rhythm in a study by Fiveash et al., (2020) who primed 7-9 year old children with 8, 16 or 32 second regular or irregular sequences. Being primed by a regular rather than irregular sequence facilitated grammatical judgements of subsequent sentences. Longer priming sequences improved performances with 32 second priming sequences found to be optimal.

Clinically, Ladányi, et al., (2020) conducted a review of associations between speech, language processing and musical rhythm in children with developmental speech, language, attentional and motor disorders and argued that there is a significant body of evidence that individuals with atypical rhythm are at a higher risk of developmental speech and language disorders, called the "Atypical Rhythm Risk Hypothesis". Thus it can be suggested that while rhythm perception and entrainment are not the only ability needed for language acquisition, they do play an important role.

More direct evidence of infant entrainment to linguistic rhythm comes from a near infrared spectroscopy study showed that newborns demonstrated prosodic grouping of native intensity, pitch and duration markers of their native language and if infants were bilingual then both their native languages (Abboub, Nazzi & Gervain 2016). Further, to control for a potential confound that infants experience their native language throughout gestation, newborn infants have also been shown to be able to detect translational probabilities in speech of an artificial language and neurally entrain to the rhythm of the artificially generate syllables in under 2 minutes (Fló et al., 2022). As natural language is rhythmic but not always isochronous, further studies have investigated speech-brain tracking to natural speech and song, which is discussed in sections 2.4 and 2.5.

2.3.3 Conversations and turn taking structures

Conversations and social interactions may include vocalisations and linguistic components, but also more heavily incorporate visual, affective and touch components. Also, while the goal of entrainment to auditory rhythms and vocalisations may be to entrain concurrently, social interactions largely form an anti-phase turn taking structure representing different skills that the infant must learn. Behaviourally, infants have been shown to imitate social partners from birth (for review see Kinsbourne et al., 2002) showing reciprocal interactions. Infants have also been shown to exhibit a proto turn-taking behaviour shortly after birth. In a face-to-face scenario 2-4 day old infants were spoken to by their caregiver, and the caregiver and infant vocalisations were video coded. The results suggested that there was a 1 second post caregiver vocalisation offset period that could form a neonatal period of social contingency (Dominguez et al., 2016).

However, turn-taking cues for adults and infants have been shown to be different. While adults use lexico-syntactic cues to predict turn taking, this feature is still developing in infants, which means infants must rely on prosodic and rhythmic cues within infant directed speech to facilitate turn taking predictions. Kalashnikova & Kember, (2020) studied 1- and 3-year-old infants with an anticipatory looking paradigm where gaze was recorded while observing lexico-syntactically complete IDS or ADS conversational snippets that were either prosodically complete or incomplete. The results showed that 1 year old infants anticipated more turns that were signalled by prosodic cues in prosodically complete IDS but not ADS utterances. 3-year-olds however, were able to rely on prosodic cues in both IDS and ADS. The authors suggest that prosodic information in IDS supports initial turn taking and that prosody remains important through infant language development.

Along with linguistic turn-taking structures, more broadly, infants under 6 months have been shown to begin to use “universal interactional structures”, showing the emergence of adult like communication patterns of eye contact and turn taking structure (Keller et al., 1988; Lense et al., 2022). Caregivers too enhance their own behavioural cues when infants are in this age range by reducing blinking, eye motion, displaying neutral facial affect and increasing wide-eyed positive affect in time with their vocalisations (Lense et al., 2022).

Strong turn taking structures have in turn been shown to be beneficial to infant development. Research by Jaffe, (2001) used an automated dialogic coding scheme to measure infant-caregiver, infant-stranger, caregiver-stranger interactions. The results demonstrated that higher levels of coordination in vocal dialogue predicted attachment and cognition outcomes at 12 months. Differentiation of caregiver-infant vs stranger-infant interactions was informative and suggested differentiation of mothers. Timing patterns of 4-month-olds was seen as procedural/performance knowledge. More recently, infants that showed more frequent turn-taking behaviour was related to better interpersonal neural synchrony and later vocabulary size (Nguyen, Zimmer & Hoehl, 2023).

2.3.4 Infant directed song

As described in section 2.2.5, music and infant directed singing are habitual parts of an infant's daily life and are a prominent method of interaction during infant-caregiver one on one time (Mendoza & Fausey, 2021). Infants have been shown to prefer ID singing over ID speech at 6-10 months in both native and non-native languages (Tsang, Falk and Hessel, 2016). Infant directed song, like infant directed speech has exaggerated properties that may be useful to help infants process this information. Investigating the spectral properties of play songs vs lullabies when sung with ID vs non-ID singing, Trainor et al., (1997) showed that similar to ID speech, ID versions of both types of song had slower tempo, more energy at lower frequencies, longer inter-phrase pauses and higher pitch and jitter factors.

As well as processing simple auditory rhythms, infants at 8 months are also able to successfully process complex musical rhythms with multiple layers of rhythms overlayed on top of one another (Cantiani et al., 2022). Research by Moritz et al., (2013) investigated infant musical training and proficiency has shown that kindergartners who received more musical training had better rhythm perception and that higher rhythm skills were related to increased phonological segmentation skills. This was present both in infants who had better rhythm skills at the start of the year and those who received musical training showed a significant improvement at the end of the year. The authors argue that this demonstrates an association between rhythm sensitivity and oral language acquisition.

Similarly, a review by Virtala and Partanen (2018) highlight early results that music interventions and parental singing promote infants' early language development and neural processing, which could be used as an intervention for infants at-risk of language and auditory development deficits. Neurally, research by Marie, Magne & Besson (2011) in an adult population have shown that musician training aided automatic detection of syllable temporal structure, measured by the P200, which in turn impacted word detection, measured by the N400 and detection of metric violations, as measured by the P600.

2.4 Entrainment to anisochronous stimuli

The above studies investigated isochronous stimuli with a steady or changing tempo to investigate infant entrainment. However, the vast majority of rhythms that an infant may experience in their environment are not likely to be isochronous (for review of rhythmic vs irregular rhythms and adult studies see Rimmele et al., 2018). Even complex rhythms, such as those used by Cantiani et al., (2022) comprised of multiple isochronous rhythms layered upon one another. However, natural sounds may have an underlying dominant rhythm, but the individual beats are likely presented ahead or behind this isochronous dominant rhythm. This presents a problem to the developing infant if entrainment capabilities do not hold true for stimuli that are not perfectly isochronous.

However, there is limited research that has investigated entrainment to anisochronous rhythms. Trehub & Thorpe (1989) demonstrated in a behavioural study that infants between 7-9 months old could categorise rhythms when random variations were included. Lense et al., (2022) demonstrated that when the temporal regularity of an audio-visual singing head stimulus was manipulated this disrupted infant’s entrainment behaviour towards the stimulus. Similar to Lense et al., (2022), Kayser et al., (2015) showed in adults that manipulating pauses in speech did not impact evoked potentials but did impact auditory delta band entrainment to the speech envelope. Suggesting at least for adults that the delta band tracking seen is a top-down entrainment to speech rhythms, whereas this was not similarly investigated in the infant study.

More work investigating entrainment to irregular stimuli has been conducted in adults, where it has been shown that higher levels of expectation and predictability modulates sensory processing of irregular stimuli (Cravo et al., 2013, Keitel, Thut & Gross 2017; Stefanics et al., 2010). In a more naturalistic adult study Kaneshiro et al., (2020) investigated live music entrainment played regularly or by temporally disrupting the music with randomisation of phase spectra, reversal or shuffling measures. The stimuli that retained their rhythm show significantly higher behavioural ratings and neural tracking than phase scrambled controls.

Adult studies that have investigated jittered stimuli, those that have an overall isochronous rhythm, but the individual beats are presented ahead or behind the rhythm, have found entrainment to isochronous but not jittered stimuli (Lehmann, Arias & Schönwiesner, 2016; Tavano, et al., 2022). While other adult studies using a mismatch negativity paradigm have shown that temporally jittered beats did not affect participants detection of deviant beats, but simply had smaller deviant amplitudes than isochronous beats (Bouwer et al., 2016). A replication of this study by Háden et al., (preprint) with a newborn Infant population replicated the same findings as the original study. (see section 4.1.3 for further details).

This mix of results makes it difficult to assess whether there is true entrainment to irregular stimuli, especially those that are jittered. There has also been little work that has investigated how incremental changes in temporal irregularity impact entrainment rates in either adults or infants. This is especially important as so much of our social lives revolves around generally regular rhythms but temporal irregularity of individual stimuli, such as in speech processing.

2.5 How might infant entrainment to rhythm impact infant speech-brain tracking

As an example of potentially the most prevalent quasi-rhythmic stimulus in our lives, naturally produced speech, presents an interesting case study to investigate. Infants are exposed to language from birth and in a muted sense before birth. It is also an incredibly important social structure that allows us to communicate between one another. Multiple studies have demonstrated that speech has layered rhythmic components and that infant directed speech is especially rhythmic (Leong et al., 2017, Leong and Goswami, 2015;

see section 2.3.2). Rhythmic intonation of words in infant directed language has been seen as an important part of language acquisition (e.g. Mehler et al., 1988).

Rhythm may be important in language acquisition and development because it provides a simple and easily predictive framework of language, and it has been suggested that predictive processing informs our understanding of language comprehension and in turn language development (for review see Zettersten, 2019). Infants from 2 years old, have been shown to use multiple linguistic cues in a “looking while listening” paradigm to forward predict responses including: using verb semantics, grammatical gender and coarticulatory cues (for review see Zettersten, 2019). Further, infants at 2 years old recognise word meanings to correctly use verbs to anticipate nouns in a looking paradigm (e.g. *he eats* is likely to be followed by a food noun; Mani and Huettig, 2012). Further, 6-12 month olds have been shown to form auditory word expectations which allow them to more rapidly process further syllables consistent with their knowledge (for review see Zettersten, 2019).

Bergeson & Trehub 2002 showed acoustic analysis of recordings, that caregiver speech and song is relatively consistent over short time periods in terms of pitch and tempo, so much so that the authors suggested that caregivers showed an absolute pitch and tempo. However, over longer time periods, for example over 12 months, research has shown across languages using in home recordings that the tempo of infant directed speech increases to become closer to the tempo of adult directed speech (Narayan & McDermot, 2016).

As well as the structure of infant directed speech showing multiple layers of rhythmicity, infants in a sucking paradigm have shown a preference for socially relevant information such as speech vs spectrally matched, complex non-speech (Vouloanos &

Werker, 2007), and infant directed vs adult directed speech in a visual preference paradigm (Cooper and Aslin, 1990; Pegg, Werker and McLeod, 1992). In fact, infant directed speech has been shown to provide a host of benefits, such as maintaining attention to language, fostering social interaction between caregivers, and heightening the salient points of native languages (Cox et al., 2023; Cristia 2013; Golinkoff et al., 2015). Increased infant brain activity has also been shown to ID vs AD speech, for familiar words at 6 and 13 months and even unfamiliar words at 13 months (Zangl and Mills, 2007). While Fló et al., (2022), showed that sleeping neonates were able to show neural entrainment to transitional probabilities in speech, centred around the first syllable of words after only 2 minutes of exposure. Similarly, in a frequency tagging paradigm infants at 6 months were able to neurally entrain to artificially created 3 syllable words in a stream of computer spoken syllables with no pauses to differentiate words using statistical learning (Choi et al., 2020).

Recent studies have used multivariate temporal response functions (mTRFs) as measures of cortical tracking, by either mapping cortical activity onto features of the original stimuli, such as volume amplitude or spectral features of audio, or luminance or flicker in visual stimuli (for review see Jessen et al., 2021). This is known as a “backwards” model because it attempts to reconstruct features of the original stimulus and has the advantage that the idealised results are known and can be compared to the result obtained. An alternative, a “forwards model”, maps features of the stimulus onto subsequent cortical activity. This can be done for each electrode, which has the advantage of providing spatial information without a priori knowledge of which regions may encode the stimulus the most. Using a forwards model, infants at 7 months old have been shown use low frequency cortical tracking to track recorded infant directed speech but not adult directed speech (Kalashnikova et al., 2018), as well as cartoons (Jessen et al., 2019).

In a similar continuous speech brain coherence paradigm infants at 9 months old have been shown to cortically track both ID and AD speech, but with a significant preferences for ID speech (Menn et al., 2022). It is not clear whether the extra 2 months of developmental time caused infants to show cortical tracking of AD speech, or whether this was due to the stimuli being live and likely much more interesting to the infants. Further cortical tracking studies using cross correlations have shown that infants track phase and power of the speech amplitude envelope in different ways, with phase tracking of native and non-native languages shown at birth and 6 months, while amplitude tracking was shown at birth but not at 6 months (Ortiz Barajas, Guevara & Gervain, 2021) this suggests that there are different paths that each of these features takes when an infant tracks speech.

Taken together the above studies show that by 6-7 months there is evidence for speech entrainment and speech brain tracking and that steps have been taken to investigate different types of speech. However, despite changes in tempo made to slow infant directed when compared to adult directed speech, it is unclear whether this directly facilitates entrainment to the infant directed speech or whether this is used primarily to emphasize salient information within the vocalisations.

2.6 Dysregulation to infant rhythmic entrainment

If synchrony of internal and external oscillators is possible across regular and irregular rhythms then dysregulation of already established rhythms through competing external rhythms should also be possible. To my knowledge there is no research that has directly investigated the use of competing external rhythms to corral infant rhythms to change as a form of active dysregulation of entrainment where an infant would actively be seeking to remain entrained to their original rhythm. This is possibly because there are very few scenarios where a caregiver may want to disrupt an infant's rhythm and entrain them to a competing rhythm. It would also be difficult to investigate this neurally or physiologically, as we cannot be certain that infants have full control over their endogenous neural oscillations. Behaviourally however, by 6 months infants can produce rhythmic vocalisations and movements, which can be matched or contended by behavioural caregiver rhythmic responses.

A naturally occurring negative social infant rhythm is crying, which caregivers will often attempt to soothe with rhythmic vocalisations or rocking behaviour (e.g Bobin-Bègue, 2019). This represents an intriguing naturalistic scenario where a caregiver would use a competing rhythm to actively attempt to disrupt the infant's crying rhythms. Research has investigated crying and soothing techniques, and one study has demonstrated that caregiver vocalisations overlap and compete with infant negative but not positive vocalisations (Yoo, Bowman & Oller, 2018).

However, to my knowledge there are no studies that have investigated competing rocking and infant crying rhythms. In low arousal scenarios, rocking infants has been shown to synchronously entrain respiration rates (Sammon & Darnall, 1994). Research has shown that there are prosocial effects of synchronous rocking or bouncing motions, and the opposite for asynchronous vestibular motions (for a review see Cirelli et al., 2018). Studies in 14 month olds have shown that infants bounced synchronously vs asynchronously with music and an experimenter were more likely to help that experimenter (Cirelli, Einarson & Trainor, 2014), that infants were more likely to help a "friend" of the in sync experimenter (Cirelli, Wan & Trainor, 2016), but that this did not carry over to a neutral party (Cirelli, Wan & Trainor, 2014). This effect was also shown without music to drive the synchronicity (Cirelli et al., 2017). However, further research with infants at 9 and 12 months showed that younger infants did not show a social preference for synchronicity, whereas the older infants showed the same social preference (Tunçgenç, Cohen & Fawcett, 2017).

There is also a need for naturalistic research to quantify natural crying and soothing rhythms, as well as the interaction between crying and soothing rhythms as a form of active disruption of rhythmic activity. This would help our understanding and may begin to answer currently unanswered questions, for example, is there an optimal frequency to disrupt child crying? Do caregivers change their rocking frequency relative to infant crying frequency? Do rhythms actively disrupt crying or would random rocking quieten infants faster?

2.7 Mechanisms of infant entrainment

Multiple methods of rhythmic entrainment have been put forward and tested, primarily in the neural domain (for a review of adult studies see Zoefel, ten Oever & Sack, 2018). One of the primary questions remaining is whether infant entrainment is the result of infants being passive responders, contingently respond to external stimuli; whether infants employ internal neural and physiological oscillators to entrain to external rhythms, known here as oscillatory entrainment; or whether multiple mechanisms are employed simultaneously (e.g. Meyer, Sun & Martin, 2019, 2020).

Entrainment derived from contingent responding represents a readily available mechanism based off of simple stimulus-response that infants would show a rhythmic response to rhythmically presented stimuli from birth. However, a purely contingent responding mechanism of entrainment disregards the internal oscillatory activity that is present, while also implying a deterministic approach to infant interactions.

Mechanisms of infant entrainment may be partly or solely driven by oscillatory entrainment. Within oscillatory entrainment, there is further segregation between “dumb” or “smart” oscillators. Dumb oscillatory entrainment represents a bottom-up driven mechanism. Internal oscillations begin at a particular natural frequency but are caused to change their dominant frequency due to the activity generated from repeated evoked contingent responses to external stimuli that attenuates that frequency to entrain with the external oscillator. While smart oscillators are top-down controlled oscillatory mechanisms that can be guided to focus or inhibit on particular stimulus streams (for an in-depth review see Wass, Perapoch-Amadó & Ives, 2022).

2.7.1 Evidence for contingent responding

Contingent responding to rhythmic stimuli is the easiest explanation for a variety of infant responses. We know that infants show event related potentials across sensory domains (for review see de Haan, 2013) and multiple ERP responses have been shown in response to isochronously presented stimuli without habituation (e.g. Mancini et al., 2011). Therefore, when infants are presented with a rhythmic stimulus, a rhythmic response is expected regardless of the dominant mechanism of entrainment. This makes it difficult to disentangle entrainment driven by contingent responding vs entrainment resulting from oscillatory entrainment. Multiple studies in animal and adults (Ayala, Lehmann & Merchant, 2017; Capilla, et al., 2011; Galambos & Makeig, 1981; Hari, Hämäläinen& Joutsiniemi, 1989; Teichert et al., 2022), reviews (Keitel, Quigley & Ruhnau, 2014; Haegens & Golumbic, 2018) and reinterpretations of previous findings (e.g. Novembre & Iannetti, 2018) have suggested that responses to isochronous rhythms can be explained as the alignment, sometimes called a superposition, of multiple evoked responses which sum together to form an oscillatory pattern.

Studies that have examined responses to stimulus offsets (e.g. in adult studies, Aharoni et al., 2023; Lakatos et al., 2013; Lerousseau et al., 2021) and missing beats (e.g. Winkler et al., 2009) presented during isochronous stimuli have shown a decay after stimulus offset and a mismatch negativity response to missed beats. This has been suggested as evidence that there is a marshalling of activity from an internal oscillator to entrain to the external stimuli. However, there is evidence at least in the neural domain that evoked responses can "reverberate" at particular frequencies, suggesting that evoked responses are characterised by a linear dampened response, i.e. an oscillatory response that echoes with diminishing power repeatedly until fully decayed (see figure 4.1) as seen in the visual domains with the triple-flash illusion (e.g. Gulbinaite, Ilhan & vanRullen, 2016) and dampened harmonic oscillators shown in the auditory domain (Lerousseau et al., 2021). This suggests that these methodologies are not sufficient to differentiate between entrainment mechanisms.

2.7.2 Evidence for oscillatory entrainment

Endogenously driven oscillatory mechanisms rely on there being internally generated rhythms that can be temporally entrained to external rhythms. Infants and adults have been shown to generate internal physiological oscillators, ranging from long term yearly and circadian cycles (for review see Farhud & Aryan, 2018; Hofstra & de Weerd, 2008; Mirmiran & Kok, 1991), through to daily repeating (e.g. Bobin-Bègue, 2019; Refinetti & Menaker, 1992) and short-term cycles such as cardiovascular and respiratory oscillations (for review see Cohen & Taylor, 2002). Neural oscillators have been shown in adult studies over five orders of magnitude which helps promote processing of sensory information, bias input selection and promote long-term consolidation of information (Buzsáki and Draguhn, 2004). Infants have been shown to have lower dominant neural frequencies than adults, at 4-8Hz compared to the dominant 8-12Hz frequency shown in adults (Cellier et al., 2021). This has been shown to change over developmental time, with a switch to a dominant frequency closer to adult frequency bands at around 7-8 years old.

A purely contingent responding mechanism also implies that internal oscillations are predominantly background noise in relation to entrainment responses seen. However, this is contradictory to adult studies and reviews that have shown the benefits of top-down processing on entrainment, (for review see Ding and Simon, 2014, Haegens and Golumbic, 2018; Lakatos et al., 2019; Park et al., 2015; Poeppel and Assaneo, 2020, Thaut, 2013).

A solely contingent responding entrainment mechanism also implies a deterministic approach to infant interactions, with infant's reflexively responding to external stimuli. However, infants have shown behavioural volitional control over sensory inputs, as infants look away or cover their ears in response to over stimulating or negative stimuli and at older timepoints when infants have more verbal and motor control they can move away from overwhelming sensory situations and seek help by conveying their displeasure to caregivers. This suggests that, if top-down neural or physiological inhibition were possible, infants would attempt to inhibit negative or overwhelming stimuli using these processes. This has been seen in research using modified versions of the cocktail party effect for infant participants (Newman & Jusczyk, 1996; Newman, 2005), which suggested across the two studies that between 5 and 13 months infants have some capacity to attend to target auditory stimuli and inhibit noise and that this improves with age.

Increasingly, studies have discussed neural activity not as a passive, stimulus-driven devices but rather more controlled by top-down influences, such as prediction, attention selection and expectation (Engel, Fries and Singer, 2001). One investigative path has studied infant response to beat and meter within isochronous stimulus streams. Winkler et al., (2009) demonstrated that newborn infants can neurally detect the beat in music by showing that infants develop an expectation for the onset of rhythmic cycles despite there being no distinguishing spectral or stress features. After several cycles, the downbeat (onset) of the cycle was omitted, which elicited neural responses associated with violation of sensory expectations. Further 7 months old infants have been shown to mentally construct duple meter in scenarios where there is an ambiguous duple/triple meter (Cirelli et al., 2016). Behaviourally younger infants at 6 months have been shown to be able to "tune in" to a wider range of metrical structures than infants at 12 months and adults (Hannon & Trehub, 2005).

Further, neural entrainment has been shown to be salience driven, with stronger entrainment to intelligible vs unintelligible but otherwise identical speech streams (van Bree et al., 2021). Other studies have investigated attention driven top-down processing as a mediator for entrainment, with predictions shown to attended but not unattended stimuli (Rimmele et al., 2018). Similarly, individual differences have seen in entrainment responses to continuous speech streams, with some adult participants entraining to the stimuli while others did not suggesting either that physical entrainment capabilities are different or that people choose to entrain to stimuli (Assaneo et al., 2019).

A review by Bánki et al., (2022) supported the view of an oscillatory entrainment as a dominant adult mechanism driving neural entrainment, but suggested that more infant entrainment research is needed. Similarly, in a review of infant behavioural, physiological and neural oscillatory entrainment by Wass, Perapoch-Amadó & Ives,

(2022) numerous infant entrainment studies are demonstrated but the evidence for infant oscillatory entrainment is lacking.

2.8 Summary

In summary, infants have been shown to neurally entrain to a wide range of rhythmic stimuli seemingly from very early in infant development. Infants have been shown to entrain to simple and mildly complex rhythms in the first few months of infancy, and to complex and naturalistically complex stimuli in the following months. However, to my knowledge, no research has been conducted into whether infants neurally entrain to particular rhythms over others when all else is held constant. Further research is required to investigate whether infants entrain to some rhythms over others.

The underlying mechanisms for infant neural entrainment are currently unknown, there is evidence of entrainment which could be driven by contingent responding, dumb or smart oscillatory entrainment. There is some evidence of endogenously driven oscillatory entrainment even from birth. However, much of this research relies on isochronous stimuli which could be flawed based on a linear dampened response to these stimuli. Further research is required using alternate stimuli to investigate whether infants exhibit oscillatory entrainment, and whether this changes over developmental time.

Similarly, the importance of rhythm in applied contexts such as language development has been investigated in a patchwork fashion with some research that has investigated the rhythmic structure of infant directed language or music, and others that have investigated infant entrainment responses to particular frequencies. However, there is limited research that has examined infant entrainment to a range of rhythms found in naturalistic, continuous stimuli. Further research is needed to examine the role of rhythm in infant entrainment to language.

Limited research has investigated the effects of actively attempting to disrupt infant entrainment to external rhythms, especially as there are limited situations where actively disrupting infant rhythm could be seen as beneficial. However, investigating disruption of infant produced rhythms could tell us more about infant entrainment to the outside world and whether external or internal rhythms are the stronger force for infant entrainment. If an infant is trying to ignore entrainment but entrains to external rhythms anyway, this could represent the importance of externally driven contingent responses. Research is required which looks at competing rhythms from this perspective.

Chapter 3 – At Which Low Amplitude Modulated Frequency Do Infants Best Entrain? A Frequency Tagging Study

This chapter is a publication of a pre-print article, submitted for publication (Ives et al., pre-print). Subheadings, figure placement, figure and table style, and citation style have been adapted to conform to the thesis format. The supplementary materials for this chapter are available in Appendix B.

Overview

Infants are remarkably attuned to the rhythms of the world around them, displaying early capacities for rhythmic entrainment that lay the foundation for various aspects of cognitive and social development. As discussed in Chapter 2, rhythmic entrainment is a fundamental aspect of human cognition, facilitating synchronization with external rhythms and promoting social coordination. From the earliest stages of life, infants exhibit a remarkable sensitivity to rhythmic stimuli, displaying spontaneous movements and vocalizations that align with external rhythms. This innate capacity for entrainment serves as a precursor to more complex forms of rhythmic processing, laying the groundwork for the development of language, social interaction, and motor coordination.

Chapter 3 of this thesis delves into infant and adult entrainment to rhythmic patterns with a steady tempo. While previous research has shed light on infant rhythmic processing capabilities, significant gaps remain in our understanding. Specifically, little is known about how infants entrain to different frequencies, whether there are topological differences in infant/adult entrainment processing, and procedural aspects of infant entrainment research such as the most appropriate analysis methods and stimulus duration required for significant entrainment responses.

To address these knowledge gaps, this chapter focuses on exploring infant responses to the frequency of isochronous stimuli while controlling for other features. By investigating temporal and spatial responses and comparing analysis methods, we aim to gain insights into the mechanisms underlying infant rhythmic processing and potential differences from adult entrainment mechanisms. Through a comprehensive examination of these factors, we aimed to advance our understanding of infant entrainment and its implications for cognitive and social development.

Abstract

Previous infant entrainment research has shown neural entrainment to a wide range of stimuli and amplitude modulated frequencies. However, it is unknown whether infants show stronger neural entrainment to some frequencies over others, and how low amplitude modulated frequency differs between infants and adults. The current study seeks to address this by testing the neural entrainment of N=51 4–6-month-old infants and N=53 adult caregivers while they listened to sinusoidally amplitude modulated beep stimuli at rest (no sound), 2, 4, 6, 8, 10 and 12 Hz. Analysis examined differences across power and phase, across regions of interest predetermined by previous literature and across segmented time windows. Results showed that infants entrained most strongly at 2 and 4Hz, while adults strongly entrained to frequencies up to 12Hz; that there was no significant difference in power and phase; that entrainment was strongest in the occipital temporal and slightly left fronto-central in adults and right fronto-central and left occipito-temporal in infants. Segmenting by time window did not show any significant increase or decrease in entrainment over time, but longer time windows showed a stronger entrainment response. In conclusion, it is important to choose appropriate stimulation frequencies when investigating entrainment between stimulation frequencies or across ages; whole head recording is recommended to see the full extent of activation; there is no preference on power vs phase analyses; and longer recordings show stronger effects.

Word count: 229

Keywords: Entrainment, Frequency Tagging, Infant EEG, Low amplitude modulated frequencies

3.1 Introduction

Neural activity in mammalian brains during rest and excitation shows an oscillatory structure (Buzsáki, 2006). Oscillations take on the role of pacemaker networks and resonators, each responding to particular firing frequencies (Llinás, 1988). Llinás (1988) describes how these oscillatory networks help specify connectivity during development, help with motor coordination, timing and help with global states, e.g., sleep-wake or attentional state. Buzsáki and Draguhn (2004) demonstrated in the human brain that across five orders of magnitude cortical networks oscillate to help promote processing of incoming sensory information. These neural oscillations help bias input selection, promote long-term consolidation of information and help with time perception.

Person-person and person-environment neural entrainment has been well studied in the adult literature (e.g., Buzsáki and Draguhn, 2004; Glass, 2001; Hoehl, Fairhurst & Schirmer, 2021; Thut, Schyns & Gross, 2011) and to some extent in the infant literature (e.g. Cantiani et al., 2022; Nguyen, Zimmer, & Hoehl, 2023; Wass et al., 2020; Wass, Perapoch Amadó & Ives, 2022). Neural entrainment is measured as matching periods of oscillatory activity, measured either concurrently or simultaneously across a dyad, as a result of a shared stimulus or social interaction, producing a bidirectional influence between the dyad (Wass et al., 2020).

Entrainment studies in adults have suggested that entrainment helps attention selection (Besle et al., 2011; Lakatos et al., 2008; Thut, Schyns & Gross, 2011; Ward, 2003), memory encoding (Ward, 2003), memory recall (Hickey et al., 2020), sensory selection (Schroeder & Lakatos, 2009), auditory perception (Kayser et al., 2015; Rimmele et al., 2018), music processing (Doelling & Poeppel, 2015), speech processing (Ding et al., 2015; Giraud & Poeppel, 2012; Hyafil et al., 2015), interpersonal conversational features including speech pitch, intensity voice quality and speaking rate (Levitab & Hirschberg, 2011) audio-visual processing (Schroeder et al., 2008), synchrony during face-to-face communications (Jiang et al., 2012), visual-olfactory processing (Rekow et al., 2022), prediction of future actions (Kayhan et al., 2022), coordination of movement and auditory rhythm (Varlet, Williams & Keller, 2018) and event timing (Kösem, Gramfort & van Wassenhove, 2014) (for further review see Kabdebon et al. (2022).

Many behaviours are thought to evoke neural entrainment including conversations (Pérez, Carreiras & Duñabeitia, 2017) that feature both infant directed and adult directed speech (Kalashnikova et al., 2018), singing (Weinstein et al., 2016) listening to audiobooks (Koskinen & Seppä, 2014), watching movies (Lankinen et al., 2014) and watching cartoons (Jessen et al., 2019). Entrainment is also thought to be coupled across many physiological markers, including mother-infant synchronisation of heart rhythms (Feldman, 2007a; Feldman, 2011), neural tracking of pain (Guo et al., 2020), thermal-cardiac entrainment (Mannix et al., 1997), infant-adult synchronisation of pupil dilations (Fawcett et al., 2017) and synchronisation of respiration to rocking (Sammon & Darnell, 1994).

For further review of entrainment in early social interaction, and oscillatory entrainment to early social and physical environments please see Nguyen et al., (2020); Hoehl, Fairhurst and Shirmer, (2021); Wass et al., (2020); and Wass, Perapoch Amadó and Ives (2022).

3.1.1 Developmental auditory capabilities

The ongoing auditory scene from our surrounding environment is a constant barrage of overlapping spectral and temporal information that arrives at our ears as a single stream, but which humans separate into distinct sources during auditory processing in the brain.

Näätänen et al., (2001) describe how adults use “sensory intelligence” to untangle this information to produce an auditory scene in the auditory cortex. Young infants, including newborns (McAdams, 1997; Winkler et al., 2003) and 7–15-week-olds (Demany, 1982), have been shown to separate concurrent audio streams of information, suggesting that some elements of how complex auditory sounds are processed seem likely to be present at birth. After grouping and organising the auditory input, humans have then been shown to track and apply hierarchies to the information. For example, this is commonly seen in the cortical tracking of speech structures in adults (Ding et al., 2016) and infants (Choi et al., 2020; Leong & Goswami, 2015). This is also seen across sensory modalities spanning very low frequencies (< 1Hz), including across minutes, hours and days (for a review see Wass, Perapoch Amadó & Ives, 2022).

Infants, even newborns, have been shown to be born with a range of complex auditory processing capabilities that allow them to segregate sounds and organise their auditory world. For example, infants can detect changes in pitch (Alho, 1989), prefer their native maternal language (Moon, Cooper & Fifer, 1993), and recognise and prefer their mother’s voice (De Casper & Fifer, 1980). Machine learning has been shown to reliably classify the neural response of 8-week-olds to both rhythmic and non-rhythmic speech (Gibbon et al., 2021). Using this segregated information, newborns have been shown to model acoustic regularity by learning to group repeating tones out of random patterns (Stefanics et al., 2007), and model acoustic regularity and show a neural response to violations in expected auditory sequences (Carral et al., 2005).

At a more granular level, infants have been shown to employ statistical learning as patterns to perceive acoustically relevant information. Saffran et al., (1999) demonstrated the first evidence of infant statistical learning of tone sequences in both infants and adults. Further studies have demonstrated that infants under the age of 12 months can use statistical learning to pick out words in regular computerised speech (Choi et al., 2020; Fló et al., 2022; Kabdebon et al., 2015), while cortical tracking studies have demonstrated that infants track sung speech (Attaheri et al., 2022), cartoons (Jessen et al., 2019) and infant directed and adult directed speech (Kalashnikova et al., 2018).

3.1.2 Infant vs adult neural entrainment capabilities

Auditory steady state responses (ASSRs) are cortical responses to fast isochronously repeating stimuli, which are often used to test hearing capabilities in newborns and young infants. ASSRs are a valuable technique as they can be administered to adults and infants whether they are awake or asleep (e.g., Cohen et al., 1991; Jerger et al., 1986; Wang et al., 2022). Using ASSRs, infants have been shown to cortically track simple stimuli from a very early age (e.g., Lorenzini et al., 2022). Daneshvarfard et al., (2019) demonstrated that preterm infants even at an average gestational age of 31.48 weeks exhibited a classic ASSR. While Niepel et al., (2020) showed using foetal MEG that infants as young as 30 weeks gestational age exhibit ASSRs *in utero*, suggesting that even with limited capacity to interact with the outside environment foetal brains can track and entrain to rhythmic stimuli. However, while there is evidence that infants can entrain to auditory stimuli as a method of parsing auditory information in a more efficient manner, there is clear evidence that adult and infant auditory processing is not the same, and entrainment studies using adults cannot be used as a proxy for infant auditory entrainment (for a review of dissimilarities between adult and infant auditory processing please see Saffran, Werker & Werner, 2006).

One expected contributor, especially during early infancy, is the maturation occurring in infant auditory cortices and neural networks (e.g., Adibpour et al., 2020). This, in turn, impacts infants' ability to neurally synchronise with their environment and the people around them (Shafer, Yu & Wagner, 2015; Uhlhaas et al., 2010). As an example, the original ASSR study in adults by Galambos et al., (1981) used as a technique to test auditory pathways, showed a distinctive response at 40Hz and was quickly taken up in other studies (Kuwada et al., 1986; Rees et al., 1986; Rickards and Clark, 1984; D. Stapells et al., 1984). However, when compared to adults, there is not such a strong response in infants at higher frequencies including at 40Hz (Maurizi et al., 1990; Picton, 2003). Instead, this response grows as children mature into adults (Rojas et al., 2006) due to the protracted maturation of the auditory cortex (e.g. Adibpour, 2020), where the 40Hz ASSR has been reported to be centred in response to clicks, noise bursts (Hari, Hämäläinen, & Joutsiniemi, S., 1989) and sinusoidal tones (Herdman et al., 2002).

Infants have been shown to entrain at a range of lower amplitude modulation frequencies (see section 1.4) and have been shown to mentally construct meter in scenarios where there is an ambiguous meter (Cirelli et al., 2016). The ability to entrain to these low frequency rhythms may be a direct result of low amplitude modulated stimuli targeted at infants from adult caregivers, which has been shown to consistently change to match an infant's abilities across many languages (Narayan & McDermott, 2016).

3.1.3 Why are slow amplitude modulated rhythms important?

Sounds include two key oscillatory components: a carrier frequency, which is often referred to as pitch; and an amplitude modulated frequency, often referred to as the rate, tempo or rhythm of a repeating sound. Carrier frequencies must be above 20Hz to fall into the audible range of humans. This overlaps with the amplitude modulated frequencies which start at rates below 0.1Hz and can reach into the hundreds of Hz. The point at which amplitude modulated frequencies become carrier frequencies is the point at which the peaks and troughs of the oscillations are no longer individually discernible from neighbouring oscillations.

Slow amplitude modulated rhythms are prevalent in natural sounds, which may carry critical information (Singh et al., 2003). They are a key component of human social interactions such as speech and have consistently been shown to encode temporal cues that help people forward predict social information during social interactions (Ahissar et al., 2001; Alaerts et al., 2008; Aiken & Picton, 2008; Bertoncini et al., 2011; Henry, Herrmann and Grahn, 2017; Wang et al., 2011). Greenberg et al., (2003) demonstrated that properties of speech under 5Hz generally represent a lower branch of modulated speech, used for heavily stressed syllables, while speech properties with higher frequencies between 6-20Hz generally reflect unstressed syllables. Goswami and Leong (2013) investigated the characteristics of speech and suggested that the stress, syllable and part of the phoneme presentation rate were under 20Hz. Similarly, when phase locking value (PLV) of neural signal to white noise between 4-128Hz was calculated by Liegeois-Chauval et al., (2004) the auditory cortices were shown to have the strongest response to the lowest amplitude modulated frequencies between 4-16Hz, matching the range crucial for speech intelligibility.

While in the process of decoding the complex stimuli around an infant that help infants to develop the skilled task of contingent social interactions, it seems especially important that infants are able to neurally entrain to a range of low amplitude modulated signals. This is important in speech, as shown above, but also in the effect of a wider range low amplitude and ultra-low amplitude (< 1Hz) modulated oscillators including respiratory, arousal, sleep-wake and hormone cycles (Feldman, 2006; Feldman, 2007b) on infant-caregiver synchrony. In the words of Feldman (2006) “the organization of physiological oscillators appears to lay the foundation for the infant’s capacity to partake in a temporally matched social dialogue”.

3.1.4 Research of stimuli frequency

Given the importance of low amplitude modulated auditory frequencies to infant development, it is important to ask: to which frequencies do infants best entrain? While there have been some infant-adult or infant studies, the question of entrainment to auditory stimuli has received much more attention in the adult literature.

Many MEG and EEG studies have examined these questions from the perspective of hearing acuity or auditory perception: Luo et al., (2006) investigated phase vs power neural tracking between 0.3-8Hz; Nozaradan, Peretz and Mouraux, (2012), studied neural entrainment to beat and meter in musical rhythms between 0.426-5Hz; Rees, Green and Kay, (1986) looked at the effects of modulation depth on ASSR responses between 0.4-400Hz; Wang et al., (2012) studied sensitivity to spectral bandwidth in speech processing between 1.5-31.5Hz; Picton et al., (1987) examined the effects of amplitude modulation vs frequency modulation in the context of modulation depth between 2-12.7Hz; Peelle, Gross and David, (2013) examined neural responses to speech amplitude envelope between 4-7Hz; Millman et al., (2010) investigated the spatiotemporal structure of the auditory steady state response (ASSR) at 4, 8 and 12 Hz; Henry, Herrmann and Grahn (2017) examined responses to tone duration, onset/offset duration and input patterns between 5-40Hz; Roß, Borgmann and Draganova, (2000) investigated magnitude of neural response between 10-98Hz; Kuwada, Batra and Maher, (1986) looked at the impact of steady state responses on normally hearing vs hearing-impaired participants between 25-350Hz, with the largest steady state responses found between 25-50Hz.

There have been a few studies that have examined adults and infants including: Stapells et al., (1988) who tested 3 week to 28-month-old infants with a variety of steady state responses between 9-59Hz in 5Hz steps and showed no consistent neural peak. Levi, Folsom and Dobie, (1993) presented adults and 1 month old infants with stimuli amplitude modulated between 10-80Hz. Adults showed an increased magnitude of response at 40Hz which then dropped for later frequencies, while infants showed an increased linear response between 10-80Hz. Aoyagi et al., (1994) studied auditory sweeps of frequencies rapidly presented between 20-200Hz in 20Hz steps to gauge hearing acuity as a function of age in participants between 4 months to 15 years and a cohort of adults showing that the optimal steady state response was detected at 80Hz for young infants (2-4 years old). Riquelme et al., (2006) tested 149 newborns with a range of stimuli including carrier frequencies of 500, 1KHz, 2KHz and 4KHz that were amplitude modulated between 25-98Hz to test for the presence and magnitude of steady state responses. Riquelme et al., (2006) also tested the impact of intensity of these stimuli on the infants with stimuli volumes between 20-70dB. They demonstrated that there were steady state responses between 41-88Hz across all carrier frequencies. Pethe et al., (2004; see also Savio et al., 2004) investigated the impact of age on two frequencies, 40 and 80Hz, with infants between 2 months and 14 years old. Pethe et al., (2004) found that the optimal steady state response changed from 80Hz to 40Hz for children aged 13 and above, matching adult responses. Finally, Rickards et al., (1994) investigated the impact of modulation depth on 337 sleeping newborns between 60-100Hz, finding steady state responses at all frequencies tested to varying degrees.

The above research demonstrates that there have been investigations into the impact of stimuli frequency both in adult and infant studies. However, there are two issues with the above studies in relation to the current question. First, the majority of these studies, especially the infant studies, tested frequencies much higher than the low amplitude modulation rate important in speech and speech perception. Second, most of these studies have investigated a range of frequencies for the primary purpose of hearing acuity tests, which use a small number of cycles for each stimulation frequency. While there may be evidence of entrainment at these frequencies it is difficult to measure the extent to which these are the result of endogenous oscillators entraining to the stimuli or exogenously generated stimulus responses (Haegens & Golumbic, 2018; Hoehl, Fairhurst & Schirmer, 2021; Wass, Perapoch Amado & Ives, 2021). It is expected that there will be a stimulus response to any stimuli in the environment that has a high enough intensity or salience, but this does not show the neural tracking of these stimuli after initial habituation, i.e. after the reduction of high magnitude exogenous stimulus responses, driven by a high level of attention to a new stimulus.

3.1.5 Measuring infant neural entrainment with frequency tagging

3.1.5.1 Frequency tagging

While there are multiple ways to investigate the characteristics of infant neural entrainment, one popular method is frequency tagging. Fundamentally, frequency tagging assumes that the neural response to a rhythmic stimulus will be at the same frequency as the stimulation frequency. This allows researchers to “tag” a particular response in a specific brain region. The appeal of frequency tagging is that the stimulus characteristics can be manipulated to test perceptual capabilities (e.g. complex rhythm detection, Nozaradan et al., 2017), test high and low order comprehension (e.g. hierarchy of linguistic components in speech, (Lo et al., 2022) and can be completed with multiple stimulus types in parallel (e.g. social vs non-social stimuli, Vettori et al., 2020). Frequency tagging requires minimal effort from participants, and in the case of auditory frequency tagging can be completed while participants sleep, as sleep has been shown to dull but not remove steady state responses (e.g. Cohen et al., 1991; Jerger et al., 1986; Wang et al., 2022).

3.1.5.2 Stimulus characteristics

Stimulus characteristics have been explored in previous literature. Zhou et al., (2016) have investigated the interpretation of expected responses to single stimuli vs repeated periodic stimuli, single stimulus duration and effects of fast onset/offset stimuli vs sinusoidal amplitude modulated stimuli, as well as expected harmonic responses using stimuli with varying characteristics. While Kabdebon et al., 2022 have suggested considerations for the stimulation rate in frequency tagging studies, e.g. the stimulation frequency should bear in mind the cognitive response time of the region of interest being targeted; frequencies relating to dominant neural frequencies (6-9Hz in infants, 9-12Hz in adults) should be avoided if possible due to the higher resting power at these frequencies; and that there should be at least 4-8 frequency bins between stimulation frequencies if multiple stimuli are being tested. However, while previous research has suggested frequencies to avoid to stop the response clashing with dominant neural frequencies, to the best of our knowledge, there have been no studies that advise on which frequencies show the strongest entrainment.

3.1.5.3 Analysis method

A tagged response can be seen both in the power and phase domains (for a review see Kabdebon et al., 2022). However, while considerations are given to both power and phase analyses in light of different experimental designs, it is not clear from this review if all conditions are equal whether researchers should expect a larger power or phase response, and whether both are needed or if the methods could be used interchangeably. Zoefel, ten Oever and Sack, (2018) describe one issue with using a Fourier transform to measure neural frequency tagged response. Neural mechanisms such as phase resetting of neural signals to better entrain to an ongoing stimulus violate the stationarity of the signal, so when compared to a sinusoid, e.g. during a fast Fourier transform, the power of the signal will be lowered. This may also represent an issue for phase analyses if the phase resetting produces phase angles that are opposed to one another (e.g. 180 degrees apart), as this may make the amplitude of the phase measure cancel itself out, e.g. during a biopolar or tripolar phase angle response. Investigating both power and phase analyses simultaneously could shed light on the appropriateness of each technique with infant and adult populations.

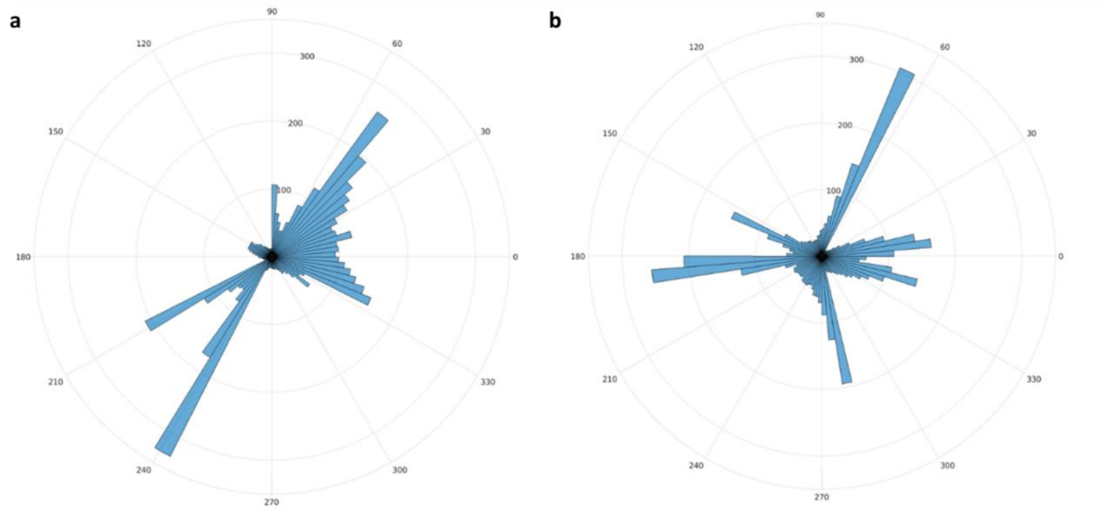


Figure 3.1 Examples of infant data that shows bipolar (a) and tripolar (b) phase responses to 8Hz stimuli. The large poles work to reduce the PLV value to a small response despite there being phase angles shown on the polar plots that are clearly preferred.

3.1.6 Auditory frequency tagging regions of interest

Previous studies have focused on particular regions of interest when conducting auditory frequency tagging. Many of these relate to auditory centres of the brain, especially when the objective of the study is to use auditory frequency tagging as a measure of auditory perception. These include Cz only (e.g. Aoyagi et al., 1994; Cohen, Rickards and Clark, 1991; Jerger et al., 1986; Mühler, Rahne and Verheya, 2013; Pethe et al., 2004; Stappells et al., 1988); Forehead/FPz only (e.g. Alaerts et al., 2009; Levi, Folsom and Dobie, 1993; Maurizi et al., 1990; Rickards et al., 1994; Riquelme et al., 2006; Savio et al., 2001); vertex area (e.g. Choi et al., 2020; Rojas et al., 2006); surrounding the zenith line from anterior to posterior (e.g. Daneshvarfard et al., 2019; Ramos-Escobar et al., 2021); and whole head recordings (e.g. Attaheri et al., 2022, Choi et al., 2020, Cirelli et al., 2016; Herdman et al., 2002; Kabdebon et al., 2015).

Choosing a region of interest is often part of a hypothesis driven approach. Investigating commonly known regions of activity (e.g. electrodes corresponding to A1, Wernicke's area, Broca's area or the perisylvian region), using previous research highlighting topographic activity or reducing unnecessary topographic artefact are seen as methods to minimise false positives. However, there is also the possibility to miss potentially interesting regional activation when investigating a topic where there is limited knowledge, especially as there are developmental changes in the topographic activity due to broader developmental change.

3.1.7 Current study aims

The current study aims to fill these gaps in the research literature by testing infants and adults with a range of low amplitude modulated frequencies to investigate their neural entrainment responses, in order to determine whether there is a preference for any particular frequency. It is expected that there will be an increased response at the naturally occurring dominant frequencies for infants (6-9Hz) and adults (9-12Hz), which will be seen in both power and phase analyses. This is thought to be true because these dominant frequencies can be flexible to help track oscillations in the environment (for a review see Zoefel, ten Oever and Sack, 2018), but increasing the spectral distance between two oscillators increases the intensity of the stimuli to bring two oscillators in alignment, which is the fundamental concept of the Arnold's tongue (Notbohm, Kurths, Herrmann, 2016; Zoefel, ten Oever & Sack, 2018).

In addition, as a secondary aim, we also examined which of several parameters show strongest entrainment. In particular, using the frequency that shows the strongest response for each participant group, we examined differences in the power and phase response; individual differences in power and phase analyses; responses over the time course of the stimulation; and with regional differences across the scalp. From this, we expect to learn lessons that will be valuable to the infant development, neural entrainment and frequency tagging communities when planning similar studies.

3.2 Method

3.2.1 Participants

55 adult-infant dyad participants were recruited as part of a wider longitudinal project. Of these, 4 infant and 2 adult datasets were rejected due to poor data quality. Further, individual conditions were rejected for some participants. As the datasets rejected were not symmetrical between the two groups, this left data collected from 49 dyads and 8 individual infant and adult datasets. The study included 7 conditions: no sound (rest), 2Hz, 4Hz, 6Hz, 8Hz, 10Hz and 12Hz. After preprocessing of EEG data, infants contributed 32, 32, 35, 34, 41, 26 and 32 participants to these passive audio conditions respectively, while adults contributed: 32, 34, 37, 40, 48, 33 and 36 participants to the same conditions. A table detailing which participants completed which conditions has been placed in the supplementary materials (SM B.2). Conditions were presented in a random order. Testing was stopped if the child fussed or woke up, contributing to the uneven sample sizes. Participant information has been included in the supplementary materials (SM B.1).

Mean adult participant age was 35.96 years (standard error, 0.46), mean infant age 166.5 days (standard error 2.43). All adult participants were female, infant participants included 35 male, 21 female. Infants had an average gestational age of 41.15 weeks.

The University of East London ethics committee approved the study. All adult participants provided informed consent for both themselves and their children according to the Declaration of Helsinki. All participants were offered a £10 shopping voucher as a monetary reward for their time. Travel and food expenses were also covered for those that requested them.

The current experiment was part of a longitudinal research project, which tests infant-caregiver dyads at 5-, 10-, 15- and 36-month timepoints. As part of the testing protocol a lab session is run split into three sessions. The data for this study was one of the sessions, the order of which varied depending on when the infants slept during the day.

3.2.2 Stimuli

Audio stimuli were generated in MATLAB 2021b, consisting of pure sinusoidal tones using a carrier frequency of 1000Hz and sinusoidal amplitude modulation at either 2, 4, 6, 8, 10 or 12Hz saved with a sampling frequency of 48KHz. Audio stimuli had a minimum amplitude of -1 and a maximum amplitude of 1, which was then controlled to the target volume of 65dB using on speaker controls and a sonometer. Each audio file was created with 240 seconds of continuous audio.

3.2.3 Procedure

Infant-caregiver dyads were invited into the baby development lab on the Stratford campus of University of East London for a full day of testing. Participants were asked to wear EEG caps (see section 3.2.4) and ECG monitors. The testing day was split into three sessions, a table top interaction session (see section 5.2.2) where participants were asked to play; a passive audio session, which was only administered if the infant fell asleep; and a screen based eye tracking task.

Data collection for this chapter was conducted during the sleeping portion lab testing session, only if the infant fell asleep. The experimental environment changed depending on how the infant was comfortable sleeping. Infants slept in their mothers' arms, in a moses basket, on a sofa, baby sling/carrier or in their pram. The room was kept quiet and in the majority of cases the lights were dimmed to help the child sleep. No auditory sleep aids were permitted.

Creative SBS 250 speakers were placed close to the participants and calibrated to 65dB at the participants' ear using a RS PRO RS-95 sonometer, chosen to be at the average volume of standard speech (e.g. Olsen, 1998).

Adult participants were also instructed to listen to the passive audio task and were permitted to sleep. Participants were asked not to do any rhythmic motions including: chewing, talking, fidgeting, waving, bouncing etc, and asked not to eat or use any electronic devices. Participants were told that if they wanted to stop for any reason or if the infant woke up then the procedure would be stopped.

3.2.4 EEG data acquisition

EEG signals were obtained using a dual BioSemi (Amsterdam, NL) ActiveTwo system configured for 64 channel recording from both participants simultaneously. Participants wore size appropriate 64 channel Electro-Cap International (Ohio, US) caps with a 10-20 electrode montage. Common Mode Sense and Driven Right Leg electrodes between Pz and Poz were used as the active reference. EEG signals were recorded at 512Hz with no online filtering using ActiView data acquisition software (version 7.07; BioSemi).

Signa Gel conductive electrode gel from Parker Laboratories BV (Almelo, NL) was used to bridge the connection between the electrodes and the participant's scalp.

3.2.5 EEG artifact rejection and preprocessing

A modified version of Marriott-Haresign et al.'s, (2021) automated EEG artifact rejection and pre-processing pipeline was implemented. First the data were high pass filtered at 1Hz using the EEGLAB (Delorme & Makeig, 2004) function *pop_eegfiltnew.m* (Widmann, 2008; default settings, filter order 3380). Second, line noise was removed using a notch filter with *pop_cleanline.m* (Mullen, 2012). Third, a low pass filter was applied at 25Hz with the same settings as the highpass filter. Fourth, the data were referenced to a robust average, calculated by first temporarily removing noisy channels using *clean_channels.m* (Kothe, 2014; default settings) and averaging the remaining channels. Fifth, after averaging all channels using the robust average, channels that were subsequently still noisy were rejected using *clean_lines.m* with a correlation threshold of 0.7 and a noise threshold of 4. Subsequently, all channels that were not rejected were put through *eBridge.m* (Alschuler et al., 2014; default settings) to identify bridged electrodes. Participants with less than 25% noisy or bridged channels had these channels interpolated using the spherical method of *eeg_interp.m* (Delorme, 2006). Participants with more than 25% noisy or bridged channels had these channels rejected, data quality metrics are shown in supplementary materials (SM B.3) for each participant and condition. Sixth, using a sliding 1 second window without overlap, epochs were rejected and zeroed out if 70% of the channels exceeded -3.5 to 5 standard deviations of a robust estimate of channel EEG power. This was completed using *clean_windows.m* (Kothe, 2010). Supplementary materials (SM B.4) show that there was no violation of stationarity using this method and no bias towards any frequency. Seventh, all of the completed conditions per participant were concatenated and an independent components analysis (ICA) was completed using *poprun_ica.m* (Delorme, 2001a). Eight, ICA components were automatically labelled with *iclabel.m* for adult participants and *iMARA.m* (Marriott-Haresign, 2021) for infant participants. Components labelled with insufficient brain data (< 0.1) were rejected with *pop_subcomp.m* (Delorme, 2001b). Data was then lowpass filtered again using the same parameters as step 3. Finally, data were re-split into their component conditions, and if channels had been entirely rejected the structure of the 64 channels was reasserted by replacing rejected channels with NaNs with the same length as the dataset, simply to make further processing in the spatial domain easier.

Individual and group level data showing the average percentage of channels interpolated vs removed and average percentage of segments zeroed out is shown in supplementary materials (SM B.3). SM B.3 also shows the total number of electrodes and seconds of EEG data included within the current analysis. Each condition recorded was considered one trial except in analysis 3. In the rest, 2Hz, 4Hz, 6Hz, 8Hz, 10Hz and 12Hz conditions, adult participants contributed 104.33, 114, 123.7, 133.75, 160.13, 110.03, 123.92 minutes artifact free EEG data, while infants contributed 111.55, 116.20, 128.27, 124.30, 146.27, 92 and 113.85 minutes of artifact free EEG data respectively.

3.2.6 Data analysis plan

The first analysis investigated whether different low amplitude modulation frequencies cause different strengths of neural entrainment in infant and adult participants. This was completed by comparing the amplitude of the signal to noise ratio value at the target frequency bin across conditions using an ANOVAN. The subsequent analyses will only use the low amplitude modulated frequency which shows the strongest entrainment response. To check whether the statistical model should account for the genetic or environmental relation between infant and adult participants a bivariate correlation between signal to noise ratio (SNR) scores was conducted and no significant correlation was shown ($p > 0.05$; please see SM B.6). As there was no statistical correlation between infant and adult participant scores no further correction was applied.

The second analysis tested the sensitivity of power and phase analyses using fast Fourier transform (FFT) and phase locking value (PLV) respectively. These will be completed on the participant samples and on individual participants.

The third analysis examined whether there is an increase or decrease in entrainment response over time. The four-minute condition was segmented into 2-minute, 1 minute and 30 second epochs. Power and phase analyses were completed to show trends of entrainment over time.

Finally, the fourth analysis studied entrainment responses using predetermined regions of interest from previous literature vs whole head neural responses.

A mixed-method design could have been implemented as an analysis strategy for this work with factors of: frequency (rest, 2, 4, 6, 8, 10 and 12Hz), participant (infant vs caregiver), space (64 electrodes), analysis type (power vs phase), and time (4, 2, 1 minute and 30 second chunks), with each major factor being investigated and post-hoc tests completed on when significant results were found.

However, this method was not used for two main reasons. First each of the analyses presented is a different research question with individual investigative aims. There was no need to include each part of each factor to answer each of these questions, for example, after investigating the frequency tagging response to different low amplitude modulated frequencies (analysis 1) it is not needed to also include these analyses in for example the temporal analysis (analysis 3) as any differences across time segments can be shown with one of the frequencies. Secondly, some of the analyses call for slightly different methods, for example in analysis 1, topographic spatial results are shown as differences across the scalp to investigate if there was a significant difference in frequency tagged responses for each frequency tested. While in analysis 3, to test whether there were differences in the underlying frequency tagged response, electrodes were tested against chance. As a result of this each analysis for each of the below sections was completed separately.

3.2.6.1 Analysis 1 – Between low amplitude modulated frequencies

Our first research question was to investigate whether different low amplitude modulation frequencies cause different strengths of neural entrainment in infant and adult participants. To address this, first for each participant x electrode x condition timeseries, a check was run to ensure that each timeseries was the same length. Any timeseries that was shorter than the required 122880 samples (512Hz x 240 seconds) were zero padded. A check was conducted to determine whether the zero padding would have an impact on the spectral composition, see supplementary materials (section SM4). An FFT was calculated using a modified *myFFT.m* (Schoof, 2017) MATLAB script.

PLV (Lachaux et al., 1999) was calculated for each of the participant x electrode timeseries against frequencies of interest 2Hz above and below the condition’s target frequency in 0.01Hz steps. For example, for the 8Hz condition, a 6-10Hz window was selected giving 401 frequency bins. For the 2Hz condition, frequencies of interest were 1-4Hz in 0.01Hz steps due to the high pass filter completed during preprocessing stage. For the rest condition, as there was no specific frequency with which to contrast any of the frequency bin could have been used, but for a level of consistency between the rest and other conditions the 2Hz frequency bin was used as a surrogate target frequency of interest. The SNR scores for FFT and PLV are shown in SM B.5 alongside the SNR scores for the 2Hz condition as a comparison. Any timeseries that was shorter than the required 122880 samples were zero padded in the same way as the FFT analysis. For each participant x electrode timeseries, a perfect sinusoid signal for each frequency bin within the frequencies of interest range was created with the same temporal length. A phase angle timeseries was calculated for the EEG timeseries and each of the frequencies of interest by taking the Hilbert transform of the data and calculating the phase angle. The difference between the neural phase angle timeseries and the frequencies of interest timeseries were taken, multiplied by the imaginary operator *i* to give a complex number and the exponential was determined for each value. PLV was calculated as the average exponential value in the timeseries. One PLV value is calculated for each frequency of interest.

After power (FFT) and phase (PLV) analyses were completed on the cleaned data sets in MATLAB 2021b (see Code Availability statement for scripts used) the resulting data were then passed through a SNR script to standardise the data across conditions by removing the 1/f component of the neural data for each electrode x participant (e.g. Vettori et al., 2019).

SNR scores were calculated using a moving window of 25 frequency bins for each dataset between the first and last frequency of interest, for each participant x condition spectral series, in 0.01Hz steps. Of the 24 bins either side of the central frequency, the two closest frequency bins, and two bins within the remaining range with the maximum absolute difference from the target were removed and the average of the remaining 20 bins was taken from the central frequency bin.

For each condition x electrode x participant, the amplitude of the SNR score at the stimulation frequency was taken. An N-ways ANOVA (ANOVAN), *anovan.m,* was completed to test the significance of the SNR amplitude across stimulation frequency conditions (rest, 2Hz, 4Hz, 6Hz, 8Hz, 10Hz and 12Hz). An ANOVAN was chosen due to the uneven participant numbers per condition. Multiple comparisons, using *multcompare.m,* were completed when an ANOVAN showed significant differences.

3.2.6.2 Analysis 2 – Sensitivity of power and phase analyses

Our second research question tested the sensitivity of power and phase analyses using FFT and PLV respectively. To investigate this, using only the stimulation frequency found to show strongest entrainment in analysis 1, FFT and PLV were calculated for each participant x electrode timeseries in the same way as in analysis 1. The resulting data were then passed through the same SNR script as in analysis 1. For the group wide analyses the data were averaged over electrodes and participants to give one FFT and PLV dataset. For individual participant analyses, the data were averaged over the electrodes only to give one data set for each of the analysis types per participant.

Next a permutation analysis was completed using each dataset. The SNR value at 2Hz was recorded as the test value. Bootstrapping with replacement was completed to create 10,000 surrogate datasets per participant and analysis type using *bootstrp.m* (Efron & Tibshirani, 1993). For each surrogate dataset the original test value was ranked against all surrogate values to give a test value rank. The 10,000 test value ranks generated with the surrogate datasets were then averaged to give an overall rank. The overall rank was considered significant if it fell in the top 5% of ranks.

3.2.6.3 Analysis 3 – Entrainment over time analyses

Our third analysis examined whether there is an increase or decrease in the entrainment response over time. To test this, using only the stimulation condition with the strongest entrainment, to see if there was an increase or decrease in entrainment over time, preprocessed data were segmented into 1 x 4 minute window, 2 x 2 minute windows, 4 x 1 minute windows or 8 x 30 second windows.

FFT and PLV were calculated for each participant x electrode x time window timeseries in the same way as in analysis 1. Subsequently SNR values were calculated in the same way as analysis 1. SNR spectral series were averaged over participant and electrode to give one dataset for infants and adults, for both power and phase analysis, and for each time window. Then, permutation analysis was conducted in the same way as in analysis 2. This generated ranked permutation results in power and phase analyses for each time window and each participant group.

3.2.6.4 Analysis 4 – Entrainment over space analyses

Our final analysis studied entrainment responses using five predetermined regions of interest from previous literature vs whole head neural responses. These five regions of interest were selected prior to the results being known. Electrodes referring to Cz; Fz-FCz-Fz; the central zenith line from anterior to posterior of the head (FPz, Afz, Fz, FCz, Cz, CPz, Pz, Poz, Oz), the zenith line plus two lines of electrodes either side (referred to as expanded zenith line; Fp1, FPz, FP2, AF3, Afz, AF4, F1, Fz, F2, FC1, FCz, FC2, C1, Cz, C2, CP1, CPz, CP2, P1, Pz, P2, PO3, Poz, PO4, O1, Oz, O2), and a square of 9 electrodes surrounding and including Cz (referred to as the vertex area; FC1, FCz, FC2, C1, Cz, C2, Cp1, CPz, CP2) were chosen as the potential regions of interest.

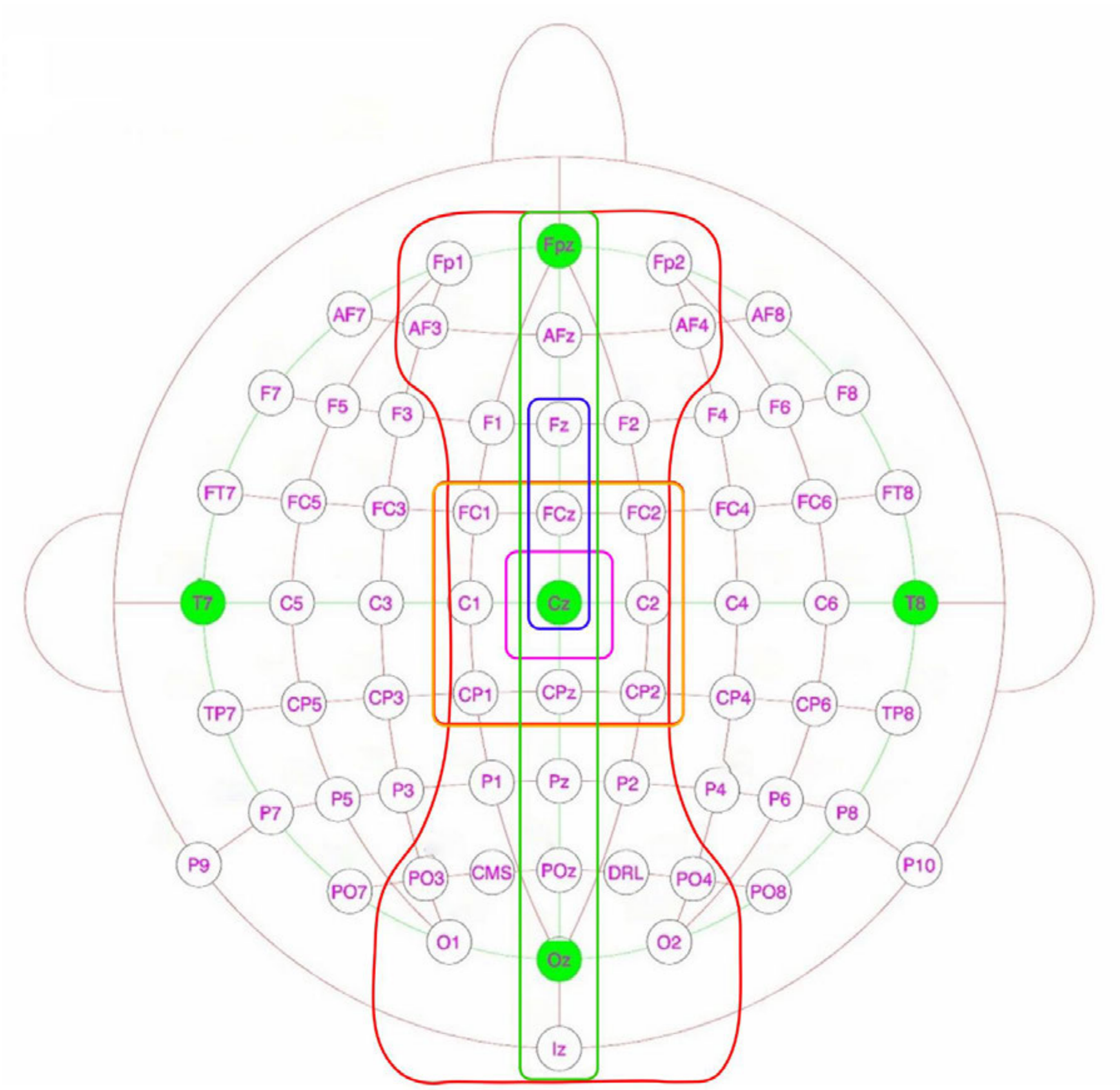


Figure 3.2 64 channel EEG cap using 10-20 montage, with highlighted areas relating to each region of interest chosen. Purple, Cz only; blue, Fz-FCz-Cz; orange, vertex area; green, zenith line; red, expanded zenith line.

For each of these regions of interest, FFT and PLV were calculated for each participant x electrode for the full 4-minute timeseries in the same way as in analysis 1. Subsequently SNR values were calculated in the same way as analysis 1. SNR spectral series were averaged over participant and electrode cluster to give one dataset for infants and adults, for both power and phase analysis, for each region of interest. Then permutation analysis was conducted in the same way as in analysis 2. This generated ranked permutation results in power and phase analyses for each region of interest and each participant group.

Along with these analyses, topographic reporting of the whole head of electrodes was also completed and is shown in the results section to compare the ground truth regions of activation vs the clusters analysed.

3.3 Results

3.3.1 Analysis 1 – Between stimulation frequencies

A comparison of the entrainment to the low amplitude modulated frequency conditions rest (no sound), 2, 4, 6, 8, 10 and 12Hz was completed in both the power and phase domains using FFT and PLV.

3.3.1.1 Power analysis

An ANOVAN comparing different SNR values of FFT results for rest (no sound) and stimulation frequencies (2, 4, 6, 8, 10 and 12Hz) for both infant (figure 3.3a) and adult (figure 3.3b) participants showed that there was a significant difference between conditions in the power domain both for infants ($F(6,63) = [41.65]$, $p < 0.001$) and adults ($F(6,63) = [150.66]$, $p < 0.001$). There were significant differences between electrodes for both infants ($F(6, 63) = [2.1]$, $p < 0.001$) and adults ($F(6, 63) = [10.46]$, $p < 0.001$). Further results investigating topographic differences between electrodes are shown in (section 3.4).

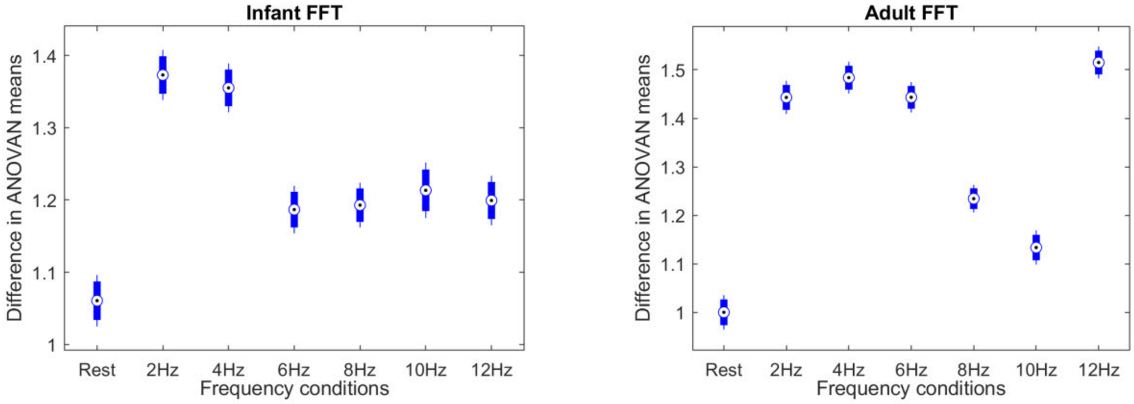


Figure 3.3 Multiple comparisons using multcompare.m between rest, 2, 4, 6, 8, 10 and 12Hz target frequency conditions, completed with an ANOVAN for infant (left) and adult (right) participants comparing SNR of FFT results. Data averaged over the whole head.

Multiple comparisons were completed using the *multcompare.m* Matlab library, which showed that all test conditions were significantly larger frequency tagging response than the rest condition, for both infants and adults ($p < 0.001$). For infants, 2Hz and 4Hz stimuli had the strongest response, being significantly higher than all other conditions ($p < 0.001$). There were no other significant conditions for infant participants. Adults showed a significantly stronger response at 2, 4, 6, and 12Hz when compared to 8Hz and 10Hz ($p < 0.001$), the 12Hz condition was also significantly larger than the 2Hz and 6Hz conditions ($p < 0.05$). There were no other significant results. Statistical differences for adult and infant participants are shown in table 3.1.

*Table 3.1. Differences of means between conditions for infant and adult SNR of FFT, dark blue and ‡ denotes p < 0.001 light blue and * denotes p < 0.05.*

Participant Group		Rest	2Hz	4Hz	6Hz	8Hz	10Hz	12Hz
Infant	**Rest**	NaN	-0.312‡	-0.294‡	-0.126‡	-0.132‡	-0.153‡	-0.069‡
	2Hz	0.312‡	NaN	0.018	0.186‡	0.180‡	0.159‡	0.174‡
	4Hz	0.294‡	-0.018	NaN	0.168‡	0.162‡	0.142‡	0.156‡
	6Hz	0.126‡	-0.186‡	-0.168‡	NaN	-0.006	-0.027	-0.013
	8Hz	0.132‡	-0.180‡	-0.162‡	0.006	NaN	-0.021	-0.006
	10Hz	0.153‡	-0.159‡	-0.142‡	0.027	0.021	NaN	0.014
	12Hz	0.069‡	-0.174‡	-0.156‡	0.013	0.006	-0.014	NaN
Adult	**Rest**	NaN	-0.443‡	-0.484‡	-0.443‡	-0.234‡	-0.134‡	-0.515‡
	2Hz	0.443‡	NaN	-0.041	0	0.209‡	0.309‡	-0.072*
	4Hz	0.484‡	0.041	NaN	0.041	0.249‡	0.350‡	-0.031
	6Hz	0.443‡	0	-0.041	NaN	0.209‡	0.310‡	-0.072*
	8Hz	0.234‡	-0.209‡	-0.249‡	-0.209‡	NaN	0.101‡	-0.281‡
	10Hz	0.134‡	-0.309‡	-0.350‡	-0.309‡	-0.101‡	NaN	-0.381‡
	12Hz	0.515‡	0.072*	0.031	0.072*	0.281‡	0.381‡	NaN

3.3.1.2 Phase analysis

An ANOVAN comparing different SNR values of PLV results between stimulation frequencies for both infant (figure 3.4a) and adult (figure 3.4b) participants both show significant differences (infant, F(6, 63) = [48.65], p < 0.001; adult, F(6,63) = [117.29], p < 0.001). There were significant differences between electrodes for both infants (F(6, 63) = [1.78], p < 0.001) and adults (sig electrodes F(6, 63) = [9.13], p < 0.001).

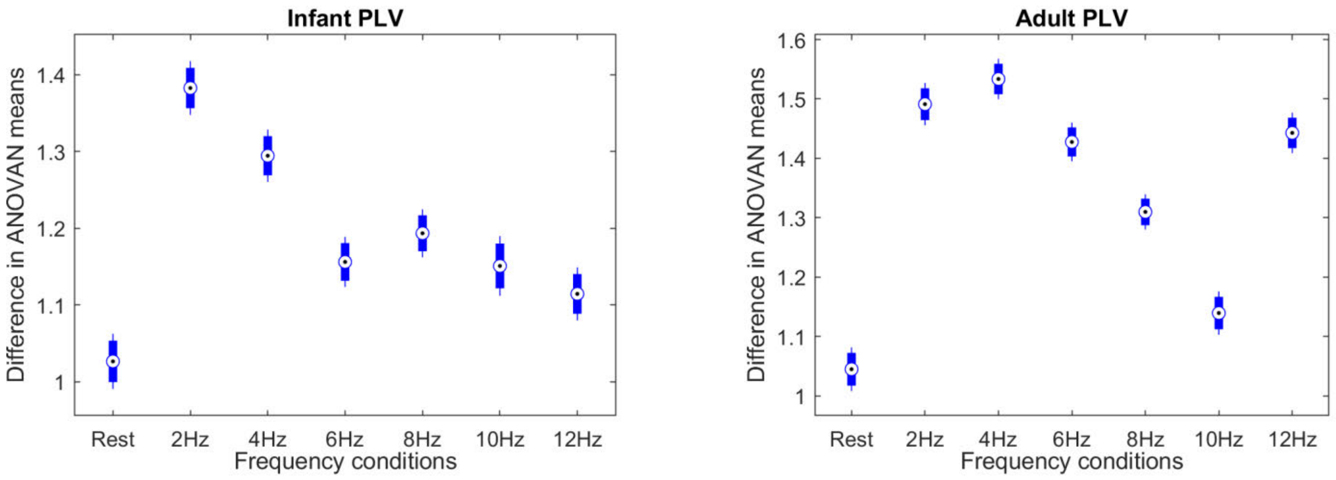


Figure 3.4 Multiple comparisons using multcompare.m between rest, 2, 4, 6, 8, 10 and 12Hz target frequency conditions, completed with an ANOVAN for infant (left) and adult (right) participants comparing SNR of PLV results. Data averaged over the whole head.

Multiple comparisons were completed using the *multcompare.m* Matlab library, which showed that all experimental conditions had a stronger PLV response than the control rest condition (all $p < 0.001$, except when compared against infant 12Hz and adult 10Hz, both $p < 0.01$).

For infant participants the 2Hz stimuli had the strongest response, being significantly larger than all other conditions ($p < 0.001$, except when compared against 4Hz, $p < 0.01$). The 4Hz condition was also larger than all conditions other than the 2Hz condition ($p < 0.001$). Finally, the 8Hz condition was found to be significantly larger than the 12Hz condition ($p < 0.01$).

For adult participants 2, 4, 6 and 12Hz conditions were found to be larger than the 8 and 10Hz conditions ($p < 0.001$), and the 4Hz condition was found to be significantly larger than both the 6Hz ($p < 0.001$) and 12Hz conditions ($p < 0.01$). Significant results are shown in table 3.2 for both adult and infant participants.

*Table 3.2. Differences of means between the rest, 2, 4, 6, 8, 0 and 12Hz conditions for infant SNR of PLV dark blue and ‡ denotes $p < 0.001$, medium blue and † denotes $p < 0.01$ and light blue and * denotes $p < 0.05$.*

Participant Group		**Rest**	**2Hz**	**4Hz**	**6Hz**	**8Hz**	**10Hz**	**12Hz**
Infant	**Rest**	NaN	-0.356‡	-0.268‡	-0.130‡	-0.167‡	-0.124‡	-0.088†
	2Hz	0.356‡	NaN	0.088†	0.226‡	0.189‡	0.232‡	0.268‡
	4Hz	0.268‡	-0.088†	NaN	0.138‡	0.101‡	0.144‡	0.180‡
	6Hz	0.130‡	-0.266‡	-0.138‡	NaN	-0.037	0.005	0.042
	8Hz	0.167‡	-0.189‡	-0.101‡	0.037	NaN	0.043	0.079†
	10Hz	0.124‡	-0.232‡	-0.144‡	-0.005	-0.043	NaN	0.036
	12Hz	0.088†	-0.268‡	-0.180‡	-0.042	-0.079†	-0.036	NaN
Adult	Rest	NaN	-0.446‡	-0.489‡	-0.383‡	-0.265‡	-0.095†	-0.398‡
	2Hz	0.446‡	NaN	-0.042	0.064	0.181‡	0.352‡	0.048
	4Hz	0.489‡	0.042	NaN	0.106‡	0.224‡	0.394‡	0.091†
	6Hz	0.383‡	-0.064	-0.106‡	NaN	0.118‡	0.288‡	-0.015
	8Hz	0.265‡	-0.181‡	-0.244‡	-0.118‡	NaN	0.170‡	-0.133‡
	10Hz	0.095†	-0.352‡	-0.394‡	-0.288‡	-0.170‡	NaN	-0.303‡
	12Hz	0.398‡	-0.048	-0.091†	0.015	0.133‡	0.303‡	NaN

3.3.2 Analysis 2 – Sensitivity of power and phase analyses

Having identified that the strongest consistently shown entrainment between participants and analyses was observed at 2Hz and 4Hz, and that there was a significantly stronger response between 2Hz and 4Hz in the infant phase analysis. The

2Hz frequency was chosen to investigate the sensitivity of power and phase analyses using FFT and PLV at the group and individual level.

3.3.2.1 Comparison of FFT and PLV at a group level

A bootstrapping analysis of the SNR scores derived from FFT and PLV analyses was conducted. Figure 3.5 shows the spectral series for adults and infants for each analysis type. Table 3.3 shows the SNR values for each participant and analysis condition. All conditions were highly significant ($p < 0.01$).

Table 3.3. Mean SNR values at the stimulation frequency of 2Hz for infant and adult participant groups for each analysis type (FFT and PLV), medium blue and † denotes p < 0.01.

	Infant	**Adult**
FFT	1.369†	1.443†
PLV	1.385†	1.491†

To investigate whether there was a difference between infant and adult participant populations, a Welch's t-test was completed between infants and adult SNR scores at the 2Hz stimulation frequency for each analysis type to test for differences between the participant groups. This was not significantly different for any of the groups ($p > 0.05$). An ANOVA comparing the mean ranks achieved for each participant's bootstrapped values between each analysis type was completed to investigate if the results were significantly different between analysis types. This was not significant for either infants or adults ($p > 0.05$).

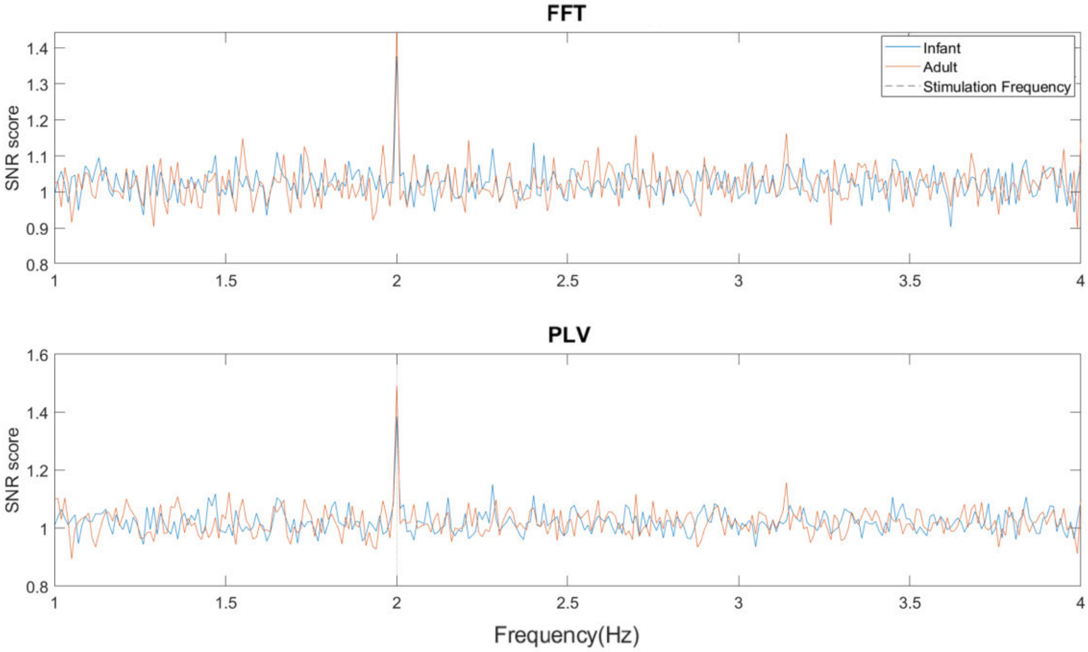


Figure 3.5 SNR scores for each of the analysis types, from 1-4Hz with a frequency resolution of 0.01Hz and a 2Hz stimulation frequency. Adults showed a strong response compared to other frequencies in all analysis types. Infants showed a response for FFT and PLV analyses that was weaker than adults. Data averaged over the whole head.

3.3.2.2 Comparison of FFT and PLV at an individual level

To investigate individual differences in frequency tagging responses and examine if the significant findings at the group level were driven by a small number of individuals or were seen across the group, analyses on a participant-by-participant basis were conducted. The analysis was the same as at the group level (section 3.2.1) but used datasets that had not been averaged over participants. Table 3.4 shows the per participant data and significance. Data quality metrics per participant are shown in Supplementary Materials (SM B.3).

There were similar numbers of adult and infant participants that were found to have a significant frequency tagging response at 2Hz (FFT, 36.4%; PLV 45.5%) compared to infant participants (FFT, 33.3%; PLV 46.6%). There were more highly significant results for adult participants across analyses (35.8%) compared to infant results (23.3%).

*Table 3.4. Mean SNR values per participant (both infant or adult) for each analysis type (FFT and PLV), black shading indicates datasets that were not present due to technical error or poor data quality, medium blue and † denotes p < 0.01 and light blue and * denotes p < 0.05*

	Infant		**Adults**	
	FFT	**PLV**	**FFT**	**PLV**
1042	1.21	1.29	2.25†	2.30†
1051	1.53	1.83†	1.30	1.12
1052	1.32	1.51*	2.47†	2.43†
1053	0.92	0.79	1.67*	1.59†
1054	0.93	0.92	0.79	0.95
1055	■	■	1.83†	1.81†
1059	■	■	1.30	1.22
1061	1.23	1.27	1.79†	1.65†
1062	1.27	1.00	1.71†	2.01†
1063	■	■	1.12	1.09
1065	■	■	1.31	1.34
1068	1.63†	1.68†	2.02†	2.50†
1073	1.57*	1.67†	■	■
1078	■	■	1.37	1.29
1079	2.06†	1.89†	■	■
1082	1.11	1.01	■	■

1083	2.40†	2.26†	0.88	1.09
1084	1.54*	1.55†	1.44	1.26
1085	1.35	1.13	2.14†	2.21†
1086	0.94	0.90	1.01	1.08
1087	1.03	1.01	1.95†	2.61†
1089	0.79	1.00	2.13†	2.07†
1091	1.06	1.16	1.13	1.24
1092	1.37*	1.26	0.98	0.86
1093	1.29	1.39*	1.01	0.87
1096	1.41*	1.61†	0.97	1.09
1097	1.19	1.37*	1.25	1.04
1099	1.54*	1.55†	1.12	0.92
1102	0.93	1.00	0.68	1.00
1103	2.85†	2.79†	1.40	2.05†
1104	1.03	0.99	1.61*	1.48*
1108	1.79†	1.67*		
1109	1.19	1.37*	1.25	1.04
1114			1.28	1.01
1117			1.28	1.61†
1128	0.86	0.92	1.73†	1.91†
1129	1.07	1.02	1.30	1.62*

3.3.3 Analysis 3 – Entrainment over time analyses

Similar to analysis 2, the 2Hz stimulation condition was chosen, as this had shown a consistently strong entrainment response for all participants and analyses (see analysis

1). To investigate whether there is a progression or reduction in entrainment as measured by different analyses, a comparison of four-time lengths was made for power (FFT) and phase (PLV) analyses. Conditions contained 4-minute recordings which were split into either: 1 x 4-minute windows, 2 x 2-minute windows, 4 x 1-minute windows or 8 x 30-second windows. A bootstrapping analysis was used to compare the test value derived from the 2Hz frequency bin for each of the temporal epochs in the same way as in analysis 2. Table 3.5 shows the SNR score at the 2Hz stimulation frequency across segments for each participant group, analysis type and time segments.

*Table 3.5. Averaged SNR value at the 2Hz stimulation frequency split by participant group, analysis and time window. † denotes $p < 0.01$ and * denotes $p < 0.05$.*

<table>
<tr><th>Participant group</th><th>Analysis</th><th>Time window</th><th colspan="8">Test values</th></tr>
<tr><td rowspan="8">Infant</td><td rowspan="4">FFT</td><td>4 minutes</td><td colspan="8">1.369†</td></tr>
<tr><td>2 minutes</td><td colspan="4">1.183†</td><td colspan="4">1.290†</td></tr>
<tr><td>1 minute</td><td colspan="2">1.143†</td><td colspan="2">1.099*</td><td colspan="2">1.135†</td><td colspan="2">1.136†</td></tr>
<tr><td>30 seconds</td><td>1.088*</td><td>1.03</td><td>1.019</td><td>1.122†</td><td>1.049</td><td>1.104*</td><td>1.125†</td><td>1.148†</td></tr>
<tr><td rowspan="4">PLV</td><td>4 minutes</td><td colspan="8">1.385†</td></tr>
<tr><td>2 minutes</td><td colspan="4">1.204†</td><td colspan="4">1.252†</td></tr>
<tr><td>1 minute</td><td colspan="2">1.143†</td><td colspan="2">1.104*</td><td colspan="2">1.131†</td><td colspan="2">1.133†</td></tr>
<tr><td>30 seconds</td><td>1.107†</td><td>1.004</td><td>1.024</td><td>1.082*</td><td>1.030</td><td>1.127†</td><td>1.112*</td><td>1.108*</td></tr>
<tr><td rowspan="8">Adult</td><td rowspan="4">FFT</td><td>4 minutes</td><td colspan="8">1.443†</td></tr>
<tr><td>2 minutes</td><td colspan="4">1.277†</td><td colspan="4">1.237†</td></tr>
<tr><td>1 minute</td><td colspan="2">1.125*</td><td colspan="2">1.125†</td><td colspan="2">1.183†</td><td colspan="2">1.124†</td></tr>
<tr><td>30 seconds</td><td>1.089</td><td>1.128*</td><td>1.125*</td><td>1.303†</td><td>1.143†</td><td>1.109*</td><td>1.188†</td><td>1.019</td></tr>
<tr><td rowspan="4">PLV</td><td>4 minutes</td><td colspan="8">1.491†</td></tr>
<tr><td>2 minutes</td><td colspan="4">1.273†</td><td colspan="4">1.297†</td></tr>
<tr><td>1 minute</td><td colspan="2">1.082</td><td colspan="2">1.227†</td><td colspan="2">1.187†</td><td colspan="2">1.141†</td></tr>
<tr><td>30 seconds</td><td>1.022</td><td>1.067</td><td>1.073</td><td>1.21†</td><td>1.097*</td><td>1.086*</td><td>1.162†</td><td>1.049</td></tr>
</table>

3.3.4 Analysis 4 – Entrainment over space analyses

To test the spatial location of the frequency tagging response in both infants and adults, and to test the validity of regions of interest selected in previous research. A bootstrapping permutation analysis was conducted that was the same as in analysis 2 and applied to each region of interest (electrode locations are listed in section 2.6.4). Table 3.6 shows which areas of interest showed significance, with all areas being significant for both FFT and PLV for adults ($p < 0.01$). In infant data only the Cz only region in the PLV analysis was not significant ($p > 0.05$), all other regions were found to have a significant frequency tagging response ($p < 0.01$).

*Table 3.6. Averaged SNR values at 2Hz for each region of interest per analysis and participant type. Medium blue and † denotes p < 0.01 and light blue and * denotes p < 0.05.*

	SNR of FFT		**SNR of PLV**	
	Infant	**Adult**	**Infant**	**Adult**
All electrodes	1.369†	1.443†	1.385†	1.491†
Cz only	1.298†	2.025†	1.111	1.696†
Fz-FCz-Cz	1.335†	1.235†	1.818†	1.901†
Zenith line	1.317†	1.245†	1.507†	1.529†
Expanded zenith line	1.295†	1.270†	1.495†	1.526†
Vertex area	1.267†	1.207†	1.707†	1.732†

To compare these regions to the significant clusters seen across the whole head, topoplots of SNR scores for FFT and PLV analyses for participant groups are shown in figure 3.6 (a-d), and significance values are mapped with the topoplots in figure 3.6 (e-h).

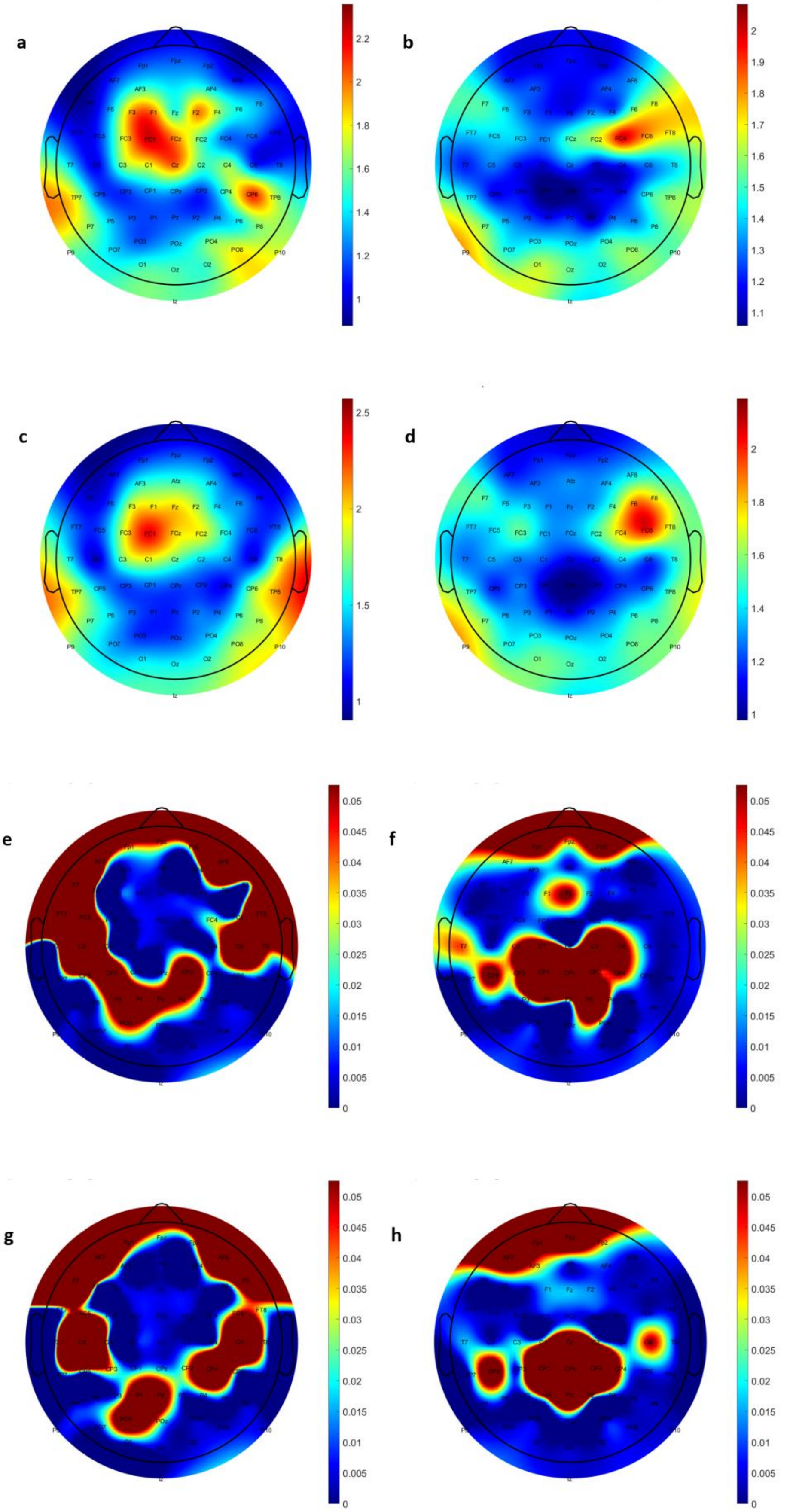

a
b
c
d
e
f
g
h
2.2
2
1.8
1.6
1.4
1.2
1
1.9
1.7
1.5
1.3
1.1
2.5
0.05
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0

Figure 3.6 Topoplots showing the signal to noise ratio value at 2Hz (a-d) and areas of significance (e-h) $p > 0.05$ shown in red, $p < 0.05$ to $p < 0.001$ shown in various shades of blue as shown by the colour bar. Topoplots show a standard 64 channel 10-20 montage. Adult data shown on the left (a, c, e, g), infant data shown on the right (b, d, f, h), FFT data shown in plots a, b, e and f while PLV data is shown in plots c, d, g and h.

3.4 Discussion

3.4.1 Overview

In this study we examined how low amplitude modulated frequencies impact neural entrainment in young (4–6-month-old) infants and their adult caregivers. We also investigated the sensitivity of analysis methods and spatiotemporal characteristics of infant and adult entrainment to the stimulation frequency that showed the strongest neural entrainment. To the best of our knowledge, no previous research has demonstrated a direct comparison of these important low amplitude modulated frequencies under the same conditions.

One issue when comparing across frequencies is that the 1/f nature of neural data makes it difficult to compare absolute values in response to different stimulation frequencies (e.g. Cellier et al., 2021). Another issue is that the range of values expected for each analysis is different (e.g. Gross et al., 2021). To combat this, all results were subjected to a signal to noise ratio calculation to remove the average of the surrounding data and normalise the results. Below when referring to different frequency or analysis conditions these are all in fact the signal to noise ratio versions of these data so that they can be compared.

3.4.2 Between low amplitude modulated frequencies

The results comparing low amplitude modulated frequency conditions found that infants showed the strongest entrainment in both power and phase domains to the 2Hz and 4Hz stimuli when compared to all other conditions. Adults showed a larger range of strong entrainment in the power and phase domains across the 2, 4, 6 and 12Hz stimuli. While many stimulation frequencies have been used in a range of auditory frequency tagging studies (see section 1.4), to our knowledge this is the first study to directly compare this range of low amplitude modulated stimulation frequencies as a measure of infant and adult neural entrainment in the auditory domain and shows that not all low amplitude modulated stimulation frequencies are responded to equally, and that there are developmental differences to be considered. This current result contradicts our hypothesis that there would be an increased response to neural frequencies related to the resting states of the infant (6-9Hz) and adult (9-12Hz) brains. This was thought to be the case due to the increased power at these frequencies and previous research by Notbohm, Kurths and Sack, (2016) which demonstrated that increasing spectral distance between a neural oscillator and stimulus frequency requires an increase in the stimulus intensity to bring the two oscillators into alignment. While there was some increase in the neural entrainment response seen at these dominant frequencies to the stimulation frequency, for infants this was dwarfed by the response to the 2Hz and 4Hz conditions. For adults while there was a strong entrainment response to 12Hz, there was not a strong response seen at 10Hz.

This contradicts previous work (e.g., Kabdebon et al., 2022) that suggested it is best to avoid stimulation frequencies that correspond to these dominant neural frequencies without an appropriate control. However, the current data highlights that there could be an issue when comparing between age groups or when comparing stimuli at a range of stimulation frequencies. Across frequencies, researchers may not be able to differentiate how much of their result is due to the experimental manipulation and how much is due to differences in neural responses at different frequency stimulations, potentially due an underlying endogenous oscillatory activity that favours some frequencies over others.

This is in part similar to previous work in the visual domain investigating the effect of frequency in SSVEPs to a limited extent (for a full review see Köster et al., 2023, especially sectioned 3.6.1 and 4.3.2), which has suggested that there is no relationship between the strongest SSVEP responses and resting dominant frequencies (Birca et al., 2006). However, in contrast to the current results, in the small number of studies that have directly compared frequency of stimulation (e.g. Christodoulou et al., 2018; experiment 1) there was no significant trend found suggesting lower frequency stimuli promote a stronger neural response, with the only significant difference being between the highest frequency tested (15Hz) being significantly lower than two others (10 and 12Hz), while there was no significant differences found between any condition and the 6Hz stimuli.

In the current study, it is interesting that both infants and adults responded more strongly to the two lowest amplitude modulated stimulation frequencies that were presented, which raises further questions. While 2Hz and 4Hz were the stimulation frequencies for infants that showed the strongest response in this study, it is not clear from these results whether this is because 2Hz condition truly shows the strongest response of all possible low amplitude modulated frequencies, or whether this shows the strongest response because this study did not investigate the 0.1-4Hz frequency range in enough detail. Future studies should investigate the neural response to stimuli between 0.1 and 4Hz to further our understanding.

We also considered confounding explanations for the strength of neural response at the lowest AM frequencies. Perhaps the result at 2Hz represents the resting heart rhythm of the infants and adults, who have an average heart rhythm of 2.15Hz and 1.2Hz respectively (Ostchega et al., 2011). Similarly, it could be questioned whether the results are due to the 1/f distribution of the neural signal. However, for both theories, if this underlying activity was present after the signal to noise ratio calculations, this would also have been seen in the resting data collected.

Another possibility is that the result could be attributed to the language heard by participants, as listening native rather than languages with different tempos have been shown. English has many elements that follow a 2Hz structure including a 2Hz "stress" rate (Goswami and Leong, 2013; Leong, 2014), and infant directed speech also shows similar prosodic stress components at 2Hz (Leong and Goswami, 2015). While not all participants were native English speakers, all certainly had proficient English enough to live in the UK and it is to be expected that the infants will have been heavily exposed to English in their environments. Further studies are required in cultures that do not share this property with the English language.

A learned entrainment response is also supported by the larger range of strong entrainment responses seen in adults, who may have been exposed to a larger range of low AM frequencies throughout their lives and would have become more experienced at entraining to a wider range of stimulation frequencies. If entrainment to a range of frequencies could be learned through lived experiences, it would suggest that the responses are an endogenous oscillatory response, as evoked responses should be the same for all participants across all frequencies.

3.4.3 Sensitivity of power and phase analyses

On a group level, investigation of FFT and PLV analyses as a method of detecting frequency tagging responses showed that FFT and PLV demonstrated a significant frequency tagged neural response at the target stimulation frequency of 2Hz. This is consistent with previous studies (see section 1.5.3). There were no significant differences found between infant and adult results at the 2Hz stimulation frequency for FFT or PLV analyses.

On a participant-by-participant level, the results showed that despite the group showing a strong overall entrainment to the condition, not all participants had a significant frequency tagging response at 2Hz and that a significant result using one analysis did not always translate to a significant result across all analysis types. Adult participants had more individual participants that showed a highly significant frequency tagging response than infants.

Individual participant details relating to age, infant gestation and gender are shown in supplementary materials (SM B.2). Infant participants that showed a neural response were not of one age group, gender or having had a particularly long or short gestation period. Data quality metrics are also shown in the supplementary materials (SM B.3). Analyses suggested that data quality was not a significant driver of whether significant entrainment was observed.

3.4.4 Entrainment over time analyses

To investigate whether there is an increase or decrease in entrainment as measured by the frequency tagging response, data were segmented into 4, 2 or 1 minute or 30 second chunks and bootstrapping analyses were conducted on each segment for each participant group. For infant FFT and PLV analyses, data remained significant throughout the 4-, 2-

and 1-minute segments. Data were also significant for the 1st, 4th and 6th to 8th of the 30 second chunks. Adult results showed all the 4-, 2- and 1-minute segments for both FFT and PLV analyses were significant, except the first minute chunk of the PLV analysis. The 2nd-7th and 4-7th segments in FFT and PLV analyses respectively were also found to be significant.

As the data for these segments is the same but divided into smaller and smaller pieces, it could be suggested that taking the spectral series from a smaller timeseries decreases the response seen to the stimulation frequency, which in turn reduces the likelihood of seeing a statistically significant result. This is evidenced by smaller segments of highly significant epochs not being subsequently significant. As expected, then, the length of the stimulation is important to seeing a statistically significant result.

3.4.5 Spatial analyses

Previous studies have used specific regions of interest to investigate frequency tagging. These areas are often chosen to satisfy hypothesis driven research, but in many cases are also centred around auditory regions of the brain. Five ROIs were found in previous literature (see section 1.6), these were each tested to show if they would demonstrate a frequency tagging response. The ground truth frequency tagging response was also shown in topoplots. In adult data all five of the regions of interest showed significant results, while infant data showed significant results for all regions except when the Cz electrode was used alone in the PLV analysis.

Interestingly, when examining the topoplots showing regions of the brain that showed significant vs non-significant entrainment, each of the predetermined regions tested overlapped with regions that were not significant for both adults and infants, across analysis types. This suggests that previous results may be underpowered due to the entrainment differences between electrodes.

Further data showed that while there was a broad range of significant electrodes that showed significant entrainment, there were clusters of increased activity in adult data in a slightly left lateralised fronto-central region and in a temporo-occipital cluster spread across both hemispheres. In infant data a right lateralised fronto-central-parietal cluster was shown along with a left lateralised occipital-temporal cluster. This may suggest that while there is some overlap between the predefined clusters that previous studies have used, there are other areas that are of interest outside the usual regions of interest. The lateralisation seen in infant data and the slight lateralisation on the opposite side of the adult data is also interesting as it may represent a developmental change in response to these stimuli. While no developmental conclusions can be drawn from the current results it is worth further study to see if there is a gradual or sudden shift in significant regional clusters over time.

3.5 Conclusion

Careful choice of low amplitude modulated stimulation frequency should be an important consideration when designing a frequency tagging experiment. Based on current findings choosing a 2Hz or 4Hz low amplitude modulated frequency offers a comparable stimulation frequency between young infant and adult participants. Other stimulus and analysis choices are also worth considering, including having sufficient stimulus length and conducting whole head recordings where possible. There were no significant differences between power and phase analyses.

Chapter 4 – Do Adults And Infants Entrain To The Stimulus Beat Or The Isochronous Rhythm? Evidence From A Novel Jittered Audio Approach

The manuscript is currently in preparation for publication. The supplementary materials for this chapter are presented in Appendix C.

Overview

Chapter 4 of this thesis builds directly upon the findings and methodologies established in Chapter 3, where infant and adult entrainment to rhythmic patterns were explored with a steady tempo. Natural sounds, such as live music or speech, often contain subtle variations in rhythm that deviate from perfect isochrony. Understanding how infants respond to these irregularities is crucial for understanding the mechanisms underlying their entrainment abilities and their implications for cognitive development. While previous studies have touched upon the topic of entrainment to irregular rhythms, the research area remains relatively unexplored.

The primary objective of Chapter 4 is to investigate how infants entrain to increasingly jittered sounds while maintaining the same level of experimental rigor and control established in Chapter 3. By introducing controlled variations in rhythm, we aim to simulate varying levels of real-world complexities and assess infants' capacity to synchronize with these fluctuating rhythms.

Depending on the mechanisms infants employ to synchronize with external rhythms, we anticipate distinct patterns of neural activity in response to jittered stimuli. For example, if infants rely primarily on contingent responding, we may observe neural responses that mirror the fluctuations in rhythm, reflecting their passive alignment with the external auditory input. Alternatively, if oscillatory entrainment mechanisms play a significant role, we might expect to see more robust and sustained neural entrainment, even in the presence of rhythmic variability.

Overall, Chapter 4 represents a significant advancement in our understanding of infant entrainment by exploring their responses to stimuli that simulate the complexity of the real world while maintaining controlled experimental settings. By attempting to bridge the gap between basic research and real-world auditory experiences, this chapter contributes valuable knowledge to the field of infant development and lays the groundwork for future investigations into the complexities of rhythmic entrainment.

Abstract

Neural entrainment mechanisms are hotly contested between evoked contingent responding vs endogenous oscillatory entrainment. Limited studies have investigated neural entrainment to “jittered” audio, audio with a consciously perceived isochronous rhythm but jittered individual beats presented ahead or behind the rhythm. Of those, none have investigated whether the level of temporal irregularity impacts neural entrainment in an infant population. The current study investigated whether participants showed a consistent or linear decreasing neural entrainment response as jitter increases, using a frequency tagging approach, followed by inter trial coherence (ITC) of power amplitude and phase alignment to individual beats vs the isochronous rhythm. Consistency was examined between early infancy with (N=51) 4–6-month-old infants versus (N=53) adult caregivers. Results showed that infants displayed a consistent neural response in power and phase domains as jitter level increased, consistent with an oscillatory entrainment mechanism, whereas adults showed a non-linear decrease in neural response indicative of a mixture of contingent responding and oscillatory entrainment. Results of the ITC analyses demonstrated some oscillatory entrainment and contingent responding, however there was no clear trend between jitter conditions, and results suggested multiple strategies of neural entrainment. No spatial differences between beat and rhythm neural processing were found. Overall, while the results do not point to purely contingent responding or oscillatory entrainment from either participant group, results may suggest individuals utilise both strategies to fully perceive and understand irregular stimuli. Differences between adults and infants may suggest greater infant neural entrainment sensitivity to a wider range of stimuli.

Word count: 248

Keywords: Infant entrainment, jittered audio, contingent response, oscillatory entrainment

4.1 Introduction

Neural entrainment is a popular research subject that has been shown extensively with adults and in limited infants studies, across a range of stimuli, and has been shown to have a wide range of benefits for neural processing (for review see Haegens and Golumbic, 2018; Henao et al., 2020; Lakatos, Gross & Thut, (2019); Oblesser & Kayser, 2019; Saberi & Hickock, 2022; Wass, Perapoch Amadó & Ives, 2022; Zoefel, ten Oever & Sack, 2018).

While debate remains as to the driving mechanism of neural entrainment between exogenously driven contingent responses to external stimuli vs endogenously driven oscillatory entrainment between internal oscillations to external stimuli (for example Meyer, Sun and Martin 2019; 2020) two fundamental issues remain. First, the majority of neural entrainment studies rely on isochronous beats to investigate entrainment to stimulus beat and rhythm, allowing for interpretation using both contingent responding and oscillatory entrainment models. Second, the vast majority of entrainment studies have been conducted in adults, potentially missing developmental clues and comparisons that may help us determine not only what the primary neural entrainment mechanisms are, but also how these were developed over time. This study aims to explore both of these issues using a range of jittered but still rhythmic stimuli, in both infant and adult populations.

4.1.1 Infant neural entrainment

4.1.1.1 Infant understanding of beat and meter

Infants are often shown to be able to perceive rhythm much earlier than they are able to produce rhythmic responses or indeed synchronise these responses to their rhythmic environment (for review see Repp and Su, 2013). For example, children have been shown to prefer rhythmic over non-rhythmic drumming (Einarson & Trainor, 2016). However, infants themselves are unable to synchronise their movements with the beat until later on in development, with rhythmic movement that is not synchronised with the beat appearing in the first two years (Zentner & Eerola, 2010; Rocha & Mareschal, 2016), while simpler rhythmic movements such as synchronous drumming shown reliably at 30 months and in social settings with the infant's caregiver at 18 months (Kirschner & Tomasello, 2010; Yu & Myowa, 2021).

Similarly, perception of rhythms are the precursors of practical applications in auditory and speech processing. Nazzi, Bertoncini & Mehler (1998), demonstrated that newborns were able to classify languages based on global rhythmic properties. While Fló et al., (2022) demonstrated that newborn infants were able to detect translational probabilities in speech of an artificial language and neurally entrain to the rhythm of the artificially generated syllables in under 2 minutes.

Infants as early as 2-4 months old have been shown the be able to successfully discriminate auditory sequences with slight variations in tempo, at a level similar to adult participants (Baruch & Drake, 1997). While preverbal infants at 7 months can entrain and classify a range of rhythmic and melodic patterns such as duple vs triple meters (Hannon & Johnson, 2005). When infants of the same age are presented with rhythms with ambiguous meters, infants have been shown to prefer duple over treble meter (Cirelli et al., 2016). However, this was mediated by infant engagement in music classes and whether the parents were musically trained (Cirelli et al., 2016). Further studies have demonstrated more direct intervention. Bouncing 3-7 month old infants along with a duple or treble meter influences their preferences to that meter (Phillips-Silver & Trainor, 2005).

Interestingly, Hannon and Trehub (2005) showed that Western infants tested at 6 months were able to detect metrical violations in culturally familiar (duple, triple) and unfamiliar (quintuple, septuple) rhythmic patterns, while infants at 12 months were not able to replicate this finding. However, after brief exposure to foreign metrical rhythms the 12-month-olds were able to detect the same violations, while adults were not able to detect culturally unfamiliar metrical violations even with training. This suggests that not only are infants capable of entrainment to a range of rhythms, in an early sensitive period infants show a remarkable plasticity relative to their rhythmic environment. Further study may help researchers to gain a better understanding of these neural entrainment mechanisms and limits.

Taken together, these studies show that when isochronous stimuli are presented to infants, they are often able to entrain at a similar level to adults, and in some cases show a wider range of entrainment not yet pruned for cultural significance. This suggests that problems with infant motor synchrony are related to the neuromuscular development rather than any particular issues with understanding rhythms.

4.1.1.2 Infant neural entrainment to complex stimuli

Infants have not only been shown to be able to entrain to stimuli that are physically present, but also to stimuli that are not directly supported by any acoustic features. For example, a mismatch negativity (MMN).

Similarly, entrainment to information requiring top-down comprehension has also been shown. Using a frequency tagging paradigm Choi et al., (2020), demonstrated that infants presented with fabricated three syllable words randomly spoken using computerised speech entrained to the 3.3Hz syllable rate as well as the 1.1Hz word rate. This was despite there being no acoustic features representing the words, all syllables were spoken with a continuous inter stimulus interval, showing that infants were able to spontaneously segment the continuous audio and neurally entrain to the result.

Utilising more ecologically valid stimuli, Cantiani et al., (2022) investigated entrainment to song and speech with 8 month old infants and adults. Infants showed steady state evoked potentials (SS-EPs) to the rhythmic beats within the music. Interestingly, infants were shown to entrain to the faster beats such as the quavers and semi-quavers, while adults were more likely to entrain to slower beats. Neither infants nor adults were shown to have SS-EPs to natural spoken stimuli, this may be because it lacked a rigid rhythmic structure, which is a prerequisite for SS-EPs.

Other studies have shown neural tracking in the 1-4Hz delta range to sung-speech that was controlled with a 2Hz metronome (Attaheri et al., 2022) and cortical tracking has also been shown to natural infant directed speech, which has been shown to be more rhythmic than adult directed speech (Kalashnikova et al., 2018).

Understanding and neurally entraining to a range of multifaceted potentially acoustically unsupported rhythms and complex structures can not only help infants to process incoming sensory information but also helps infants to understand more abstract dyadic communication structures. This is seen in further studies, which have shown that interpersonal neural entrainment played a beneficial role in infant-caregiver communication (Nguyen et al., 2020a; Nguyen et al., 2020b; Piazza et al., 2019).

4.1.2 Consideration when measuring infant neural entrainment

Multiple reviews have examined the methodologies used to study entrainment in both infant and adult populations with a particular interest in differentiating exogenously driven contingent responding to external stimuli vs endogenous driven oscillatory entrainment between internal neural oscillations and external stimuli (primarily Zoefel, ten Oever & Sack, 2018; but see also Lakatos, Gross & Thut, 2019; Kabdebon et al., 2022; Meyer, Sun & Martin, 2020; Wass, Perapoch Amadó & Ives, 2022). However, to our knowledge no review has written directly on the problem that isochronicity brings to differentiating contingent responses from oscillatory entrainment.

Neural oscillations are naturally present in the brain across five orders of magnitude (Buzsáki and Draguhn, 2004) and have been demonstrated to help facilitate sensory processing, temporally structure incoming information, anticipate upcoming events, distribute attention and perception among other uses (Haegens and Golumbic, 2018). Evoked stimulus responses are also a common, unavoidable and widely studied feature of how brains interact with the outside environment generating both fast responsive neural activity and slow, low frequency evoked responses (Mazaheri & Jensen, 2010). Continuous contingent responses to rhythmic stimuli have been shown to evoke multiple ERP responses in line with stimulus-response, rather than cancelling out activity (Mancini et al., 2011).

Therefore, any model of neural entrainment must account for both mechanisms simultaneously and provide explanations of how models change depending on whether they are largely driven by contingent responding or oscillatory entrainment. Simply demonstrating an isochronously rhythmic response to an isochronously rhythmic stimulus (such as in many frequency tagging studies) cannot be conflated with showing endogenous neural oscillatory entrainment, as this response could just as easily be a repeated contingent response to the stimuli (e.g. Mancini et al., 2011).

Similarly, previous research has demonstrated that evoked responses “reverberate” through the brain with a linear dampened response, i.e. a response that repeats with diminishing power over time (see for example figure 4.1). This has been seen in the visual domain with the triple-flash illusion (e.g. Gulbinaite, Ilhan & vanRullen, 2016) and dampened harmonic oscillators shown in the auditory domain (Lerousseau et al., 2021). This questions neural response results that rely on missing beats or offsets of isochronous rhythms, as figure 4.1 demonstrates, reverberating responses to a rhythmically evoked stimuli can fill the gap of missing beats and show a consistent rhythmic decay after stimulus offset without an underlying oscillator driving the activity. Therefore, results using either missing beats or stimulus offsets with an iscochronous stimulus could equally be the result of oscillatory entrainment of neural oscillations to the stimuli or a linear dampened contingent neural response to the stimulus without any underlying endogenous oscillation.

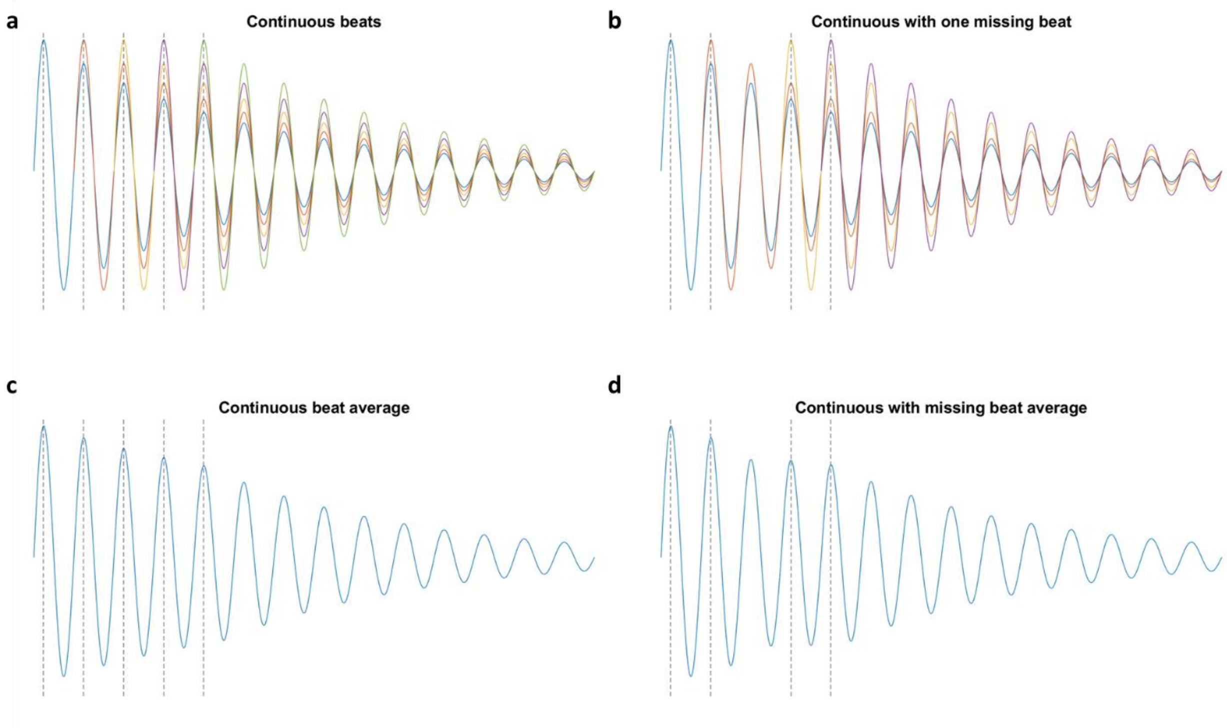


Figure 4.1 Individual (a) and averaged (c) simulated linear dampened responses to beats (shown with dashed lines) and decaying offset. Linear dampened responses of individual (b) and averaged (d) signals with a missing beat show that where a third beat would have been still shows a large response in time with the original after only two preceding beats and no change in offset activity.

Similarly, other adult approaches are not suitable for infant populations, especially very young infant populations. Techniques requiring participants to imagine a particular stimulus or to directly respond to subthreshold or undetected stimuli (for review in adult literature see Zoefel, ten Oever & Sack, 2018), cannot be administered to infant populations as complex instructions cannot be sufficiently conveyed.

Further adult studies have investigated entrainment effects that are unsuitable for developmental populations. As reviewed in Zoefel, ten Oever & Sack, (2018), adult studies have used direct brain stimulation through transcranial magnetic stimulation and transcranial alternating current stimulation to investigate the impact of generating endogenous brain rhythms. However, while safety reports (e.g. Matsumoto & Ugawa, 2017; Zewdie et al., 2020) have suggested that there are no adverse effects of these techniques on infant populations, these techniques are typically not used in during infant research.

Therefore, to investigate contingent responses vs oscillatory entrainment in infant studies requires a stimulus that is: fundamentally rhythmic, without isochronous beats, grounded in physically present stimuli, measured with non-invasive techniques, and must be administered passively to avoid issues with instructions.

4.1.3 Entrainment to jittered audio

While many studies have investigated isochronous rhythms, as seen above, it is rare to find a perfectly isochronous rhythm outside of digital sources. There is a level of irregularity in all live, human produced audio and indeed in many of the mechanically man-made sounds in the regular modern environment.

A selection of studies have examined neural entrainment to "jittered" audio stimuli, predominantly using adult participants. For the present study jittered audio is defined as stimuli that have an overall isochronously defined rhythm, but individual beats are presented ahead or behind that rhythm. Crucially, beat stimuli must be centred relative to the isochronous rhythm using a parametric or flat distribution. A parametric distribution will cluster jittered beats towards the centre of the isochronous rhythm, therefore we suggest that a flat distribution is preferrable. Jittered stimuli produce non-isochronous beats but with an isochronously defined rhythm (see figure 4.2). This paradigm is administered passively to participants, even while sleeping, making it ideal for infant studies.

Using adult participants, Lehmann, Arias & Schönwiesner, (2016) used isochronous and jittered consonant-vowel sounds to investigate SS-EPs responses, while 8% of the stimuli were omitted to also examine MMN. The stimuli had a 40ms duration and intertrial stimulus interval (ISI) of 50ms (20Hz), while the ISI of jittered audio was between 17-87ms randomly assigned using a flat distribution. Results showed SS-EPs were present in isochronous but not jittered conditions, but that both the regular and jittered conditions showed an MMN to the missing beats.

Similarly, research by Tavano, et al., (2022) showed no SS-EPs to rhythms in the delta (1.67Hz) and theta (6.67Hz) range that had been jittered up to 20% in 9 equal steps centred around the isochronous rhythm, randomly assigned using a flat distribution. The study investigated deviant beats but did not investigate differences between EEG responses to deviant beats in isochronous and jittered conditions.

Research by Bouwer et al., (2016) investigated sequential (memory based) learning vs rhythmic entrainment using ERPs to differentiate whether neural entrainment to rhythmic stimuli is driven by a prediction model. Audio consisted of bass and hi-hat drum sounds, alternated to create a two-tone metrical structure. Audio was produced with an isochronous ISI of 225ms (4.44Hz). Jittered beats were uniformly distributed with an ISI between 150-300ms. ERPs demonstrated that temporal jitter of the beats did not affect participants detection of deviant beats, but that isochronous rhythms had larger deviant amplitudes, suggesting stronger rhythm perception.

The only jittered study that the authors are aware of that is conducted with infants is a replication of Bouwer et al., (2016) by Háden et al., (preprint) using newborn infants. The study focussed on whether learning sequential properties of the beat or beat perception were the drivers of neural responses to isochronous stimuli. Repeated sequences of accented and unaccented stimulus beats were either played isochronously or jittered. The results showed similar responses as Bouwer et al., (2016) in newborns as in adults, however as these results have yet to be peer-reviewed they must be viewed with caution.

Finally, a study by Kayser et al., (2015) investigated jittered audio stimuli with adult participants in a more ecologically valid context by manipulating pauses in excerpts of actor performed speeches. The mean and standard deviation across all pauses was maintained, by expanding or shrinking low activity periods parametrically to be 30%, 60% or 90% more irregular than the original. Results showed no reduction of individual speech sound encoding, but delta band entrainment was reduced as jitter increased. The authors suggested that this shows two concurrent neural processes to entrain to individual beats and the stimulus rhythm.

While this study did not have a strictly isochronous rhythm, the study identified frequency bands representing features in the continuous speech and highlights that the jittering procedure did not impact the amplitude envelope of these frequency bands (Kayser et al., 2015, figure 1d). However, Kayser et al., (2015) figure 1b, which describes temporal changes of pauses for each condition, highlights that the parametric distribution of jittering meant that most pause changes were within 0.1 seconds for each level of jitter, suggesting the conditions may not have been as different as they might originally appear. Notwithstanding, this research provides an ecologically valid option, that could be replicated with infant friendly contributions using infant directed speech or song.

Taken together, while the above studies did not show evidence of oscillatory entrainment when investigating SS-EPs, there was evidence of neural processing related to the rhythm in jittered conditions in ERP and MMN results. It seems counterintuitive that across multiple analysis types there are inconsistencies that point to different interpretations of mechanisms behind neural entrainment. As the type of stimulus has been shown to impact steady state responses (e.g. Chapter 1; Kabdebon et al., 2022) studies could investigate alternate stimuli. Similarly, the level of jitter present in the above studies may pass an SS-EP detection threshold, therefore testing multiple levels of jitter may reveal if there are previously unseen tolerances to temporal irregularities. Further, replicating this research with infant populations are also needed to investigate whether early-stage neural entrainment plasticity extends to jittered temporal irregularities.

4.1.4 Current study aims

The current study aims to fill this gap in the research by testing infant and adult populations with a range of jittered audio to investigate neural entrainment responses. The jittered audio created has been designed to generate neural entrainment responses to both the isochronous rhythm and individual beats, with increasing levels of jitter in the audio the results should show either a linear decrease in neural responses representing a purely contingent responding mechanism, or a flat response suggesting an oscillatory entrainment mechanism. Responses between these extremes may show whether there are thresholds beyond which temporal irregularities cannot be entrained, or may elucidate on a mixture of neural entrainment mechanisms.

This study also aims to investigate inter trial power and phase coherence responses at the stimulus beat and isochronous rhythm, which will show an increased power amplitude coherence or phase alignment coherence between trials depending on whether the neural responses are tracking the individual beats of the isochronous rhythm. We hypothesize that the adult neural responses will show a stronger response at the isochronous rhythm. If infants show a high level of neural plasticity to incoming auditory stimuli we expect that infants will show a stronger response to the isochronous rhythm, however if this is not true and they are neurally underdeveloped relative to adults we would expect a stronger response to the stimulus beats.

4.2 Methods

4.2.1 Participants

The current experiment was part of a longitudinal research project, which tests infant-caregiver dyads at 5-, 10-, 15- and 36-month timepoints. As part of the testing protocol a lab session is split into three sessions. The data for this study was one of the sessions from the 5-month timepoint. The order of which varied depending on when the infants slept during the day.

The study included one no sound (rest) control condition, one 8Hz isochronous beat condition (referred to as the 0% jitter condition) and 3 jittered beat conditions: 5%, 10% and 20% jitter, all with an overall rhythm of 8Hz (see section 4.2.2).

For 5-month timepoint 55 adult-infant participant dyads were recruited as part of a wider longitudinal project, 52 adult and 50 infant participants completed at least one condition. Adult and infant average ages were adult, 36.07 years (standard error, 0.43); infant 164.82 days (standard error, 2.39). Average gestational period was 41.12 weeks (standard error, 0.46). All adult participants were female, infant participants included 32 male and 18 female. There were 31, 47, 34, 29 and 29 adult and 31, 40, 27, 23 and 22 infant datasets in the rest, 0%, 5%, 10% and 20% jitter conditions respectively.

A detailed breakdown of which participants completed each of the conditions after EEG preprocessing is shown in supplementary materials (SM C4.1). Conditions were presented in a random order, infants generally slept throughout the conditions and testing was stopped if the child fussed or woke up. Uneven sample sizes are a result of infants not sleeping during the lab session, waking up during the protocol or rejected during preprocessing of the EEG data.

The University of East London ethics committee approved the study. All adult participants provided informed consent for both themselves and their children according to the Declaration of Helsinki. All participants were offered a £10 shopping voucher as a monetary reward for their time. Travel and food expenses were also covered for those that requested them.

4.2.2 Stimuli

4.2.2.1 Isochronous beats

Audio stimuli were generated in MATLAB 2021b, consisting of pure sinusoidal tones using a carrier frequency of 1000Hz and sinusoidal amplitude modulation at 8Hz saved with a sampling frequency of 48KHz. Audio stimuli had a minimum amplitude of -1 and a maximum amplitude of 1, which was then controlled to the target volume of 65dB using on speaker controls and a sonometer. Each audio file was created with 240 seconds of continuous audio.

4.2.2.2 Jittered audio

Jittered beats were produced in a similar manner to the isochronous beats. Produced in MATLAB 2021b, with the same carrier and sampling frequencies, minimum and maximum amplitude, duration of the audio clip and controlled target volume. All jittered beats had an average amplitude modulated frequency of 8Hz.

To create the jitter, the period of each beat was 20% shorter than isochronous beats, with a 10% pause of zeros on either side of the beat. The pause on either side of the beats was then manipulated to shift each of the beats to either precede of follow the average rhythm that would be present if the beats had been isochronous (figure 4.1). The amount that the beat was shifted was based on a flat distribution of lags, with a maximum of 5%, 10% or 20% depending on the condition. This meant that if a beat was shifted to be 10ms ahead of the isochronous rhythm then there would be a corresponding beat that was 10ms after the isochronous rhythm at some point during the audio. The pairs of shifted beats may have proceeded one another, but the lag values were distributed randomly throughout the audio. Figure S1 in supplementary materials (SM4.3) shows the distribution inter-beat intervals, which demonstrates that the random distribution of these lags did not affect the overall flat distribution of lags.

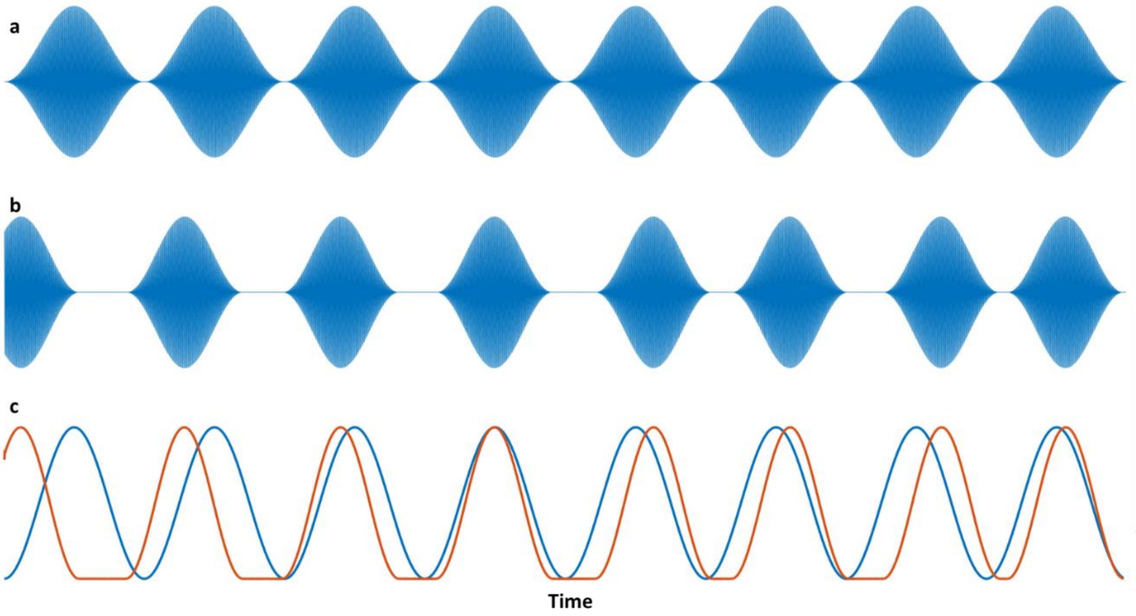


Figure 4.2 showing a) one second of 0% jitter beats at the isochronous rhythm, b) 20% jittered beats from the same second and c) the two amplitude envelopes from a (blue) and b (orange) overlaid to show the temporal difference between the jittered beat and the isochronous rhythm.

By asking adult participants to rate multiple short periods of each of these audio beats as either 0%, 5%, 10% or 20% jittered, pilot testing showed that jittered beats at 5% and

10% jitter were indistinguishable from isochronous, 0% jittered beats. Jittered beats at 20% jitter were found to be distinguishable from the others.

4.2.3 Procedure, EEG acquisition and EEG preprocessing

The procedure, EEG acquisition and EEG preprocessing were kept consistent with Chapter 1 the only difference was the conditions of the audio being presented.

Data collection occurred when an infant fell asleep within a lab testing session. In most instances, the light in the room was dimmed and a tranquil ambiance was maintained to promote the-child's sleep. No auditory sleep aids were allowed. Creative SBS 250 speakers were calibrated to 65dB close to the participants' ears, using a RS PRO RS-95 sonometer. 65Db was chosen to be close to the average amplitude of standard speech (Olsen, 1998). Adult participants were directed to rest and listen to the passive audio task. Adults were allowed to sleep during the task. Participants were asked to abstain from rhythmic movements including bouncing, waving, fidgeting, talking or chewing and asked not to use their phones or eat during the procedure. Participants could stop the session at any time.

4.2.4 EEG data acquisition

EEG signals were obtained using a dual BioSemi (Amsterdam, NL) ActiveTwo system configured for 64 channel recording from both participants simultaneously. Participants wore size appropriate 64 channel Electro-Cap International (Ohio, US) caps with a 10-20 electrode montage. EEG signals were recorded at 512Hz with no online filtering using ActiView data acquisition software (version 7.07; BioSemi). Signa Gel conductive electrode gel from Parker Laboratories BV (Almelo, NL) was used to bridge the connection between the electrodes and the participant’s scalp. Common Mode Sense and Driven Right Leg electrodes between Pz and POz were used as the active reference.

4.2.5 EEG artifact rejection and preprocessing

The EEG preprocessing and artifact rejection was similar to Chapter 1 using an 8 step process: first data were high pass filtered at 1Hz; second line noise was removed at 50Hz using a notch filter; third data were low pass filtered at 25Hz; fourth data were referenced to a robust average, calculated by first temporarily removing noisy channels using default settings of *clean_line.m*, averaging the remaining channels and removing this robust average from all channels; fifth after robust averaging all channels, channels that were subsequently still noisy were rejected with a correlation threshold of 0.7 and a noise threshold of 4, all remaining channels were checked for bridging with *eBridge.m,* all noisy or bridged electrodes were interpolated using the spherical method of *eeg_interp.m*; sixth, using a sliding 1 second window without overlap, epochs were rejected and zeroed out if 70% of the channels exceeded -3.5 to 5 standard deviations of a robust estimate of channel EEG power; seventh conditions were concatenated and an ICA was completed with *run_ica.m*; eighth automatic rejection of bad components was completed for adult participants using *iclabel.m* and for infant participants using *iMARA.m* (Marriott-Haresign et al., 2021). Finally, data was resplit into their constituent conditions.

EEG data with more than 75% channels interpolated were rejected. Participants with 25-75% of channels interpolated had the interpolated channels removed and the remaining channels were used as a partial EEG dataset.

Individual and group level data showing the average percentage of channels interpolated vs removed, as well as the average percentage of segments zeroed out is shown in supplementary materials (SM C.3). SM C.3 also shows the total number of electrodes and seconds of EEG data included within the current analysis. In the rest, 0%, 5%, 10% and 20% jitter conditions, adult participants contributed 104.33, 160.13, 118.67, 99.83, and 98.73 minutes artifact free EEG data, while infants contributed 111.55, 146.27, 100.87, 86.67, and 81.13 minutes of artifact free EEG data.

4.2.6 Data analysis plan

The first analysis aimed to investigate whether infants neurally entrain to low frequency amplitude modulated rhythms that had been jittered from their original rhythm and whether the strength of those entrainments diminished as more jitter was introduced. To test this a frequency tagging analysis (see Chapter 1) was employed across power and phase measures for rest, 0%, 5%, 10% and 20% jittered conditions. A signal to noise correction was applied and differences between conditions were calculated using an ANOVAN.

The second analysis aimed to investigate whether participants were responding to the individual beats in an evoked contingent response compared to the rhythm, which would represent an endogenous oscillatory entrainment response. To do this an inter trial coherence (ITC) analysis was performed across power and phase domains to investigate the power and phase coherence across trials to the beat or to the overall rhythm.

Considering the findings of Chapter 1, the full 4-minute testing condition was used and whole head recordings were employed.

4.2.4.1 Analysis 1 - Is entrainment seen at the isochronous rhythm?

Our first aim was simply to investigate whether infants and adults were shown to have an entrainment response to a rhythm that, in the jittered conditions, was not directly supported by the peaks of isochronous beats. To address this for each participant x electrode x conditions timeseries (trial) a similar analysis to the first analysis in Chapter 1 was conducted. Recordings that were longer than 122880 samples (512Hz x 240 seconds) were cropped to this size and recordings that were shorter were zero padded to the required length.

FFT analysis was conducted with *myFFT.m* MATLAB script (Schoof, 2017) between 6-10Hz in 0.01Hz frequencies giving 401 frequency bins.

PLV (Lachaux et al., 1999) was calculated over each trial with the same frequency range and resolution as the FFT analysis. Each frequency bin was compared to a computer-generated perfect sinusoid signal of the same temporal length as the trial. A phase angle timeseries was calculated for each sinusoid vs the EEG trial by taking the phase angle of the Hilbert transform of the data. The difference between the neural and sinusoidal phase angle timeseries was taken, multiplied by the imaginary operator *i* and an exponential was determined for each value. The averaged exponential value is used as the PLV value. A PLV value was created for each frequency bin for each trial, before being averaged across trials for each frequency bin.

A signal to noise ratio (SNR) calculation was then completed on all power (FFT) and phase (PLV) analysis results to standardise and detrend the 1/f nature of the data (Vettori et al., 2019). Using a moving window of 25 frequency bins, SNR scores were calculated by averaging 20 bins either side of the target frequency, excluding the two closest bins and the two bins with the maximum absolute difference from the target. Two further analyses were then completed using these SNR scores.

4.2.4.1.1 Analysis 1.1

Analysis 1.1 used a bootstrapped permutation analysis to investigate whether the SNR value at the stimulation frequency was stronger than the surrounding frequency bins within the 6-10Hz range. The value at 8Hz was taken as the test value, and bootstrapping with replacement was completed to create 10,000 surrogate datasets per participant and power/phase analysis. With each surrogate dataset the original test value was ranked against the 10,000 values at the 8Hz frequency in the surrogate bootstrapped datasets and considered significant if the original value fell in the top 5% of ranks.

4.2.4.1.2 Analysis 1.2

Analysis 1.2 used an N-ways ANOVA (ANOVAN) to compare the SNR scores at the stimulation frequency for each participant x electrode x condition dataset to see if there were differences in the stimulation response at the 8Hz rhythm with differing levels of jitter. An ANOVAN was chosen due to the uneven participant numbers per condition. Multiple comparisons were completed with an ANOVAN showed significant results.

As the jitter ranges are relatively temporally small and cover a flat distribution across those ranges, all jittered beats necessarily overlap with the central point where the isochronous rhythm would be, causing there to be an 8Hz frequency response even when not directly supported by the majority of the beat peaks (figure 4.3). This leads to the possibility that contingent responses to the beat would overlap in a similar manner, which could create a significant peak at the isochronous rhythm.

Therefore, if the neural response to the jittered audio shows a peak at 8Hz with significantly diminishing amplitude and increasing surrounding noise as jitter increases then this could be due to a reduction in endogenous oscillatory activity with increased difficulty or contingent responding to the jittered beats. However, if there is a relatively stable neural entrainment at the 8Hz isochronous rhythm across jitter conditions then this could only be due to an endogenous oscillatory response.

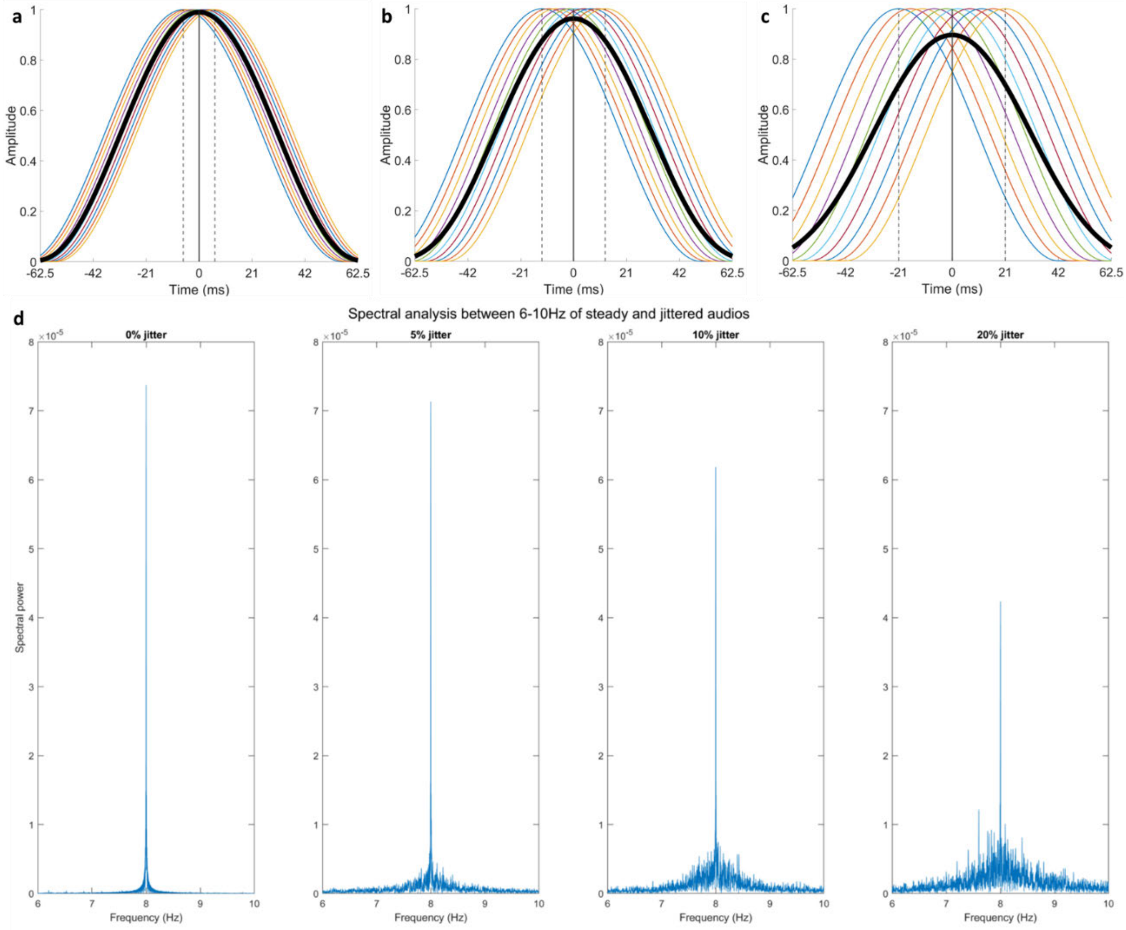


Figure 4.3a-c demonstrate the overlapping nature of the jittered beats at a) 5% b) 10% and c) 20% jitter. For each plot the solid line at 0ms represents where the isochronous rhythm would be, while the dashed lines represent the limits of where jittered peaks can be. Ten example beats are shown for each condition and the average of these beats is shown with a thick black line, which demonstrates the diminishing peak of the average jittered rhythms the jittered beats become more spread out. Plot d shows spectral plots between 6-10Hz of recorded audio for each of the audio conditions using the same equipment used in the procedure. As the amount of jitter increases the 8Hz peak in the audio recordings decreases and the amount of noise surrounding 8Hz increases.

4.2.4.2 Analysis 2 – Neural power and phase differences between the beat and rhythm

While a continually high power and phase response to the 8Hz frequency bin would show endogenous oscillatory entrainment, a diminishing response seen at 8Hz may be the result of a contingent response structure or a diminishing oscillatory entrainment response with increased difficulty.

Therefore, to investigate whether any response at the 8Hz frequency rhythm was a result of a convergence of contingent responses to individual beats or whether there was a stronger response at the isochronous rhythm, an ITC analysis across power and phase domains was conducted to compare the EEG signal at the beat or rhythm.

For the 5%, 10% and 20% jitter conditions, beat and rhythm trials were created by epoching the data 62.5ms before and after a peak to avoid stimulus overlap. The rest and 0% conditions were excluded, the rest condition because there is no stimulus with which to epoch the data, and the 0% jitter condition because the beat and the rhythm occur simultaneously, so there would necessarily be no difference between conditions. For the beat trials, the data were epoched relative to the stimulus beat peak. Rhythm trials were created by epoching the data centred around the timepoint where the isochronous rhythm would have been for each trial.

Inter trial power and phase coherence were calculated for both beat and rhythm trials. Time-frequency power and phase were extracted via complex Morlet wavelet convolution. The wavelets increased from 6-10Hz in 0.1Hz linearly spaced steps and the number of cycles increased from 3-10 logarithmically.

For the inter trial phase coherence the phase angle of this waveform was taken. For the inter trial power coherence the square of the absolute value was calculated. For both power and phase analyses, each of these results was multiplied by the imaginary operator *i* and the exponential was taken. The absolute value of the mean of the exponential over the trials was used as the inter trial coherence value for that channel x frequency x time trial.

As ITC investigates the similarity of phase alignment and power responses across trials, if there is a strong power and/or phase alignment to one condition, it necessarily can't also be as strongly aligned with the other condition. This is because relative to one condition the other would seem spread out, for example compared to the isochronous rhythm, the stimulus beats are spread out, and similarly when epoching centred around the stimulus the isochronous rhythm would appear relatively spread out. However, if there is no strong neural tracking of either condition, this would demonstrate no preference for entrainment to either the beat or the rhythm.

4.2.4.2.1 Analysis 2.1

First, to test whether the power and phase alignment results in the ITC were above chance, a control condition was produced by randomly generating surrogate stimulus timepoints, epoching the EEG data and calculating power and phase ITC as above. This was conducted over 100 permutations and the results were averaged over the permutations to give a control dataset. This was completed for each participant group x condition.

As the audio stimuli were continuous, and we expect the EEG responses to also be continuous, each trial in the control condition should contain at least one stimulus beat. Therefore, the control condition should represent EEG responses to a maximally distributed set of individual beats across the entire trial. By averaging across trials this should give the base level of neural activity for each condition.

To test the difference between real and control data, paired sample t-tests were used with a Benjamini-Yekutieli correction for false discovery rate (Benjamini & Yekutieli, 2001).

4.2.4.2.2 Analysis 2.2

Next, statistical differences between neural entrainment in the beat and rhythm conditions were investigated. Paired sample t-tests were conducted between the EEG signal of the beat and rhythm EEG epochs across each participant group x condition x power/phase ITC analyses. A Benjamini-Yekutieli correction for false discovery rate was used to correct for multiple comparisons (Benjamini & Yekutieli, 2001).

4.2.4.2.3 Analysis 2.3

Finally, to test whether there were significant differences in neural tracking within the beat and rhythm conditions, an ANOVAN was used to compare the ITC power and phase responses to each condition x participant group. Multiple comparisons using *multicompare.m* were calculated when significant differences were shown.

4.3 Results

4.3.1 Analysis 1 – Is entrainment seen at the isochronous rhythm?

4.3.1.1 Analysis 1.1

To investigate whether there was significant frequency tagging to both steady and jittered low amplitude modulated stimulus conditions at 8Hz compared to surrounding frequency bins, a bootstrapping analysis derived from SNRs of the FFT and PLV analyses was conducted. Table 4.1 shows the SNR value across the FFT and PLV analyses, for each condition and participant group. Highlighted in light and medium blue are conditions with p values less than 0.05 and 0.01 respectively.

The results showed that there was no significant frequency tagging response for any of the rest conditions vs other frequency bins for FFT and PLV analyses, for both infants and adults ($p > 0.05$). There was a significant ($p < 0.01$, except infant and adult PLV to

20% jitter condition, $p < 0.05$) frequency tagging response at the 8Hz frequency bin compared to surrounding frequency bins for all other condition.

*Table 4.1. Showing the average SNR value for adult and infant participants for control, steady and jittered frequency tagging conditions, medium blue and † denotes $p < 0.01$ and light blue and * denotes $p < 0.05$.*

	Infant		**Adult**	
	FFT	**PLV**	**FFT**	**PLV**
Rest	1.05	1.06	1.05	1.04
8Hz, 0% jitter	1.20†	1.20†	1.23†	1.31†
8Hz, 5% jitter	1.15†	1.19†	1.26†	1.27†
8Hz, 10% jitter	1.16†	1.18†	1.19†	1.17†
8Hz, 20% jitter	1.16†	1.14*	1.13†	1.11*

4.3.1.2 Analysis 1.2

Next an analysis was conducted to determine whether the neural tracking at the 8Hz frequency was consistent across jitter conditions, which would suggest endogenous oscillatory entrainment, or whether the neural tracking at 8Hz diminished across jitter conditions, which would suggest either contingent responses to the individual beats or a decrease in entrainment due to the difficulty of tracking as jitter increased.

To test for differences between conditions an ANOVAN was conducted using the values at the 8Hz rhythm frequency bin. Results showed that there were significant differences between conditions for infant FFT ($F(4, 63) = [11.42]$, $p < 0.001$), infant PLV ($F(4, 63) = [9.04]$, $p < 0.001$), adult FFT ($F(4, 63) = [39.37]$, $p < 0.001$) and adult PLV ($F(4, 63) =$

[67.69], $p < 0.001$) analyses. There were no significant differences found between electrodes for all analyses ($p > 0.05$).

Multiple comparisons were completed for each analysis type and participant group, the difference in ANOVAN means are shown in figure 4.4. For both adult and infant participants, the rest no sound condition was found to have a significantly lower frequency tagging response when compared to all other experimental conditions. These differences were highly significant ($p < 0.001$) for all conditions except rest vs 20% jittered condition in the PLV analysis ($p < 0.05$) for both participant groups. For infant participants, no further significant differences were found between groups ($p > 0.05$).

For adult participants in the power (FFT) analysis, the 0% jitter condition was found to have a significantly larger response than the 20% jitter conditions ($p < 0.001$); the 5% jitter condition had a significantly larger response than the 10% ($p < 0.01$) and 20% ($p < 0.001$) jitter conditions and there was a non-significant ($p = 0.053$) difference between 10% and 20% jittered conditions. All other comparisons were not significantly different ($p > 0.05$).

For the adult participants in the phase (PLV) analysis, there was a significantly larger response at 0% and 5% jittered when compared to 10% and 20% jittered conditions ($p < 0.001$). There were no significant differences between the 0% and 5% conditions ($p > 0.05$) or the 10% and 20% conditions ($p = 0.06$).

Overall, the results show that there was a diminishing response in both power and phase of neural tracking seen across conditions in the adult participant group suggesting either a decrease in oscillatory neural entrainment as the amount of jitter increased or that the neural signal showed adult participants were contingently responding to the stimuli. However, this pattern was not seen for infant participants in either the power or phase domain, who were shown to have no significant differences between conditions, which may represent endogenous oscillatory neural entrainment to the rhythm.

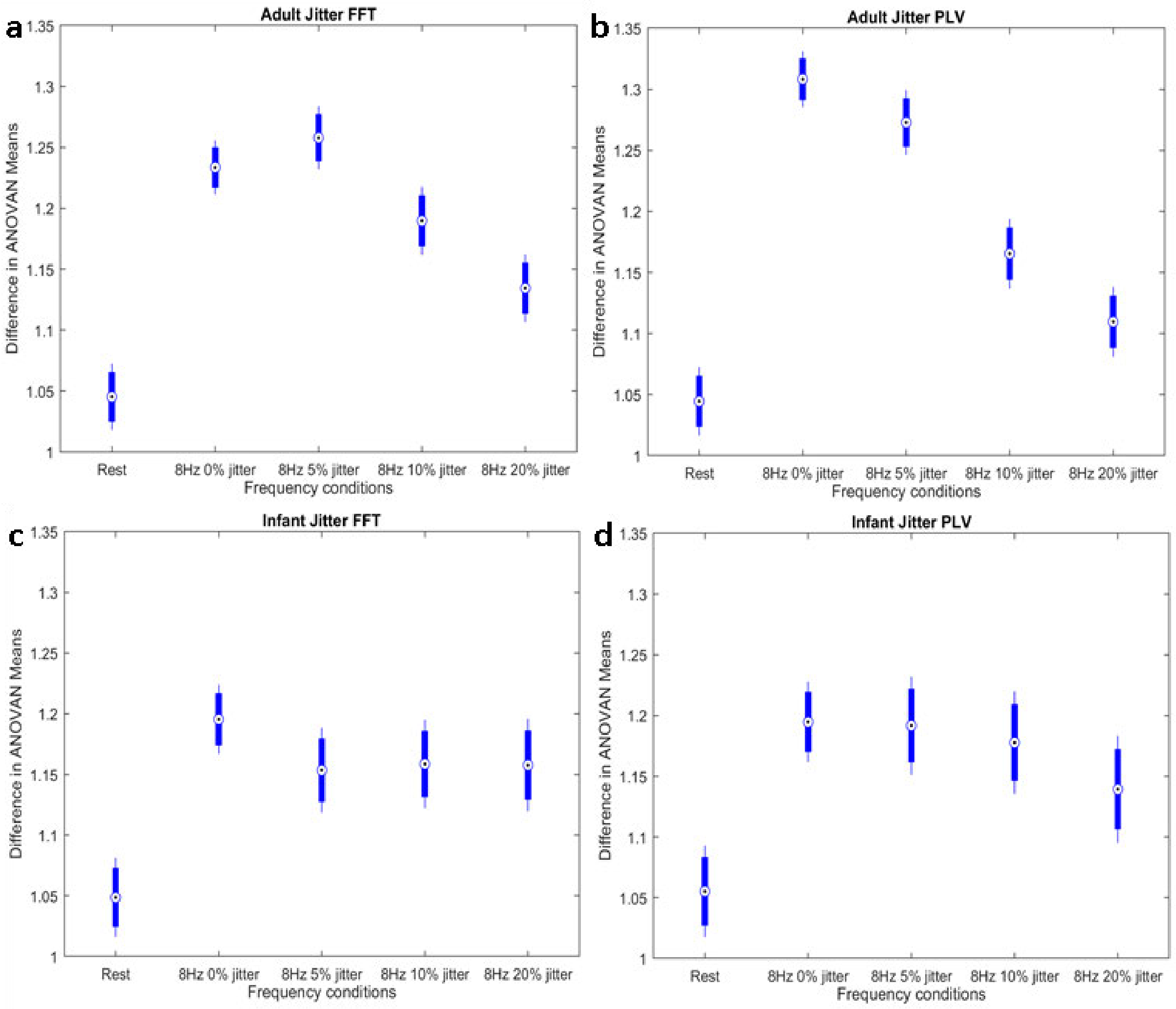


Figure 4.4 ANOVAN multiple comparisons using multcompare.m. Infant (a, b) and adult (c, d) power (a, b) and phase (c, d) analyses are shown for the target conditions rest (no sound), 8Hz 0% jitter (i.e., steady beats), 5%, 10% and 20% jitter. Data averaged over the whole head.

4.3.2 Analysis 2 – Neural responses to the beat vs the rhythm

To investigate whether significant differences in adult responses were associated more to the stimulus beat rather than the isochronous rhythm and to further investigate whether the infant responses were a result of oscillatory entrainment to the isochronous rhythm, inter trial power and phase coherence (ITC) analyses were run.

4.3.2.1 Analysis 2.1 – Inter trial power coherence vs control

First to test the assumption that all data showed a response significantly higher than chance, each experimental condition was compared to a permuted surrogate dataset using pairwise t-tests, p values have been corrected using a Benjamini-Yekutieli false discovery rate correction for multiple comparisons (Benjamini & Yekutieli, 2001).

When comparing the power and phase alignment to the stimulus beats compared to the permuted control condition, all experimental conditions showed significantly more power and phase alignment than their respective control conditions: infant power at 5% jitter ($t(1367) = 18.64$, $p < 0.001$); infant power at 10% jitter ($t(1265) = 15.98$, $p < 0.001$); infant power at 20% jitter ($t(1162) = 18.49$, $p < 0.001$); infant phase at 5% jitter ($t(1367) = 6.32$, $p < 0.001$); infant phase at 10% jitter ($t(1265) = 7.03$, $p < 0.001$); infant phase at 20% jitter ($t(1162) = 7.44$, $p < 0.001$); adult power at 5% jitter ($t(1977) = 28.30$, $p < 0.001$); adult power at 10% jitter ($t(1702) = 28.25$, $p < 0.001$); adult power at 20% jitter ($t(1703) = 31.54$, $p < 0.001$); adult phase at 5% jitter ($t(1977) = 13.96$, $p < 0.001$); adult phase at 10% jitter ($t(1702) = 4.73$, $p < 0.001$); adult phase at 20% jitter ($t(1702) = 5.20$, $p < 0.001$).

When comparing the power and phase alignment to the isochronous rhythm compared to the permuted control condition, all experimental conditions also showed significantly more power and phase alignment than their respective control conditions: infant power at 5% jitter ($t(1367) = 18.54$, $p < 0.001$); infant power at 10% jitter ($t(1265) = 15.92$, $p < 0.001$); infant power at 20% jitter ($t(1162) = 18.49$, $p < 0.001$); infant phase at 5% jitter ($t(1367) = 6.07$, $p < 0.001$); infant phase at 10% jitter ($t(1265) = 7.25$, $p < 0.001$); infant phase at 20% jitter (t(1162) = 5.33, $p < 0.001$); adult power at 5% jitter (t(1977) = 28.28, $p < 0.001$); adult power at 10% jitter (t(1702) = 28.23, $p < 0.001$); adult power at 20% jitter (t(1703) = 31.55, $p < 0.001$); adult phase at 5% jitter (t(1977) = 15.33, $p < 0.001$); adult phase at 10% jitter (t(1702) = 4.55, $p < 0.001$); adult phase at 20% jitter (t(1702) = 3.12, $p < 0.01$).

Overall, the results demonstrate that as expected neural tracking to responses at the stimulus beat or the rhythm is above chance for all conditions and participant groups.

4.3.2.2 Analysis 2.2

Next to examine whether participants showed a significantly stronger neural tracking to the stimulus beat vs the isochronous rhythm, pairwise t-tests were calculated between beat and rhythm conditions for each jitter condition and participant group, p values have been corrected using a Benjamini-Yekutieli false discovery rate correction for multiple comparisons (Benjamini & Yekutieli, 2001).

Infants showed a significant tracking of the beat over the rhythm in the phase analysis at 20% jitter (t(1162) = 4.86, $p < 0.001$) and adults showed a significant neural tracking of the rhythm over the beat at in the phase analysis at 5% jitter (t(1977) = -3.20, $p < 0.05$). All other comparisons were found to be non-significant ($p > 0.05$).

This demonstrates that while there was significant power and phase alignment compared to chance (analysis 2.1), there were only a few cases of significant neural entrainment to either the beat or the rhythm when comparing between the conditions. This may suggest that there is either not enough neural tracking of either the stimulus or the rhythm specifically to elicit a significant difference, or that there are multiple strategies being employed which allow the brain to track both the beat and the rhythm simultaneously. One potential explanation could be that separate areas of the brain track different aspects of the auditory stimuli.

4.3.2.3 Analysis 2.3

Finally, to investigate whether there were significant differences between conditions and to examine whether there were topographical differences across the scalp, an ANOVAN was calculated for the beat and rhythm conditions and for power and phase analyses separately.

There were significant differences found between conditions for each ANOVAN, inter trial power coherence to the stimulus beat ($F(5, 63) = 67.21$, $p < 0.001$), and isochronous rhythm ($F(5, 63) = 67.25$, $p < 0.001$), inter trial phase coherence to the stimulus beat ($F(5,63) = 9.16$, $p < 0.001$) and isochronous rhythm ($F(5,63) = 16.62$, $p < 0.001$). There were no significant differences found between electrodes for any of the ANOVAN analyses calculated ($p > 0.05$). Figure 4.5 shows the differences between ANOVAN means for each of the ANOVAN analyses.

Interestingly, inter trial power coherence to the beat stimulus and isochronous rhythms showed different individual means but the same pattern of significant interactions. Multiple comparisons revealed that the infant 5% jitter condition showed significantly less power amplitude alignment than any other condition ($p < 0.001$), while infant 10% jitter showed significantly more power alignment than all other conditions ($p < 0.001$; except when compared to adult 10% jitter, $p < 0.05$). Adult 5% jitter was found to have less power than either the adult 10% ($p < 0.01$) or 20% jitter ($p < 0.05$) conditions. All other interactions between conditions were not significant ($p > 0.05$).

Inter trial phase coherence to the beat stimulus showed no significant differences between infant jitter conditions ($p > 0.05$). Between adult jitter conditions the 5% jitter condition had significantly more phase alignment than the 10% and 20% jitter conditions ($p < 0.001$). Between participant groups, the infant 20% jitter condition had more phase alignment than the adult 10% jitter condition ($p < 0.05$) and the adult 5% jitter condition had significantly more phase alignment than the infant 5% jitter condition. All other interactions were not significantly different ($p > 0.05$).

Inter trial phase coherence to the isochronous rhythm also showed no significant differences between infant conditions ($p > 0.05$). Adult jitter conditions showed that the 5% jitter condition was had significantly more phase alignment than the adult 10% and 20% jitter conditions. Finally in cross participant interactions the adult 5% jitter condition was shown to have higher phase alignment than the infant 5% ($p < 0.001$), 10% ($p < 0.01$) and 20% ($p < 0.001$) conditions and the infant 10% jitter condition was found to have more phase alignment than the adult 10% ($p < 0.05$) and 20% jitter ($p < 0.01$) conditions. All other interactions were not significant ($p > 0.05$).

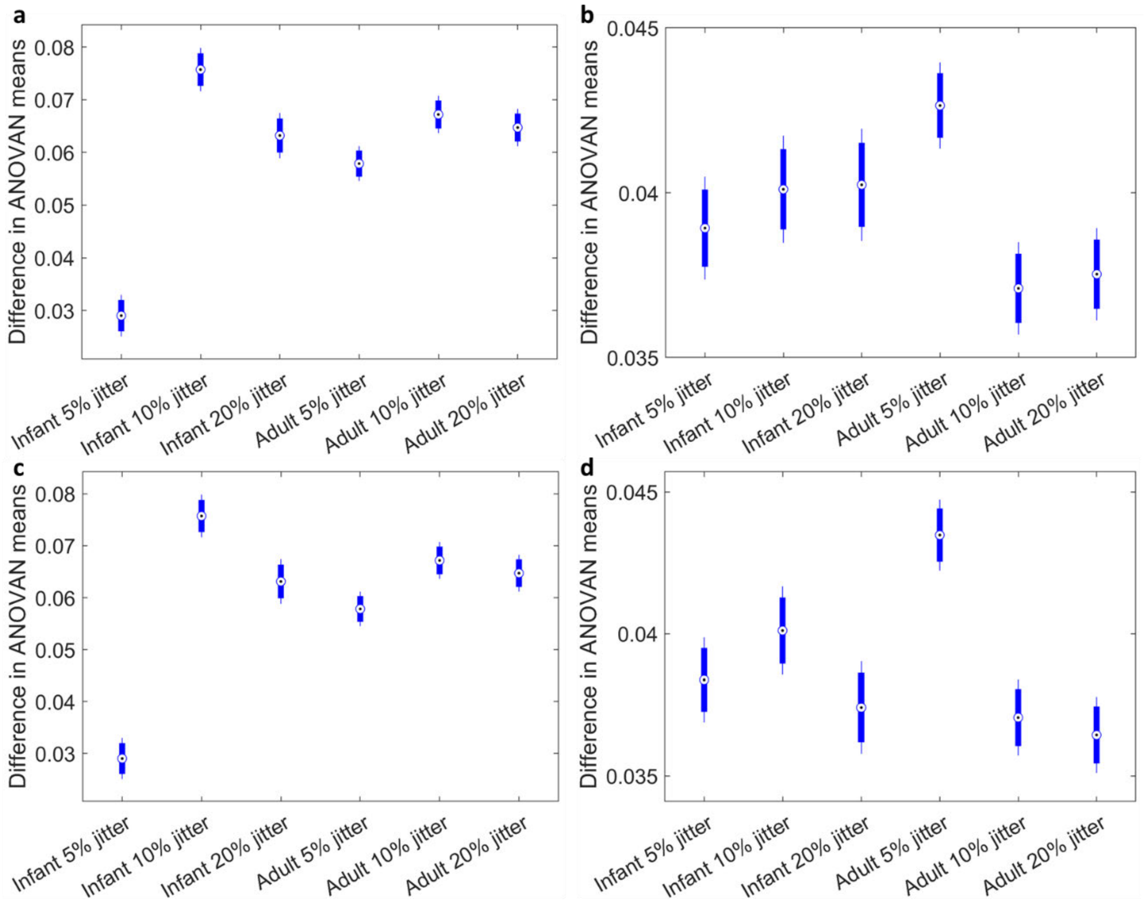


Figure 4.5. Showing multiple comparisons for infant and adult participants across the 5%, 10 and 20% jitter for a) power amplitude to the stimulus beat, b) power amplitude response to the isochronous rhythm, c) phase alignment to the stimulus beat and d) phase alignment to the isochronous rhythm. Data averaged over the whole head.

4.4. Discussion

4.4.1 Overview

In this study we examined how entrainment to low amplitude modulated rhythms in adult and infant (4-6 month old) participants are impacted by increasing irregularity in rhythmic stimuli. We investigated whether neural entrainment to jittered stimuli showed patterns of contingent responding to the stimulus beats or whether there is evidence of neural entrainment. Differing patterns of activity are seen in adult and infant populations. To the best of our knowledge, no previous research has used these methods to investigate the mechanisms of entrainment.

4.4.2 Frequency tagging analyses

Analysis 1.1 compared frequency bins at the target frequency of 8Hz to surrounding frequency bins between 6-10Hz, for conditions with increasing levels of auditory jitter. The results showed that there was a significant response at the target 8Hz frequency compared to surrounding frequency bins for all experimental conditions and no neural response to the 8Hz frequency bin in the control, no sound, condition. This is in contrast to previous results, which both showed no steady state responses at the target rhythm. This may be due to the level of jitter and jitter administration in each of the previous studies: 20% jitter in only 9 equally spaced bins (Tavano et al., 2022) and 55-65% jitter (Lehmann, Arias & Schönwiesner, 2016). Also, the stimuli used were very short, allowing for minimal overlap with the target rhythm rate, which may have contributed.

Analysis 1.2 examining the neural response at 8Hz across conditions, showed that there were no significant differences seen between infant experimental conditions. However, there was a significant uneven decreasing amplitude across conditions as jitter increased for adult participants.

If responses to the auditory stimuli were driven solely by contingent responding, we would expect that the neural response would decrease linearly with the amount of jitter present in the audio. This would match the linear decrease in activity at the target stimuli seen in the audio itself (see figure 4.3). However, if neural responses were largely driven by an endogenous oscillatory entrainment mechanism, we would expect participants to be able to maintain a steady neural response to the isochronous rhythm of the audio across conditions. As infants showed a steady neural response at 8Hz to the auditory stimuli, it could be argued that this represents oscillatory neural entrainment to the stimuli.

An alternative explanation for the infant activity could be that they are showing general neural activity due to the sound without entraining to the stimuli. However, as frequency of the neural activity is very specific to 8Hz, it suggests that there isn't general noisy neural activity being found. Also, previous research (e.g. Chapter 1) has demonstrated that similar infants, in the same conditions, are capable of entraining to a range of 0% jitter low amplitude modulated stimuli. As this study found that there was no significant difference between the neural entrainment response at 0% jitter compared to other conditions it can be assumed that there was a similar level of entrainment throughout.

A further explanation of these results could be that infants have a higher tolerance for jittered stimuli compared to adults. This may be because infants experience the world in a less temporally precise way due to cortical immaturity and or their inexperience with processing complex stimuli. This may suggest that infants must rely more heavily on internally generated oscillators related to the average rhythms around them rather updating their internal representations of complex rhythms on a beat-by-beat basis. Due to this, it may be that as infants age their neural temporal precision increases as a function of increased myelination and cortical development as well as with more experience with enhanced sensory perception and information integration.

Results from the adult participants showed a general decrease in neural entrainment as the level of jitter increased. However, it was also shown that decreases between conditions were not consistent, which would have been expected from a purely linear decrease driven by contingent responding alone. interestingly in both power and phase domains there was no significant differences between 0% vs 5% or 10% vs 20% jitter conditions, although 10% vs 20% conditions in both power and phase domains showed a trend towards significance. It would have been expected that there was a larger not smaller decrease in neural amplitude between 10% vs 20% when compared with 5% vs 10%. It is also interesting that the largest amplitude decrease in consecutive conditions was between 5% and 10% conditions that were both rated as indistinguishable from isochronous rhythms, rather than between 10% and 20% conditions.

While the adult results do not point to a purely contingent responding mechanism, the results could suggest a mixture of contingent responding and oscillatory entrainment to the jittered audio. The results could also suggest that there are a mix of responses seen either: between participants, between levels of jitter or across brain regions.

Finally, different neural entrainment responses across conditions between the participant groups point to a general developmental difference in neural entrainment. Infants may be presenting an increased sensitivity to temporally irregular stimuli, similar to early entrainment sensitivities similar to those reported by Hannon and Trehub (2005) and Cantiani et al., (2022).

4.4.3 Inter trial power and phase coherence analyses

Results from analysis 2 further investigated neural responses between conditions, spatial locations and attempted to tease apart responses to the stimulus beats vs the isochronous rhythm. An intertrial power and phase coherence analysis was used to investigate the coherence of power and phase alignment responses over trials centred around the stimuli and the isochronous rhythm. Analysis 2.1 demonstrated that both the stimulus beat, and isochronous rhythm conditions had a significantly larger response than chance. Analysis 2.2 compared beat and rhythm conditions and showed in the phase domain that infants in the 20% jitter condition had a stronger phase alignment to the beat over the rhythm, while adults showed the opposite response in the 5% jitter condition.

Overall, these results showed that there are limited differences between the neural entrainment responses to the beat and isochronous rhythm. There is a small amount of evidence for different mechanisms being used to entrain to the beat and rhythm, with oscillatory entrainment to the isochronous rhythm seen with adults, while infants showed entrainment to the stimulus beat. However, as this crosses not only conditions but also participant groups, no conclusive generalisations can be drawn.

Given the significantly above chance power and phase alignment seen in analysis 2.1, we were expecting that there would be some activity centred around the central timepoint for both beat and rhythm conditions in analysis 2.2. However, if there was a strong bias towards contingent responding or oscillatory entrainment, we would have expected a stronger phase alignment towards that preference. By showing that there were limited differences between segmented neural responses to the stimulus beat and isochronous rhythm it suggests that there are a mixture of mechanisms being employed when entraining to this jittered audio.

Finally, analysis 2.3 investigated differences between conditions in the beat and rhythm conditions, as well as investigating spatial differences across the scalp. The power domain demonstrated similar responses to the beat and rhythm conditions, infant neural responses in the 5% jitter condition were lower than all other conditions, while responses to the infant 10% jitter condition was higher than all other conditions. Adult participants showed a lower response to the 5% jitter condition when compared to the adult 10% and 20% jitter conditions. The phase domain showed neural phase alignment that was largely the same as analysis 1.2, with no significant differences between infant conditions, and adult conditions showing a larger phase alignment response to the 5% jitter than all other conditions. No differences between electrodes were seen.

As there were no topographical differences shown between the conditions, this suggests that the mixed results seen in the above analyses are not a result of different spatial regions each tracking different aspects of the auditory stimuli.

It was not surprising to see that the phase results were similar to analysis 1.2, as the inter trial phase coherence is quite similar to PLV, the main differences being that ITC investigates across epoched trials, while PLV looks at the entire continuous signal. Whereas dissimilarities between ITC power and FFT results in analysis 1.2 were expected as the FFT to the entire signal shows spectral power, while ITC of power shows the alignment of power amplitude responses over trials, which is a fundamentally different response.

It is interesting that other adult and infant results showed a low power amplitude alignment in the 5% jitter condition, and a high power amplitude alignment particularly in the 10% jitter condition but also the 20% jitter condition. It is unclear why the power amplitude alignment in the 5% jitter condition would be significantly lower than other conditions, especially as the clustering of beat stimuli would be the most concentrated of the three conditions. Further study is required to determine whether this is anomalous or a fundamental response.

4.4.4 Limitations and future directions

One potential limitation is that this study did not investigate individual differences between participants to examine whether there were groups that tracked either the beat or the rhythm. If present, these groups may potentially be derived from individual differences in neuroanatomy or brain processing. Adult participants could have differed in the level of attention or inhibition shown towards the jittered signal, either to actively pay attention to it or to ignore it passively or actively. Similarly, the sleep stage that infants were in was not categorised, although the procedure generally started 10 minutes after the infant fell asleep. Individual differences were not explored due to the sample size, with more participants the groups could have been further subdivided. Future studies should explore a larger participant pool with the goal of investigating whether there are individual differences and whether there are particular traits that could characterise them.

Results from analysis 2.3 demonstrate that alignment of power amplitude responses was not a sensitive measure to differences between the beat and rhythm conditions, this is likely because the average power amplitude would rise relative to the beat or rhythm leading to an average increase of power amplitude over time up to the centre point and then a decrease (similar to figure 4.3a-c). Therefore when the inter trial power coherence was calculated, there was likely a similar pattern of power amplitude increases and decreases seen.

Finally, future studies may wish to explore different stimulus frequencies. Recent studies (Chapter 1) have demonstrated that infants and adults show a particularly strong neural frequency tagging response at 2Hz and 4Hz. Using lower stimulus frequencies also gives a larger inter beat interval between beats to manipulate. However, future researchers may also need to increase the stimulus length to account for the drop in the number of beats.

4.4.5 Conclusion

Overall, while the results do not point to a conclusive pattern of entrainment to the stimulus beats or to the isochronous rhythm. The results do point to an interesting trend of activity that does not strictly follow either a contingent responding or oscillatory entrainment structure.

Chapter 5 – Do Infants Show More Neural Tracking To More Rhythmic Speech And Song?

The manuscript is currently in preparation for publication. The supplementary materials for this chapter are presented in Appendix D.

Overview

Chapter 5 of this thesis marks a significant departure from the controlled experimental settings of previous chapters, transitioning into a more naturalistic and ecologically valid context of infant social interaction over tabletop play. Along with changes in oscillation frequency (Chapter 3) and regularity (Chapter 4), infants must contend with entraining to stimuli in a range of environments and with subtle differences in the way in which their caregiver interacts with them.

Building upon the knowledge established in earlier chapters, this chapter explores the dynamic interplay between infants and caregivers in everyday interactions, with a specific focus on the role of rhythm in shaping infant speech brain tracking. As highlighted in Chapter 2, speech and infant-directed speech (IDS) are characterized by rhythmic properties that play a crucial role in early language acquisition and social interaction.

Leveraging this understanding, Chapter 5 investigates how variations in rhythmic patterns during caregiver-infant interactions influence infants' neural responses to speech stimuli. By investigating the naturally occurring changes in the rhythmic properties of caregiver speech and song, we can assess how variations in rhythm influence infants' neural responses. By examining the effects of rhythm on infant speech brain tracking, we aim to investigate the mechanisms underlying language tracking at varying ages with the hope of shedding more light on development and social communication in infancy.

Overall, Chapter 5 represents a crucial step towards understanding the role of rhythm in shaping infant speech perception. By investigating how variations in caregiver speech rhythm influence infants' neural responses, we aim to uncover the underlying mechanisms of language in early infancy. Through a combination of naturalistic observation and advanced neuroimaging methods, this chapter contributes valuable insights into the intricate interplay between rhythm, speech, and infant brain development.

Abstract

Infant directed (ID) speech and song are important in language acquisition, employing acoustic features that make speech perception easier for preverbal infants. ID vocalisations have a slower speech rate, which quickens as infants develop and gain experience. Neural studies show infants track prosodic, stress and syllabic frequencies in ID language. However, stimuli have predominantly been recorded, manipulated during or after recording, and are often repeated. This presents issues of ecological validity due to the quasi-chaotic nature of social interactions and natural rhythmic structures. Limited studies have investigated continuous neural mapping in response naturalistic dyadic interactions and to our knowledge no studies have investigated neural tracking of ID vocalisation rhythmicity across developmental time. This study investigated infant-caregiver dyads at 5 (N=58), 10 (N=29) and 15 (N=17) months old with naturalistic ID speech and song during free-flowing tabletop play interactions. Infant tracking of caregiver audio was analysed with continuous multivariate temporal response functions. Results showed above chance delta band speech tracking, and theta and alpha band song tracking. Split by age, older infants showed delta speech brain tracking (SBT) in the delta band, while song-brain tracking was shown in the youngest timepoint and again in the 15 month infants. This suggests different mechanisms of language tracking across development. However, no significant correlation between the amount or frequency of identified oscillatory audio activity vs SBT was found. The present study discusses the future directions of research investigating neural tracking of infant rhythm and the importance of further naturalistic infant SBT research.

Word count: 248

Keywords: infant speech brain tracking, neural entrainment, rhythmic audio

5.1 Introduction

Infant directed (ID) vocalisations hold a privileged place in infant language acquisition (for review see Golinkoff et al., 2015) that is seen cross-culturally (e.g. Broesch & Bryant, (2014); Tsang, Falk and Hessel (2016); Poeppel & Assaneo, 2020), across vocalisation type (e.g. speech vs song: Kalashnikova et al., 2018, compared to Nguyen et al., (preprint)) and throughout infant development (e.g. Narayan & McDermot, 2016).

However, until recently the majority of neural measures of ID vocalisations relied on recorded and repeated stimuli to measure cortical responses, which lacks ecological validity. Modern speech-brain tracking (SBT) techniques have been developed that allow for continuous mapping of cortical responses to incoming stimuli, yet these too have largely been used in response to recorded stimuli. The current study builds on limited recent dual EEG studies using free flowing dyadic play interactions to investigate infant neural tracking across developmental time and ID vocalisation type (speech and song) to compare how the rhythmicity of audio impacts infant neural tracking in the developing brain.

5.1.1 Infant directed speech

ID speech has been widely studied in the infant literature (for review see Golinkoff et al., 2015; Cox et al., 2023) and is characterised by *"fewer words per utterance, more repetitions and expansion, better articulation, and decreased structural complexity. Prosodic modifications include higher overall pitch, wider pitch excursions, more distinctive pitch contours, slower tempo, longer pauses, and increased emphatic stres*s" (Cooper and Aslin, 1990).

ID speech has been linked to the promotion of infant attention to language, helping to foster social interactions between infants and caregivers and demonstrates to infants the various important aspects of their native language by heightening salient aspects of speech (Golinkoff et al., 2015). ID speech modulations have also been shown cross-culturally. Broesch & Bryant (2014) demonstrated that there are similarities in ID speech acoustics across cultures, with higher pitch and slower speaking rate shown in western and “traditional” cultures such as those in Fiji and Kenya (see also Tsang, Falk and Hessel (2016) and Poeppel & Assaneo, 2020).

Infants have been shown repeatedly to prefer ID over AD speech from as early as 1 month (e.g. Cooper and Aslin, 1990), with higher neural activity seen to ID speech at 6 and 13 months when comparing ID and AD speech, even when comparing familiar and unfamiliar words (Zangl and Mills, 2007, see also Kalashnikova et al., 2018). This increased neural activity is also seen in neural tracking of ID speech, which is thought to benefit infant language perception and acquisition (e.g. Fló et al., 2022; Menn et al., 2022).

Research investigating ID speech oscillatory rhythms by Leong, et al., (2014; see also Goswami, 2017) showed that ID speech enhances temporal rhythmic structure of the auditory amplitude envelope, which helps the infant’s brain by supporting multi-scale neuronal oscillatory sampling of the speech signal. Leong et al., (2014) studied natural maternal ID speech to infants at 7, 9, 11 and 19 months old and were able to extract consistent stress, syllable, and phoneme rates. These demonstrated a stress dominated profile in ID compared to AD speech, with a high stress-syllable synchronisation compared to syllable-phoneme rates. The authors suggest that when speaking to infants, caregivers pattern their syllables to be more regularly in time with prosodic stress, while allowing phonemes to vary more.

Cross-cultural and longitudinal research supports the observations seen by Leong et al., (2014). ID speech rate has been shown to be slower across cultures, and in non-western languages, however tracked longitudinally the tempo was shown to increase as infants aged and became more proficient at language eventually being close to AD speech rates (Narayan & McDermot, 2016).

Taken together the above studies show that ID speech is a consistently present and important element of infant language development. Speech rate has been shown as a useful feature of ID speech, one that is adapted as infant language experience develops. Increasing tempo may reflect improvements in neural processing or increasing experience and proficiency with language stimuli. By comparing ID speech with acoustically similar ID vocalisations but different dominant tempos, such as ID speech and song, evidence can be gathered on whether infants directly entrain to different ID vocalisation rhythms.

5.1.2 Infant directed song

While ID song has received less attention than ID speech in the literature, research has shown: differences in acoustics and behaviour between ID and AD song are similar to differences in ID and AD speech; acoustic differences between ID speech and song; and differences between types of ID song (for a review see Chen-Hafteck, 2021; Markova, Nguyen & Hoehl, 2019; Trehub & Russo, 2020).

Infants have been shown to prefer ID singing over ID speech at 6-10 months in both native and non-native languages (Tsang, Falk and Hessel, 2016). Similar to differences between ID and AD speech, ID song has been shown to have greater pitch-amplitude coupling, a greater dynamic range but fewer expressive variations in timing than non-ID performances (Nakata & Trehub, 2011). Further research by Falk (2011), showed similarities melodic and intonational contours between ID speech and song across French, German and Russian nursery rhymes. However, there were differences between play and soothing contexts.

There are multiple types of ID song, two main groups are play songs and lullabies, which have been shown to have distinctive infant directed messages as rated by adult listeners (Rock, Trainor, & Addison, 1999). Infant play songs were shown to be more smiling and prominent consonants, be rated as more brilliant and be more rhythmic. Lullabies were rated as more soothing, smooth and airy (Rock, Trainor, & Addison, 1999).

Investigating the spectral properties of play songs vs lullabies when sung with ID vs non-ID singing, Trainor et al., (1997) demonstrated that similar to ID speech, ID versions of both types of song had slower tempo, more energy at lower frequencies, longer inter-phrase pauses and higher pitch and jitter factors. Between the ID song types pitch variability was higher and rhythm exaggerated in the ID version of play songs but not in lullabies.

As both ID speech and song have been shown to have many similar acoustic characteristics, have both been shown to capture infant attention and have both been shown to impact language acquisition, these ID vocalisation types offer an interesting opportunity to investigate neural tracking differences, such as level of rhythm in the vocalisations, based on the change of vocalisation type.

5.1.3 Measuring neural tracking of infant directed speech and song

Traditional methods of infant speech/song-brain tracking often require recorded audio, or video, manipulated stimulus characteristics and many repetitions to ensure repeatability of neural responses that can be averaged over trials. However, this is unlike anything that an infant is likely to encounter in their normal lives, which introduces potential confounds based on issues with habituation to the stimuli, lack of interest and confusion. These can all potentially lead to neural responses that do not truly reflect the neural responses to naturalistic stimuli.

A recently developed technique called multivariate temporal response function (mTRF) (for an infant mTRF methodological review Jessen, Obleser, & Tune, 2021; see also Crosse et al., 2016) uses continuous linear regression to map stimuli onto the neural responses or attempt to recreate the stimulus from the neural responses. There are two broad mTRF model types, those that use the original input signal to map the neural responses, known as “forwards models” or those that use the neural responses to recreate the stimuli are called “backwards models”.

There are limited studies that have used this technique with infants, however, early infant SBT studies using a forwards mTRF model assessed differences in cortical tracking of recorded ID and AD speech and demonstrated stronger low-frequency cortical tracking of ID speech envelope over AD speech envelope in 7 month olds (Kalashnikova et al., 2018). Similarly, investigating infant neural tracking of cartoons, Jessen et al., (2019) demonstrated in a forwards model that infants showed successful neural tracking of both the auditory stimuli and the level of movement in the cartoon, but showed no tracking of changes of luminance.

Using a different cortical tracking approach, Ortiz Barajas, Guevara & Gervain (2021) demonstrated in newborn and 6-month-old infants that there was a difference in amplitude (power) and phase tracking of recorded ID speech. Infants at both ages showed cortical tracking to the phase of the speech, but amplitude tracking was only seen just after birth. The authors suggested that this could represent two different neural mechanisms of oscillatory synchronisations which may play different roles in speech perception.

Attaheri et al., (2022) investigated delta and theta SBT to sung speech, recorded by an actor who listened to a 2Hz click track to improve the stability of the rhythm. The results of the backwards model showed that there was a greater modulation energy in ID than AD speech centred around the 2Hz frequency band, which the authors suggest is evidence of entrainment for language acquisition.

Taken together, the above studies begin to build a picture of infant SBT using continuous mapping methods. However, the above studies continued to use recorded and in one case manipulated recordings as their stimuli. In some cases, this removed aspects (such as visual inputs) that would be present in natural interactions, fails to reflect the quasi-chaotic nature of social interactions and removes any contingent responding from the caregiver within the interaction. A move towards more naturalistic, and free flowing SBT would allow researchers to demonstrate whether SBT is a phenomenon present when responding to controlled stimuli or a natural part of infant neural processing.

In a more naturalistic setting, Menn et al., 2022 using other infant SBT methods, gave caregivers novel objects to describe to their infants using either ID or AD speech. The researchers quantified the rhythms present in the audio and measured SBT to both the prosodic (1-2.5Hz) and syllabic (centred at 5.15Hz for AD speech and 5.5Hz for ID speech) rates of naturally occurring speech. The results demonstrated that 9-month-old infants showed more "speech-brain coherence" to both ID and AD speech at both the prosodic and syllabic rate, but that there was higher SBT to ID prosodic stress rate over syllabic rates.

Phillips et al., (in prep) demonstrated in a proof-of-concept study investigating methods of mTRF analysis with naturalistic ID speech during a tabletop play interaction that 9-12 month old infants showed delta modulation SBT similar to Attaheri et al., (2022). While, an investigation into naturalistic ID singing using mTRF analyses by Nguyen et al., (preprint) examined differences in neural tracking between play songs and lullabies showed above chance tracking of ID singing, with superior tracking of lullabies over play songs in 7 month old infants. The authors suggested that acoustic features of tempo and pitch of the lullabies compare to play songs modulated infant neural tracking.

The above studies demonstrate that continuous modelling approaches can show dynamic neural tracking of stimuli in both controlled and in a limited capacity in naturalistic settings to both ID speech and song. However, more research is required to investigate neural tracking during a wider range of infant development, examine further stimulus types such as a direct contrast between ID speech and song, and begin to test how naturally occurring acoustic and visual features of the infants' environment impact their neural tracking.

5.1.4 Current study

The current study aims to fill this gap by using a forwards model mTRF approach to measure SBT to naturalistic ID speech and song in infants at 5, 10 and 15 months. As shown in section 5.1.3 there are limited previous studies that have investigated continuous tracking of neural responses to stimuli. This is methodologically important during naturalistic studies as participants do not repeat themselves in a way that would allow the study to be segmented into neat trials (such as would be needed for ERPs for example). While other continuous analysis methods exist (see section 5.1.3), mTRF methods are emerging as a robust method of continuous naturalistic analysis with standardised mTRF Toolbox methods.

SBT in the delta, theta and alpha frequency bands will be examined to determine whether there are changes in SBT frequency band tracking across vocalisation type and age. As there has been limited research that have investigated speech brain tracking, this study will use the previously tested frequency bands from Attaheri et al., (2022) and Phillips et al., (under review), which consisted of delta (1-4Hz), theta (4-8Hz) and alpha (8-12Hz) to replicate previous findings and then build upon these with infant directed song.

This study will investigate whether any differences found between vocalisation type are driven by how rhythmic the vocalisations are by studying whether caregivers that display more rhythmic vocalisations induce more infant SBT. We predict, in line with previous findings (Attaheri et al., 2022; Philips et al., in prep) that SBT will be shown in the delta frequency band in the 10- and 15-month-old infants, with increasing responses seen. As above chance speech brain tracking has been shown in response to ID singing (Nguyen et al., preprint) and that ID song is generally slower in tempo than ID speech Trainor et al., (1997) we expect that infants will show song-brain tracking in the same delta band frequency band as ID speech. We predict that as above chance ID song has been shown in 7 month olds that delta band tracking will start earlier at 6 months and be present throughout development increasing in strength. We predict that the increased rhythmic nature of ID song will increase neural tracking, so ID song will have a stronger response than ID speech. Similarly, we expect that increased rhythmicity in general will increase neural tracking regardless of the play condition.

5.2. Methods

5.2.1 Participants

81 adult-infant participant dyads were recruited as part of a wider longitudinal project, which tests infant-caregiver dyads at 5-, 10-, 15- and 36-month timepoints. Data in this study is taken from the first 3 timepoints. A detailed breakdown of which participants completed each of the conditions after EEG preprocessing is shown in supplementary materials (SM5.1). Conditions were completed in a random order and testing was stopped if the child fussed too much or the caregiver asked to stop.

There were two play conditions: ID speech with 48, 22 and 10 participant dyads at visits 1-3 respectively; and ID song with 42, 23 and 14 participant dyads at visits 1-3 respectively. Average adult age at each of the three visits was 35.75, 35.26 and 36.08 years with a standard error of 0.46, 0.86 and 1.28 respectively. Average infant age at each visit was 5.58, 10.1 and 15.53 months with a standard error of 0.19, 0.28 and 0.21. All adult participants were female, infant participants across the ID speech and ID song conditions included 35, 17 and 8 male and 22, 12 and 10 female participants at each of the visits. Further participant details are included in supplementary materials (SM D.2).

The University of East London ethics committee approved the study. All adult participants provided informed consent for both themselves and their children according to the Declaration of Helsinki. All participants were offered a £10 shopping voucher as a monetary reward for their time. Travel and food expenses were also covered for those that requested them.

5.2.2 Procedure

Participants were seated facing one another across a table, infants were either sat in a highchair or on a researcher’s lap. Audio from the interactions was recorded using a

ZOOM H4n Pro Handy Recorder with lapel worn microphones and a Sennheiser EW 112P G4-R receiver. Triggers that were later used to synchronise EEG and audio signals (see section 2.4) were produced at the start and end of each recording using a custom written MATLAB script that sent a signal simultaneously to the EEG recording software and activated a 3000-4000Hz buzzer within range of the microphones for 1 second.

Participants completed the two play conditions examined in this study as part of a battery of play conditions, which were presented in a random order. In the ID speech condition, participants were given 3 toys to play with for a period of up to 10 minutes (for the 5, 10 and 15 month visit total time was 469.17, 226.26, and 178.59 minutes, with an average time of 5.52, 5.39 and 5.95 minutes respectively) and the caregivers were instructed to play as they would at home. In the ID song condition, caregivers were given a hand puppet and asked to sing nursery rhymes or other songs to the infant with the help of the puppet (for the 5, 10 and 15 month visit total time was 384.07, 199.99, and 142.44 minutes, with an average time of 4.86, 4.88 and 5.48 minutes respectively). During either procedure there were no restrictions on the language spoken, there were no auditory aids such as music or a click track.

5.2.3 EEG Acquisition and preprocessing

5.2.3.1 EEG acquisition

EEG signals were obtained using a dual BioSemi (Amsterdam, NL) ActiveTwo system configured for 64 channel recording from both participants simultaneously. Participants wore size appropriate 64 channel Electro-Cap International (Ohio, US) caps with a 10-20 electrode montage. EEG signals were recorded at 512Hz with no online filtering using ActiView data acquisition software (version 7.07; BioSemi). Signa Gel conductive electrode gel from Parker Laboratories BV (Almelo, NL) was used to bridge the connection between the electrodes and the participant's scalp. Common Mode Sense and Driven Right Leg electrodes between Pz and POz were used as the active reference.

5.2.3.2 EEG preprocessing and artifact rejection

EEG preprocessing and artifact rejection was similar to chapter 1 using an 8 step process: first data were high pass filtered at 1Hz; second line noise was removed at 50Hz using a notch filter; third data were low pass filtered at 25Hz; fourth data were referenced to a robust average, calculated by first temporarily removing noisy channels using default settings of *clean_line.m*, averaging the remaining channels and removing this robust average from all channels; fifth after robust averaging all channels, channels that were subsequently still noisy were rejected with a correlation threshold of 0.7 and a noise threshold of 4, all remaining channels were checked for bridging with *eBridge.m,* all noisy or bridged electrodes were interpolated using the spherical method of *eeg_interp.m*; sixth, using a sliding 1 second window without overlap, epochs were rejected and zeroed out if 70% of the channels exceeded -3.5 to 5 standard deviations of a robust estimate of channel EEG power; seventh conditions were concatenated and an ICA was completed with *run_ica.m*; eighth automatic rejection of bad components was completed for adult participants using *iclabel.m* and for infant participants using *iMARA.m* (Marriott-Haresign et al., 2021). Finally, data was resplit into their constituent conditions.

EEG data with more than 75% channels interpolated were rejected. Participants with 25-75% of channels interpolated had the interpolated channels removed and the remaining channels were used as a partial EEG dataset.

Individual and group level data showing the average percentage of channels interpolated vs removed, as well as the average percentage of segments zeroed out is shown in supplementary materials (SM D.3). SM D.3 also shows the total number of electrodes and seconds of EEG data included within the current analysis. In the ID speech condition at 5-, 10- 15-months, infant participants contributed 158.55, 63.08, and 33.87 minutes of artifact free EEG data, while in the ID song condition infants contributed 134.45, 72.30, and 158.55 minutes of artifact free EEG data.

To prepare the data for the TRF analysis (see section 2.6), EEG data for each participant dyad x condition were filtered into delta (1-4Hz), theta (4-8Hz) and alpha (8-12Hz) frequency bands using *eeg_popfiltnew.m*. The signals were down sampled to 128Hz to improve computational efficiency and to match the audio signals. The EEG signals were then z-scored.

5.2.4 Audio preprocessing

To synchronise the audio with the EEG data, triggers identifying the start and end of the audio were manually coded and the audio segmented into individual trials. Next the audio envelope was calculated, in order to replicate the equal spectral spacing generated by the cochlear according to Greenwood's (1990) equation, a 3rd order Butterworth filter using *filtfilt.m* was used to separate the audio from the caregiver's microphone into 9 frequency bands. The signal from each frequency band was then Hilbert transformed using *hilbert.m* to generate the analytic signal. The envelope of the absolute values of these analytic signals was then taken and averaged to produce the audio envelope. Finally, the audio envelope was downsampled to 128Hz to improve computational efficiency and the resulting envelope was z-scored.

5.2.5 Cycle-by-cycle analysis

While attempting to isolate rhythmic periods of activity in the audio envelopes it was determined that a spectral decomposition gave too many false positives across the frequency range, especially during times when the audio was silent. This is because a spectral decomposition gives a value in each frequency bin regardless of whether there is an oscillation, or even audio, it also assumes that oscillations are sinusoidal, which falsely represents sharp onsets such as those from some syllables.

Therefore, to determine timepoints where there were true oscillations in the signal, a cycle-by-cycle analysis (Cole & Voytek, 2018) was employed to verify when oscillations were present using signal characteristics in the temporal domain. The cycle-by-cycle analysis uses five parameters to determine whether oscillations are present: minimum number of cycles, minimum amplitude threshold, amplitude consistency, period consistency, and monotonicity, which looks at the symmetry of the gradients of the rise and fall of the oscillations. A test set of data with known oscillations (sourced using visual and auditory inspection of data), was used to calibrate the cycle-by-cycle analysis parameters. The parameters included were: minimum number of cycles 2, amplitude threshold of 0.5, amplitude consistency of 0.3, period consistency of 0.2 and monotonicity of 0.7. The frequency range was limited between 1-12Hz to capture the full range of frequency bands used in the multivariate temporal response function (mTRF) analysis (see section 2.6).

5.2.6 mTRF analysis

Forwards mTRF modelling represents the linear mapping of the speech audio stimulus on the neural response at each electrode site. All preprocessing was completed for each participant dyad, electrode and play condition (ID speech and ID song). Preprocessing, mTRF training and testing were completed using parameters that were similar to previous literature (Phillips et al., in prep; Attaheri et al., 2022; Gross et al., 2013).

The preprocessed audio and EEG data for each trial were segmented into 10 equally spaced folds and one of these folds was randomly assigned to the test dataset, the rest were assigned to the training dataset. mTRF analysis was carried out using the mTRF toolbox (Crosse et al., 2016) using a forwards TRF model, which can be expressed with the following equation where the neural response *r(t, n)* is a linear convolution of the input signals at time *t* across time lag τ, and channel *n*. The linear convolution of–the stimulus property *s(t - τ)* with an unknown TRF *w(τ, n)*, with *ε(t, n)* representing mTRF responses unexplained by the model.

$$r(t,n) = \sum_{\tau} w(\tau,n)s(t-\tau) + \varepsilon(t,n)$$

Individual models were computed over time-lags from -50 to 250ms (Attaheri et al., 2022; Philips et al., in prep). The optimal regularisation parameter for the mTRF model of each participant was calculated by training the models, using the training data only, over λ ridge parameters ranging from 10^{-7} to 10^{7} (Jessen et al., 2021). The optimal regularisation parameter identified for that participant was then used to train the mTRF model using *mTRFtrain.m*. The mTRF model was then tested on the test dataset using *mTRFpredict.m* to reconstruct the neural response. The predictive accuracy between the neural signal and the predicted result from the model were calculated via Pearson's correlation (r), model values were averaged, leaving one value for each participant dyad x electrode pair. This value was taken as the measure for predictive accuracy of the SBT.

To generate a control dataset, a random permutation of the audio and EEG data were calculated for each participant x electrode x condition. This was achieved by randomly shuffling the EEG and audio datasets between participants and calculating the mTRF measure of predictive accuracy as above. As the data may not have been the same length, the longer of the datasets was first cropped to the same length. 100 of these random permutations were completed and the resulting r values of predictive accuracy were then averaged over the number of permutations. This gave a control dataset with one value per participant dyad x electrode to match the test dataset.

5.2.7 Data analysis plan

The first analysis investigated whether there was above chance SBT in either the ID speech or ID song conditions, whether there were significant differences across frequency bands and whether there were significant differences between play conditions. Spatial differences were also investigated to examine whether topographic changes across the cortex account for any SBT found. To test whether there was above chance SBT, ANOVAs were calculated in each of the frequency bands of interest between the mTRF r values for conditions vs controls at each electrode. To test between frequency bands within each play condition, ANOVAs in the ID speech and ID song conditions compared the mTRF r values between each of the frequency bands at each electrode. To test whether there were significant differences between play conditions ANOVAs were calculated between ID speech and ID song conditions at each frequency band at each electrode side.

The second analysis investigated whether there was higher SBT with more rhythmic speech or song within each of the play conditions. To determine the amount of rhythmic audio in each condition, the results from the cycle-by-cycle analysis showing the number of rhythmic oscillations detected as a percentage of total trial time and the frequency of these rhythmic oscillations was correlated against the mTRF r values for each participant x frequency band. To further investigate whether there are age or frequency band effects the age points and frequency bands will be split and tested.

The third analysis investigated whether there are differences in SBT seen across infant development and frequency bands of interest, as well as testing for developmental changes in spatial responses. To test this the experimental and control data were split by age, so that SBT could be determined as above or below chance for each age range. ANOVAs were conducted for each play condition x frequency band x electrode location.

5.3. Results

5.3.1 Analysis 1

The first analysis aimed to investigate whether infant SBT was found in either play condition (ID speech or ID song) and for any frequency band of interest (delta, 1-4Hz; theta, 4-8Hz; alpha, 8-12Hz) when compared to the corresponding control datasets and to what extent there was SBT between frequency bands and play conditions. A secondary aim was to investigate whether there was significant SBT in particular spatial regions by examining differences between electrodes.

5.3.1.1 Analysis 1.1

Analysis 1.1. investigated whether there was above chance SBT in each play condition and frequency band of interest.

The results, shown graphically in figure 5.1, demonstrated for the ID speech condition that there was significant above chance SBT in the delta frequency band ($F(1,63) = 7.01$, $p < 0.01$), but not in the theta or alpha frequency bands ($p > 0.05$). While, the ID song condition showed the opposite response with no SBT shown in the delta frequency band ($p > 0.05$), but significant tracking shown in the theta ($F(1,63) = 19.45$, $p < 0.001$) and alpha ($F(1,63) = 42.39$, $p < 0.001$) frequency bands. There were no significant differences between electrodes for any condition ($p > 0.05$).

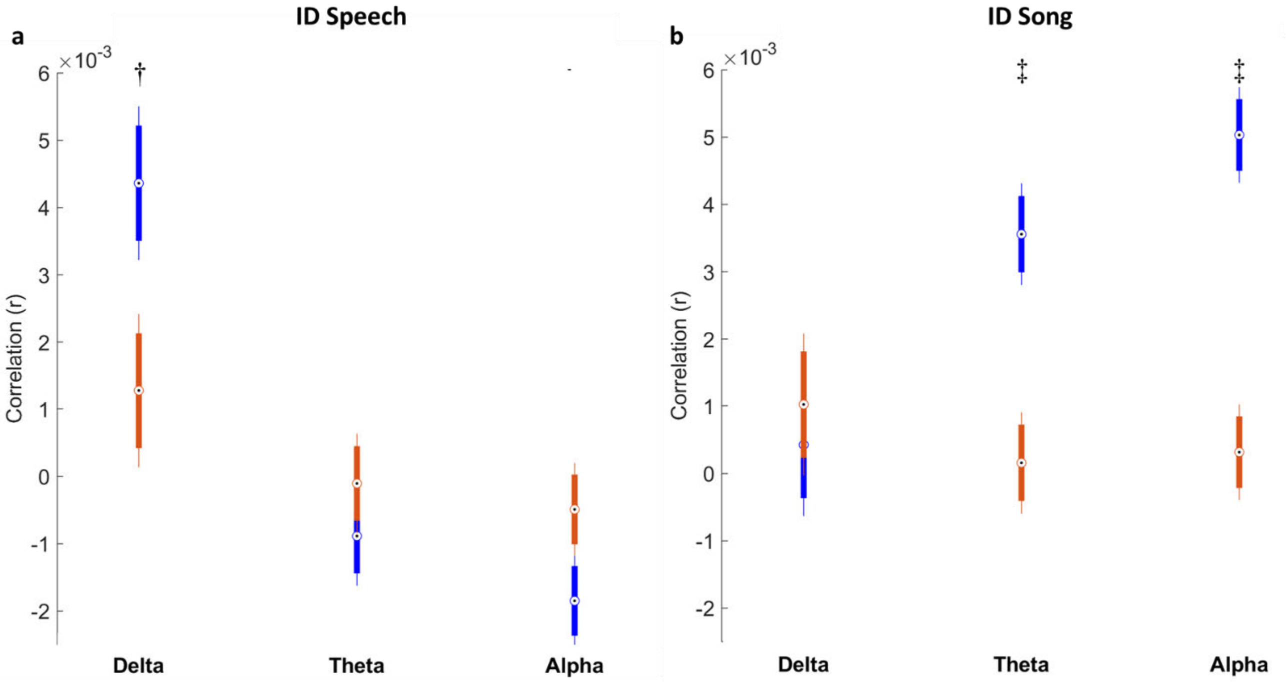


Figure 5.1 Mean mTRF values for experimental (blue) vs control (orange) conditions for ID speech (a) and ID song (b) play conditions for the delta, theta and alpha frequency bands. † denotes $p < 0.01$ and ‡ denotes $p < 0.001$. Data averaged over the whole head.

These results demonstrate that different types of play conditions with different audio types show SBT in varying frequency bands. One main difference between the play interactions is the form of vocalisation, i.e. ID singing vs ID speech, which may directly influence infant SBT. Other differences between the play conditions and to discern whether there are developmental differences are also investigated below.

5.3.1.2 Analysis 1.2

Analysis 1.2 investigated whether there were significant differences in SBT between frequency bands within each play condition and whether there were differences between electrodes representing spatial differences in SBT at different frequency bands.

The results showed that there was a significant difference between frequency bands in both the ID speech ($F(2, 63)=9.58$, $p < 0.001$) and ID song ($F(2,63)=4.82$, $p < 0.01$) conditions. There were no significant differences found between electrodes for either condition ($p > 0.05$). Multiple comparisons in the ID speech condition showed that there was significantly more SBT in the delta frequency band when compared with the theta ($p < 0.01$) and alpha ($p < 0.001$) frequency bands. There was no significant difference found between theta and alpha frequency bands ($p > 0.05$). Multiple comparisons in the ID song condition showed that there was a significant difference between the delta and alpha frequency bands ($p < 0.01$) and there were no other significant differences ($p > 0.05$).

5.3.1.3 Analysis 1.3

Analysis 1.3 investigated whether there were significant differences between the play conditions in each of the frequency bands of interest, and whether there were spatial differences seen between electrodes.

The results showed that there was a non-significant trend for more SBT in the ID speech condition vs the ID song condition in the delta band ($F(1,63) = 3.83$, $p = 0.0503$) and that there was significantly more SBT in the ID song condition in both the theta ($F(1,63) = 12.11$, $p < 0.001$) and alpha ($F(1,63) = 31.53$, $p < 0.001$) frequency bands. There were no significant differences between electrodes for any condition ($p > 0.05$).

These results for both analysis 1.2 and 1.3 are as expected, given that there was no significant SBT in the ID song delta frequency band and ID speech theta and alpha frequency bands, so when comparing to the other frequency condition it was anticipated that there would be significant differences or in the case of delta tracking a non-significant trend.

5.3.2 Analysis 2

5.3.2.1 Analysis 2.1

The second analysis investigated whether differences in SBT were the result of the level of rhythmicity of the audio. As participants in the ID song condition were asked to sing, there was an assumption that this condition contained my rhythmic audio than ID speech.

To test this, a cycle-by-cycle analysis was employed to identify oscillatory activity in the temporal domain. Table 5.1 shows descriptive statistics for each play condition. Figure 5.2a shows violin plots of the mean duration of oscillatory audio identified by the cycle-by-cycle analysis, which was summed across oscillatory bursts. Figure 5.2b shows the average frequency identified by the cycle-by-cycle analysis within the 1-12 Hz range (used to match the mTRF analyses).

A two-sampled t-test was used to examine whether there was a statistically significant difference between the amount of oscillatory audio identified in each play condition. The results showed that there was significantly more oscillatory audio in the ID song condition than the ID speech condition ($t(166) =-2.7729$, $p < 0.01$). Similarly, statistical significance between the mean frequency identified in each play condition was also calculated. The results showed that there was a significantly higher frequency identified in the ID speech condition when compared with the ID song condition ($t(166) =16.371$, $p < 0.001$).

Table 5.1. Descriptive statistics showing for each play condition the mean duration of the trials, summed amount of identified oscillatory from cycle-by-cycle analysis and the average frequency identified between 1-12Hz. Standard deviations given for each value.

	Mean duration (s)	**Oscillatory audio duration (s)**	**Mean frequency (Hz)**
ID Speech	331.76 (std, 96.58)	76.83 (std, 25.39)	5.03 (std, 2.48)
ID Song	293.13 (std, 77.64)	88.02 (std, 26.89)	3.76 (std, 1.94)

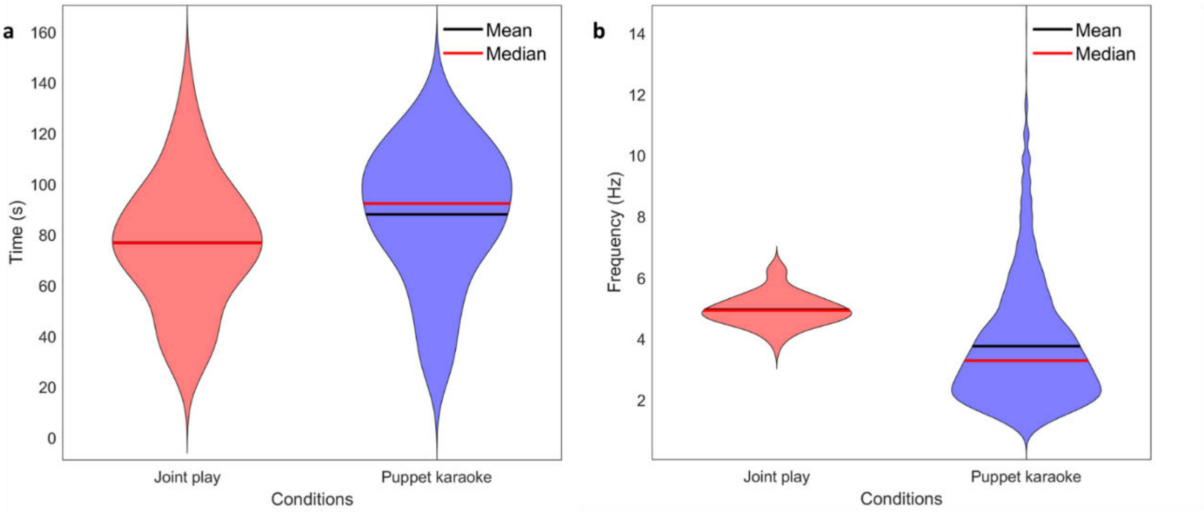


Figure 5.2 displaying for each play condition a) the amount of time in seconds detected as rhythmic by the cycle-by-cycle analysis and b) the frequency in Hz detected by the cycle-by-cycle analysis.

Play conditions were grouped for each frequency band and correlations between the level of SBT as measured by the TRF score vs the amount of oscillatory audio as measured by the cycle by cycle analysis across 1-12Hz, the percentage of oscillatory audio and the dominant oscillatory frequencies identified were shown to not be significantly correlated. This was true across all frequency bands.

5.3.2.2 Analysis 2.2

To examine this further and investigate whether rhythmicity of the audio impacts SBT at different ages, the data were split into each age group and the same analysis as in 5.3.2.1 was conducted. Table 5.2 shows descriptive statistics for each play condition and age point.

Table 5.2. Descriptive statistics showing for each play condition and age point the mean duration of the trials, summed amount of identified oscillatory from cycle-by-cycle analysis and the average frequency identified between 1-12Hz. Standard deviations given for each value.

	Age (months)	**Mean duration (s)**	**Oscillatory audio duration (s)**	**Mean frequency (Hz)**
ID Speech	5	313.58 (std, 89.5)	73.78 (std, 25.26)	5.13 (std, 2.52)
	10	337.38 (std, 99.76)	75.35 (std, 24.39)	4.91 (std, 2.43)
	15	384.87 (std, 100.79)	89.71 (std, 25.18)	4.89 (std, 2.41)
ID Song	5	282.30 (std, 75.48)	84.30 (std, 25.09)	3.67 (std, 1.90)
	10	279.86 (std, 83.10)	83.38 (std, 30.03)	3.78 (std, 1.97)
	15	350.89 (std, 45.85)	107.97 (std, 17.23)	3.91 (std, 1.96)

Two-sampled t-tests were used to examine whether there was a statistically significant difference between the amount of oscillatory audio identified in each play condition at each age. The results showed that there was significantly more oscillatory audio in the ID song condition than the ID speech condition at the 5 ($t(90) =-2.0005$, $p < 0.05$) and 15 ($t(26) =-2.2339$, $p < 0.05$) month timepoints, but there was no significant difference at the 10 month age point ($p > 0.05$).

Statistical significance between the mean frequency identified in each play condition was also calculated. The results showed that there was a significantly higher frequency identified in the ID speech condition when compared with the ID song condition at all 3 ages (5 months ($t(90) =12.6093$, $p < 0.001$); 10 months ($t(46) =8.285$, $p < 0.001$); 15 months ($t(26) =6.4781$, $p < 0.001$)).

In each age range, the play condition TRF scores were grouped by frequency band and correlated against the amount and the frequency of the oscillations identified across the whole 1-12Hz range, which showed no significant correlations for any age range ($p > 0.05$).

5.3.2.3 Analysis 2.3

A further investigation explored whether oscillations identified in particular frequency bands correlated with an increase in SBT in the corresponding TRF frequency band (e.g. oscillations identified between 1-5Hz vs the delta band TRF scores). The cycle-by-cycle analysis was restricted to match each of the TRF frequency bands, otherwise the analysis was similar to section 5.3.2.1. Table 5.3 shows descriptive statistics for each play condition and frequency band.

Table 5.3. Descriptive statistics showing for each play condition and each frequency band the mean duration of the trials, summed amount of identified oscillatory from cycle-by-cycle analysis and the average frequency identified between. Standard deviations given for each value.

	Frequency band	**Mean duration (s)**	**Oscillatory audio duration (s)**	**Mean frequency (Hz)**
ID Speech	Delta	331.76 (std, 96.58)	57.79 (std, 19.29)	2.75 (std, 1.05)
	Theta	331.76 (std, 96.58)	42.32 (std, 17.84)	6.11 (std, 1.85)
	Alpha	331.76 (std, 96.58)	17.44 (std, 9.80)	9.33 (std, 2.23)
ID Song	Delta	293.133 (std, 77.64)	75.38 (std, 23.66)	2.53 (std, 0.97)
	Theta	293.133 (std, 77.64)	28.99 (std, 13.56)	5.66 (std, 1.78)
	Alpha	293.133 (std, 77.64)	8.58 (std, 7.07)	9.06 (std, 2.23)

Two-sampled t-tests were used to examine whether there was a statistically significant difference between the amount of oscillatory audio identified in each play condition in each frequency band. The results showed that there was significantly more oscillatory audio in the ID song condition than the ID speech condition when examining the delta frequency band (t(166) =-5.2874, $p < 0.001$) but significantly less oscillatory audio in the ID song condition compared to the ID speech condition in the theta (t(166) =5.4403, $p < 0.001$) and alpha (t(166) =6.704, $p < 0.001$) frequency bands.

However, there were no significant correlations between the amount of each band of oscillations found using the cycle-by-cycle analysis and the TRF score in any of the frequency bands ($p > 0.05$).

5.3.3 Analysis 3

The third analysis aimed to investigate whether there were significant differences in SBT in each of the frequency bands and play conditions over developmental time, to examine whether developmental changes were the primary drivers behind the results seen in analysis 1. Spatial differences were also investigated to study whether there were developmental changes seen across ages.

The results for each play condition and frequency band showed that there were significant differences in SBT across age ranges: ID speech, delta ($F(5,63) = 7.0831$, $p < 0.001$), theta ($F(5,63) = 3.2524$, $p < 0.01$), alpha ($F(5,63) = 7.833$, $p < 0.001$); ID song delta ($F(5,63) = 3.4255$, $p < 0.01$), theta ($F(5,63) = 4.2039$, $p < 0.001$), alpha ($F(5,63) = 19.309$, $p < 0.001$). There were no significant differences found between electrodes for any condition ($p > 0.05$). Multiple comparisons are shown graphically in figure 5.3.

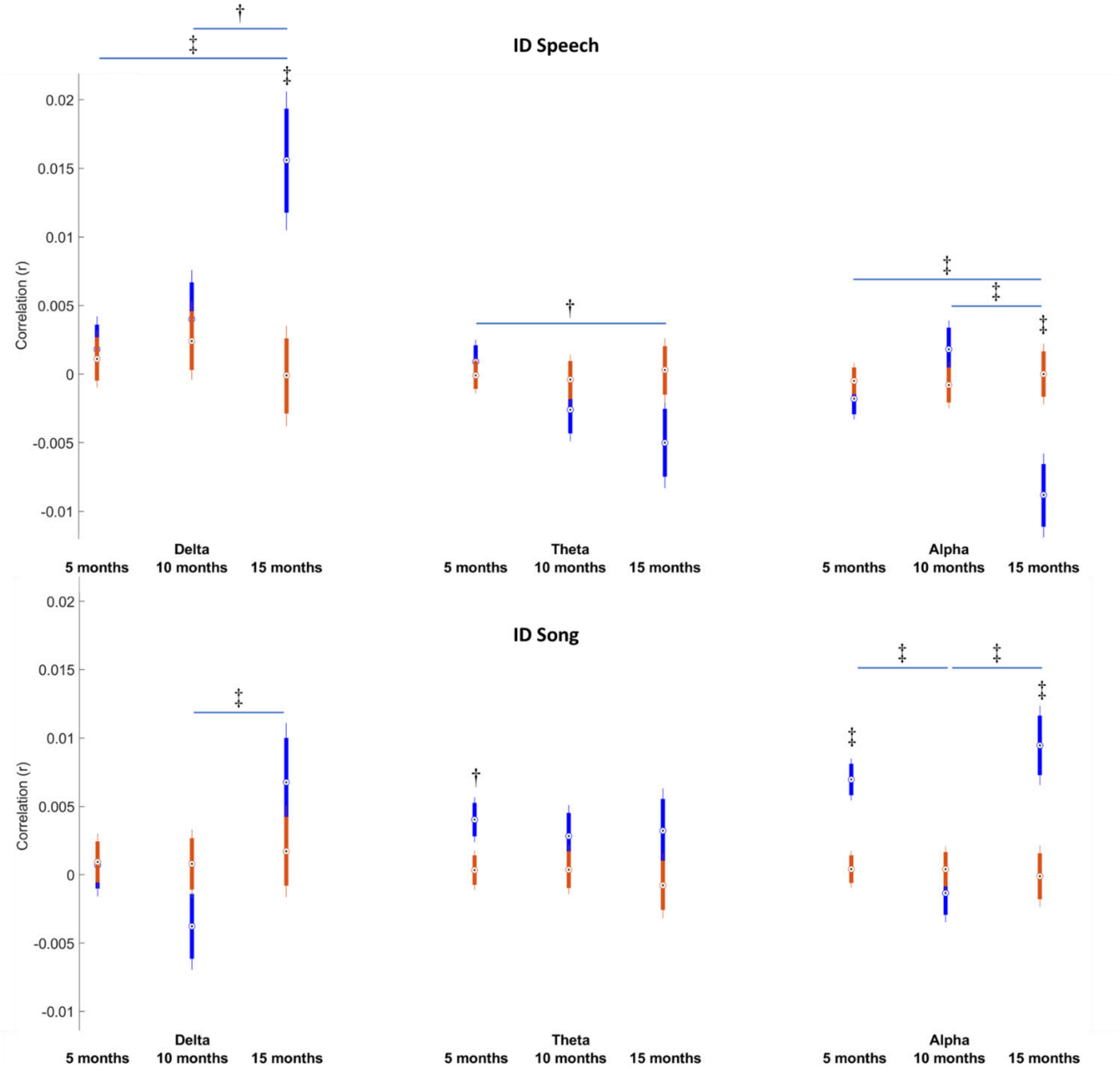


Figure 5.3 Mean mTRF values for experimental x control conditions x age group for ID speech and song play conditions in each of the three frequency bands of interest: delta, theta and alpha. † denotes $p < 0.01$ and ‡ denotes $p < 0.001$, statistical significances shown graphically between experimental vs control and between experimental conditions but not between experimental vs control between ages. Data averaged over the whole head.

Multiple comparisons between experimental and control conditions at each age point, showed that there was above chance SBT in the ID speech delta band at 15 months only ($p < 0.001$), while there was also a reduction in alpha band speech-tracking compared to chance at 15 months ($p < 0.001$). In the ID song condition, there was above chance SBT in the theta band at 6 months ($p < 0.01$), and alpha band at 6 months ($p < 0.001$) and 15 months ($p < 0.001$). There were no other significant differences found between other experimental and control conditions.

Multiple comparisons between experimental conditions for play conditions and frequency bands showed that in the ID speech condition there was significantly more SBT in the delta band at 15 months compared to both 6 months ($p < 0.001$) and 10 months ($p < 0.01$); there was significantly more theta band SBT at 6 months compared to 15 months ($p < 0.01$), and that there was significantly more SBT in the alpha band at 6 ($p < 0.001$) and 10 ($p < 0.001$) months compared to 15 months. In the ID song condition, there was significantly more SBT in the delta band at 15 months compared to 10 months ($p < 0.001$), and there was significantly more SBT in the alpha band at 6 ($p < 0.001$) and 15 ($p < 0.001$) months than at 10 months. There were no other significant differences found between age points in different play and frequency conditions.

Other significant interactions showed in the ID speech condition that there was significantly more SBT in the delta frequency band at 15 months compared to control data at 6 ($p < 0.001$) and 10 ($p < 0.001$) months; there was significantly less SBT in the theta frequency band at 15 months compared to the control data at 6 months ($p < 0.05$); and there was significantly less SBT in the alpha frequency band at 15 months compared to control data at 6 ($p < 0.001$) and 10 ($p < 0.001$) months. Finally, in the ID song condition there was significantly more SBT in the theta frequency band at 6 months compared to control data at 10 ($p < 0.01$) and 15 ($p < 0.05$) months; and there was similarly more SBT in the alpha frequency band at 6 months compared to control data at 10 ($p < 0.001$) and 15 ($p < 0.001$), more SBT at 15 months compared to control data at 6 ($p < 0.001$) and 10 ($p < 0.001$) months.

5.4. Discussion

5.4.1 Overview

In this study we examined infant SBT to a range of naturalistic play scenarios that encouraged adult caregivers to produce both ID speech and ID singing. We investigated the sensitivity of infant SBT in delta, theta and alpha frequency bands and across developmental time with infant timepoints at 6, 10 and 15 months. We examined whether differences in SBT were the result of the changes in rhythm present in the audio, age related developmental differences, and spatial differences across the scalp. To the best of our knowledge, no previous research has used these methods to attempt to investigate differences in infant SBT.

5.4.2 In which frequency bands do infants show speech brain tracking?

In analysis 1.1 we investigated whether infant SBT, as measured by mTRF r values, was present above chance in each of the play conditions (ID speech and ID song) and in each of the frequency bands of interest (delta, 1-4Hz; theta, 4-8Hz; and alpha 8-12Hz). We predicted that speech brain tracking would be present in the delta band for both ID speech and song in line with previous literature (Attaheri et al., 2022; Philips et al., in prep; Nguyen et al., preprint; Trainor et al., 1997).

The results showed SBT in opposing frequency bands across the play conditions. In the largely ID speech condition there was significant SBT in the delta frequency band, whereas in the largely ID song driven ID song there was significant SBT in the theta and alpha frequency bands. This was later supported in analyses 1.2 and 1.3 examining mTRF values across frequency bands and directly across play conditions.

This is interesting as it potentially demonstrates that the rhythmic auditory features between conditions drives SBT frequencies. However, while significant differences between the identified dominant oscillatory frequencies were shown between play conditions, these dominant frequencies did not match the frequency bands where significant SBT was identified. In the ID speech condition with significant delta band SBT, the average dominant frequency was a 5Hz, whereas the ID song condition showed significant SBT in theta and alpha bands but oscillations in the audio showed a much wider range, with mean and median dominant frequencies in the delta band range. Further, when split by frequency band there was no significant impact of the oscillations found in each of the three frequency bands on SBT within the same frequency band.

Further to this, correlations between the SBT predictive accuracy values and the amount of rhythmic auditory oscillations, as well as the frequency of those oscillations were not found to be significantly correlated. Suggesting that auditory oscillations and the auditory frequencies used are not directly leading to increases in SBT.

5.4.3 Changes in neural tracking with development

To investigate whether developmental changes in infant SBT, analysis 3 split the data by infant participant age and measured whether there was above chance SBT and differences across timepoints. We predicted that speech brain tracking would be present in the delta band at 10 and 15 months, in line with previous literature (Attaheri et al., 2022; Philips et al., in prep) we also expected that delta band SBT would increase with age. We also predicted that we would see delta frequency song brain tracking from 6 months, with increasing power as infants aged. ID singing is generally sung at a slower tempo, which we believed would facilitate song brain tracking and entrainment has previously been shown at 7 months (Nguyen et al., preprint).

The results showed that in the ID speech condition the delta band SBT seen was largely driven by the 15-month-old infants, as there was significantly more SBT than chance and vs previous age points, whereas there was no above chance SBT in 6 and 10 month old infants. Interestingly, there was also a developmental change in the theta and alpha frequency bands in the ID speech condition, with both showing a significant decrease over time and against chance. Suggesting a negative correlation between neural activity in these frequency bands and SBT.

In the ID song condition, there was a significant increase in delta band SBT between 10 and 15 months, which may suggest that there was a general increase in delta band SBT across play conditions. In the theta band while each experimental condition had a higher average than chance only the 6 month old condition was significantly higher than chance, and in the alpha band both the 6 and 15 month old condition were significantly higher than chance, but this was not seen in the 10 month old condition.

Taken together there is a clear trend in SBT to the ID speech condition with increasing delta and decreasing theta and alpha, however we did not see the same results in the SBT to ID singing in the ID song condition with a wider range of significant SBT seen. This may point to different mechanisms related to other differences in the auditory characteristics such as the pitch, auditory complexity, volume etc of the audio.

It should be noted that any age effects must be viewed with caution given the smaller sample size at older ages due to the longitudinal nature of the study. It is especially interesting that there was a significant increase in delta band tracking at older ages in ID speech conditions despite the smaller number of participants. However, further trends may have been found especially in the ID song condition with larger numbers of participants in these age ranges. Further, it must be noted that caregivers change the type and frequency of vocalisations as infants age (e.g. Narayan and McDermott, 2016) and so while every care has been taken to make the conditions the same across ages it would not be possible to standardise the types and frequency of vocalisations without compromising naturalistic aspects of the study.

5.4.4 Spatial differences in neural-tracking

Throughout these analyses, the spatial characteristics of the SBT were investigated by examining whether there were differences on an electrode level. It was surprising to find no significant differences across play conditions, frequency bands and even when splitting the data by age range. It was expected that there would be significantly more SBT in the auditory cortex, which was expected to increase as participants aged. We also wanted to investigate whether there were differences seen in the visual cortex, and auditory vs visual cortices, as we expected that there may have been some speech brain tracking driven by visual cues such as facial movements and expressions as well as gestures and actions with the toys and puppet.

The lack of significant differences suggests that infant forwards mTRF models are not sensitive enough to discern spatial differences and a much larger infant participant pool would be required to investigate this further.

5.4.5 Limitations and future research

While 81 participant dyads were included in the study, due to the longitudinal nature of the data collection there were more infants at younger ages, which could have biased results in the younger age ranges to have more statistical power when compared to older ages. It would have been interesting to see if the decrease in SBT seen in the theta and alpha frequency bands as infants aged would have been shown more strongly with more participants. It is also interesting that there is a significant song brain tracking response seen at both 5 and 15 months in the alpha band but not at 10 months. More participants at the 10 month age point would allow for a higher confidence level in these results and give room to break down the participants to attempt to understand this dip. Despite this, there was still significant speech brain tracking demonstrated at older timepoints, which may either suggest that there is strong SBT or that the current result do not fully encompass the extent of SBT at older ages. This is potentially seen in the unclear developmental trend in the ID song condition.

A further limitation was that there were no restrictions on the type of speech in the ID speech condition such as language used, the level of salient information including as naming, positive or negative affect vocalisations etc. Similarly, participants were not instructed to sing play songs, lullabies or even infant specific songs. These restrictions were not put in place because we wanted the interactions to be as natural as possible within the lab environment. The type of speech or song could have impacted the neural tracking that was seen, and future research would benefit from a larger cohort so further distinctions could be made.

5.4.6 Conclusion

The current research takes a step towards differentiating the mechanisms behind infant SBT, shows that there are frequency-based differences in spoken vs sung interactions, that there are developmental differences, but does not provide substantive evidence that SBT is driven by the rhythmic nature of ID song and speech.

Chapter 6 –Rocking As Soothing Behaviour Impact Of Dynamic Caregiver Rocking Characteristics On Infant Soothing, Evidence From Day Long Home Recordings

The manuscript is currently in preparation for publication. The supplementary materials for this chapter are presented in Appendix E.

Overview

Chapter 6 of this thesis explores the concept of rhythmic entrainment dysregulation in infancy—an area of research that has received minimal attention in the literature. Rhythmic entrainment dysregulation refers to the disruption of an infant's internal oscillatory rhythm by competing external rhythms, such as when an infant is crying and is soothed rhythmically by their caregiver. Unlike auditory entrainment, which can be studied using controlled experimental paradigms, investigating rhythmic entrainment dysregulation poses unique challenges due to the subjective nature of top-down infant auditory entrainment and the difficulty in discerning intentional inhibition of external stimuli.

Chapter 6 utilizes day-long home recordings with custom-made wearable devices to capture the sub-second interplay of rhythmic infant negative arousal episodes and rhythmic caregiver soothing responses through infant rocking. Continuous audio recordings have been integrated with GPS, proximity, actigraphy, and heart rate data. This holistic approach allows for a comprehensive examination of the temporal dynamics of rhythmic entrainment dysregulation and its impact on infant distress and soothing behaviour.

The findings from this study advance our understanding of infant rhythmic entrainment dysregulation and its role in caregiver-infant interactions. By investigating the effectiveness of various rhythmic soothing techniques in modulating infant distress, which may with future study be developed into targeted interventions to support infant well-being.

Overall, Chapter 6 represents an initial investigation into rhythmic entrainment dysregulation in infancy. By combining novel wearable methodologies with a nuanced understanding of caregiver-infant interactions, this chapter advances our knowledge of infant development and lays the groundwork for future research aimed at promoting infant well-being and enhancing caregiver-infant bonding.

Abstract

Parents frequently turn to rhythmic rocking to soothe their infants, but no study has found an optimal rocking frequency or technique to reduce negative arousal episodes. Previous lab studies have investigated rocking frequency, amplitude and direction of rocking; however, there are a lack of studies that have investigated this in a natural setting or investigated how dynamic characteristics such as changes in phase difference between crying and rocking oscillations impact cry duration. This study investigates characteristics of soothing rocking behaviour and oscillatory vocal crying in natural day long home recordings. Results show that infants produce descending crying frequencies throughout crying bouts and that adults dynamically change their rocking rhythms to remain out of sync with crying rhythms. However, caregivers that kept a consistent phase difference between crying and rocking oscillations were associated with a decreased crying length. Across longer timescales, caregivers were shown to track rocking rates to increases and decreases in the infant crying audio envelope oscillations and employed a turn taking rather than competing structure to their soothing interactions. The present study discusses the future directions to investigate dynamic changes in soothing behaviours and the importance of collecting data in the home.

Word count: 194

Keywords: Infant entrainment, entrainment disruption, infant crying, crying behaviour

6.1 Introduction

6.1.1 Crying rhythms

Ultrasound scans have shown homologues of infant crying as early as the third trimester of gestation (Gingras, Mitchell & Grattan, 2005), and while crying is thought to peak at around 6-12 weeks, crying remains relatively high throughout the first 12 months of infancy at an average of 1-2 hours a day (James-Roberts & Halil, 1991; Vermillet et al., 2022).

Studies investigating cry lengths have suggested that the natural rate of individual cry vocalisations usually lasts between 1100-2400ms total (Bobin-Bègue, 2019; Provasi, 2016) with a relatively consistent 200-300ms period of inspiration across the first 12 months, pauses around 200ms and a variable expiration between 600-2200ms, which generally increases with over the first 12 months to give longer cries (Bobin-Bègue, 2019; Wilder & Baken, 1975; Wolff, 1969).

6.1.1.1 Anatomy of crying

Cries have been split into three basic types, a voiced *phonation* cry, a noisy and turbulent *dysphonation* cry and a very high-pitched *hyperphonation* cry (Truby and Lind, 1965). Infants use different cry types along with changes in pitch and duration to signal a variety of needs such as hunger, sleep, pain, and attention to their caregivers (for a review see Zeskind et al., 2013).

When distressed infants produce a hyperphonation cries with the maximal volume and effort, limited only by the infant's own anatomy. The brain coordinates the brainstem, midbrain and limbic systems to produce changes in the level of arousal, respiratory cycle and a tightening of the laryngeal muscles, which raises vocal pitch (Zeskind, 2013). Thoracoabdominal synchrony of muscles in the thorax and abdomen produce the coordinated breaths to raise volume (Wilder and Baken, 1975). The acoustics of cry production is split between the shape and volume of the subglottal (lungs), glottal (larynx) and supraglottal (vocal tract) production areas (Golub, 1980).

Throughout this internal synchrony, cries have a rhythmic quality, which may be based on infants producing a cry and listening for a response, or just based on how quickly an infant can get breaths in and out of themselves (Provasi, Anderson, Barbu-Roth, 2014; Zeskind, 2013).

6.1.1.2 Individual differences in crying

Based on the brain to respiratory anatomy alone it could be argued that infant crying provides limited information to caregivers, which was argued to be true in the 1950s (Miller, 1951). However, as with any instrument infants very quickly learn how to tune their cries to their situation and the responses that they receive from their caregivers. Cries have been shown to differ between cultures, based on the health of the infant and due to internal changes in feedback.

Mampe et al., (2009) demonstrated with French and German infants that cry melody contours were significantly different, German cries showed a decreasing melodic contour while French cries showed a rising contour showing that native language and the infant’s environment play a role in shaping this behaviour.

Cries have also been shown to be systematically different between healthy and pathologically ill infants. By studying the differences in entropy and correlation dimensions from cepstrums of the infant crying Lahmiri et al., (2021) have shown that the cepstrums of healthy infant cries were exhibited lower disorder and randomness than unhealthy infants. This could mean that cry recordings could be used as biomarkers for illness.

As well as a range of cries for different situations, and fundamental differences based on health (Lahmiri et al., 2021) and culture (Mampe et al., 2009), crying is a behaviour that relies on both internal and external feedback. Using healthy infants Cullen et al., (1968) demonstrated that infant cries change when they listen to their own cries with delayed audio feedback, which decreased the produced cry length. While infants who are profoundly deaf have been shown to cry with a different melody and rhythm than their hearing counterparts. Deaf infants produced lower frequency and longer cries, which were found to change to more closely resemble hearing infant cries once hearing aids had been fitted (Möller & Schönweiler, 1999).

6.1.1.3 Caregiver responses to cries

Caregivers too show different responses based on different cry features, for example Zeskind, Klein & Marshall, (1992) demonstrated that caregivers rated cries differently based on the duration of pauses between cries, even when the pitch and intensity of the cries was kept constant. Cries with shortened pauses were rated as more “arousing, informative and aversive” as well as increasing the rated urgency of the cry. Duration of infant cries has also been shown to be an important factor in caregiver response. A study by Zeifman (2004) asked caregivers to watch a range of clips of infants progressing from fussing to crying and rate when, if at all, they would pick up a child to soothe them. Duration of fussing rather than fundamental frequency of the cry was shown to be the most predictive of caregiver responses.

Research investigating how parents should respond to infants during the first 12 months to reduce infant crying episodes often provide mixed advice for parents. Caregivers have often been told to answer infant cries “swiftly, consistently and comprehensively” (Ludington-Hoe, Cong, & Hashemi, 2002). However, there is a strong debate relating to whether parents should leave infants to self-soothe or “cry it out”. Research by Bell and

Ainsworth (1972) demonstrated that promptness and consistency of maternal response is associated with a decline in the duration and frequency of infant crying. In fact, simply responding to the cry early was more effective than other soothing efforts. This is supported by research that shows that increased carrying of infants, which decreases the time for caregivers to respond and could be considered as a pre-emptive response to distress, decreases crying time (Hunziker & Barr, 1986). However subsequent research has suggested that higher maternal responsiveness to crying is associated with greater cry durations. Those parents that practice "benign neglect" where found to decrease the rate at which infants cried and promote self-soothing behaviours (van Ijzendoorn & Hubbard, 2000).

Given the profound impact of infant distress on caregiver fatigue, parental distress and depression (e.g., Kurth et al., 2011; Murray & Cooper, 2001), it is important that research is as thorough as possible when providing information that could be used to soothe upset infants in everyday settings. The present study will focus on the characteristics of caregiver rocking behaviours in the home environment.

6.1.2 Rocking rhythms

Rocking rhythms, often referred to as vestibular stimulation, are a ubiquitous soothing technique that caregivers employ the world over (e.g. Bobin-Bègue, 2019). Vestibular stimulation begins during gestation with rocking through the mother's movements, such as walking, dancing, and exercising (for review see Provasi, Blanc & Carchon, 2021; kinematic representation of pregnant walking and the rhythms experienced by the foetus are described in Sydor et al., 2022). Evidence suggests that rocking rhythms are perceived *in utero*, with late gestation foetal heart rate accelerations observed in response to rocking movements such as sitting in a rocking chair or swaying (Lecanuel & Jacquet, 2001).

After birth, rocking has repeatedly been shown to quieten or stop crying in infants (e.g. Elliott, Fisher & Ames, 1988). Parents rated rocking as the most effective of 18 common soothing methods, equal only to feeding (Ames et al., 1987). Simply holding infants has also been shown to quickly quieten infants (Esposito, et al., 2013), and many motorised rocking machines have been developed showing that rocking alone is a positive soothing technique (see section 1.2.2.1), however both holding and rocking an infant has been shown to be the most effective (e.g. Byrne & Horowitz, 1981).

As well as being an effective method of soothing crying infants, natural rhythmic vestibular motion also causes other positive effects, for example infants that were shown to be carried more throughout the day were 43% less likely to cry between 4pm and midnight (Hunziker & Barr, 1986). Other benefits found include significant increases in neuromuscular development (Clark et al., 1989); reduced bouts of apnoea (Korner et al., 1978; Tuck et al., 1982); reduced adverse respiratory and oromotor symptoms (Zimmerman & Barlow, 2011); and decreased rates of activity, better orienting responses, fewer abnormal reflexes and higher Mental Development Index scores (Barnard & Bee, 1983). For further review see Sandler and Coren (1981).

Despite this the features of rocking have received limited investigation in natural at home settings, such that there is limited research showing the natural frequency at which infants are rocked, which directions are most prevalent, rocking durations etc. There is also little consensus on how characteristics such as direction of rocking, frequency, phase difference, dynamic responses from caregivers and turn taking behaviour soothe infants. A summary of previous work is below.

6.1.2.1 Direction of rocking

In the present study, to avoid confusion the direction of rocking will be described relative to the anatomical axes of the infant's body and the body position will also be described. Up-down rocking refers to rocking along the longitudinal axis from head to toe; left-right rocking refers to a side-to-side motion; and forward-back rocking refers to an axis from the infant's chest to back. An infant may be held vertically, for example when resting on the caregiver's shoulder or in a supine position, i.e., lying down, which will refer to an infant being face up unless otherwise stated.

Lab studies have demonstrated that the direction an infant is being rocked impacts the effectiveness of the soothing behaviour. Caregivers and infants may both have a preference for the infant's orientation in space and how they are rocked. Byrne and Horowitz (1981) and Nishimura, Kanakogi & Myowa-Yamakoshi, (2016) investigated the impact of rocking newborn or six-month-old infants across different axes and showed that up-down rocking while the infant was positioned vertically produced a "bright alert" state, while left-right rocking while the infant was in a vertical position was likely to induce calmness and sleep. Accordingly, caregivers used different rocking styles in different scenarios and were likely to use up-down rocking while the child was uncomfortable and left-right rocking while they were comfortable.

Pederson, Champagne & Pederson, (1969), demonstrated with 2-month-old infants that when in a supine position, up-down rocking was more effective than left-right rocking when attempting to soothe an infant and induce sleep. Whereas in subsequent studies (Pederson, 1973; 1975) the axis or rocking was found to not be a significant factor in soothing infants and instead the frequency and amplitude of the rocking were found to have a larger impact.

6.1.2.2 Frequency

A number of studies have investigated the impact of sustained, consistent rocking frequency on soothing capabilities. Some studies have suggested that the ideal stimulation rate is at 1-1.5Hz or 60-90 rocks/min (Pederson, 1975; Vrugt & Pederson, 1973), however this has not been consistently shown across the literature. To quantify the best rate of rocking, studies have investigated both natural and mechanised methods of rocking infants.

6.1.2.2.1 Mechanised rocking research

Mechanised rocking research generally involves a mechanised bassinette that rocks infants along one of two axes, up-down or left-right, while the infant is in a supine position. To our knowledge no research has yet investigated the effects of a forward-back rocking motion on infants in a bassinette or investigated the frequency of mechanical rocking while the child was in a vertical position.

Vrugt & Pedereson (1973) rocked 2-month-old infants at a steady rate of 0, 0.5, 1 or 1.5Hz using up-down rocking while the infant was in a supine position and demonstrated that activity and arousal levels decreased with higher rocking levels, which lead to fewer cries and more sleeping among the cohort. Similarly other studies have investigated a range of rocking, investigating rocking infants at a consistent rate for a period of time, and demonstrated that the highest frequency condition was the most effective at soothing infants, including Ambrose (1969) 0.5 and 1Hz; De Lucia (1969) 0-1.165Hz; Elliott, Fisher & Ames, (1988) 0.66-0.95Hz; Pederson (1973; 1975) 0.75 and 1Hz; Pederson and Vrugt (1973) 0-1.166Hz; and Zimmerman and Barlow (2011) 0.5-0.95Hz.

To remove doubt that the motion rather than other factors, such as the sound of the motor, were confounding the results MacKinnon et al., (1986) either rocked infants at 50 cycles/min or played the sound of the motor without any movement and found that the movement caused a change in breath-to-breath variability, while no movement caused no change.

6.1.2.2.2 Naturalistic rocking research

There are two main drawbacks of mechanised rocking research, first is that it removes the sensory stimulation received while being held by a caregiver, which has been shown to induce a calming response and secondly, because infants are often not strapped into the rockers, the mechanisms generally move slower than naturalistic rocking frequencies due to the risk of injury to the infant. For example, Nishimura, Kanakogi & Myowa-Yamakoshi, (2016) recorded in a lab setting that parents were more likely to rock their infants above 1Hz and of those rocking episodes the mean frequency was 2.6Hz, with a range from 1.56-4Hz.

Unfortunately, to the authors' knowledge there have been no natural studies that have reported on the average rocking rate of caregivers when soothing their infants.

6.1.2.2.3 Impact of contingent rocking frequency on soothing

A leading theory of the importance of rocking frequency as a contingent soothing mechanism, is that rocking entrains infants' respiration during the movement. Sammon and Darnell (1994) demonstrated in a lab setting that when caregivers rocked infants between 0.5-1Hz that infants entrain their respiration to the rocking generally in a 1:1 or 2:1 ratio depending on the speed of the rocking, suggesting that infants can be comfortable at a range or rhythms. This has also been seen in ill infants, for example an intervention by Zimmerman and Barlow (2012) showed that faster rocking increased respiration, which alleviated respiratory and oromotor symptoms. Zimmerman and Barlow noted that acceleration seemed to be the driving cause of these changes.

If respiration can be quickly entrained to rocking motions, it may make crying oscillations entrain to the rocking rhythm without conscious effort from the infant. This could then be used as a tool by parents to soothe infants by changing their rhythm to make it difficult for infants to settle into a particular frequency.

6.1.2.3 Phase

While there is some research looking at the frequency of infant rocking, to our knowledge there are no studies that have previously sought to investigate whether caregivers attempt to maintain a specific phase difference between their rocking rhythms and infant crying oscillations. For example, to reduce cry length, a caregiver may attempt to have their rocking oscillations always be out of phase so that their rocking oscillations compete with infant crying oscillations. Previous studies have suggested that "rhythmic stimulation appears to enhance crying production in the same way it enhances stepping production" Provasi, Anderson & Barbu-Roth, (2014). Therefore, it follows that caregivers would want to avoid in phase concurrent rhythmic stimulation, while infants either through conscious effort of simply through respiratory entrainment to rocking may change their crying rhythms to bring it in line with caregiver rocking. In the auditory domain, competing vocalisations either fed back to the infant (Cullen et al., 1968) or overlapping soothing vocalisations from the caregiver (Yoo, Bowman & Oller, 2018) have been shown to decrease cry length, and in the case of the soothing vocalisations, caregivers actively chose to compete with infant crying rhythms.

To our knowledge no study has attempted to investigate whether caregivers employ changes in rocking frequency specifically to maintain an out-of-phase difference between infant crying oscillations and caregivers rocking rhythms.

6.1.2.4 Turn taking

In a similar vein, to our knowledge no study has investigated whether crying bouts and rocking bouts follow a turn taking structure, i.e., after the infant produces oscillatory cries, the caregiver responds with rocking and when the rocking stops the infant starts to produce oscillatory cry vocalisations again.

Previous studies (e.g. Byrne & Horwitz, 1981) have shown the usefulness of intermittent rocking strategies. However, the pattern of those intermittent rocking behaviours has not been explored. Do parent responses to negative vocalisations take on a turn taking or an overlapping structure? Alternatively, is the vestibular stimulation consistent throughout?

Yoo, Bowman & Oller, (2018) demonstrated that as well as competing soothing vocalisations to negative infant vocalisations, caregivers also produced non-competing, turn taking structures of vocalisations when infants produced positive vocalisations. This suggests that there are different conversational structures used depending on the emotional valence of the vocalisation.

6.1.3 Current question

Despite being a large part of raising infants, little research has been conducted to investigate rocking as a soothing behaviour in natural settings. Further, while some studies have invested some characteristics of rocking behaviour such as duration, frequency, amplitude and direction of motion, to our knowledge no study has investigated the effects using day long recordings or investigated dynamic characteristics including phase difference, changes in frequency over time or concurrent vs turn taking soothing structures.

In the present study we investigate day long at home recordings of continuous audio, proximity and actigraphy measurements to investigate how caregivers soothe 6 month old infants.

6.2. Methods

6.2.1 Participants

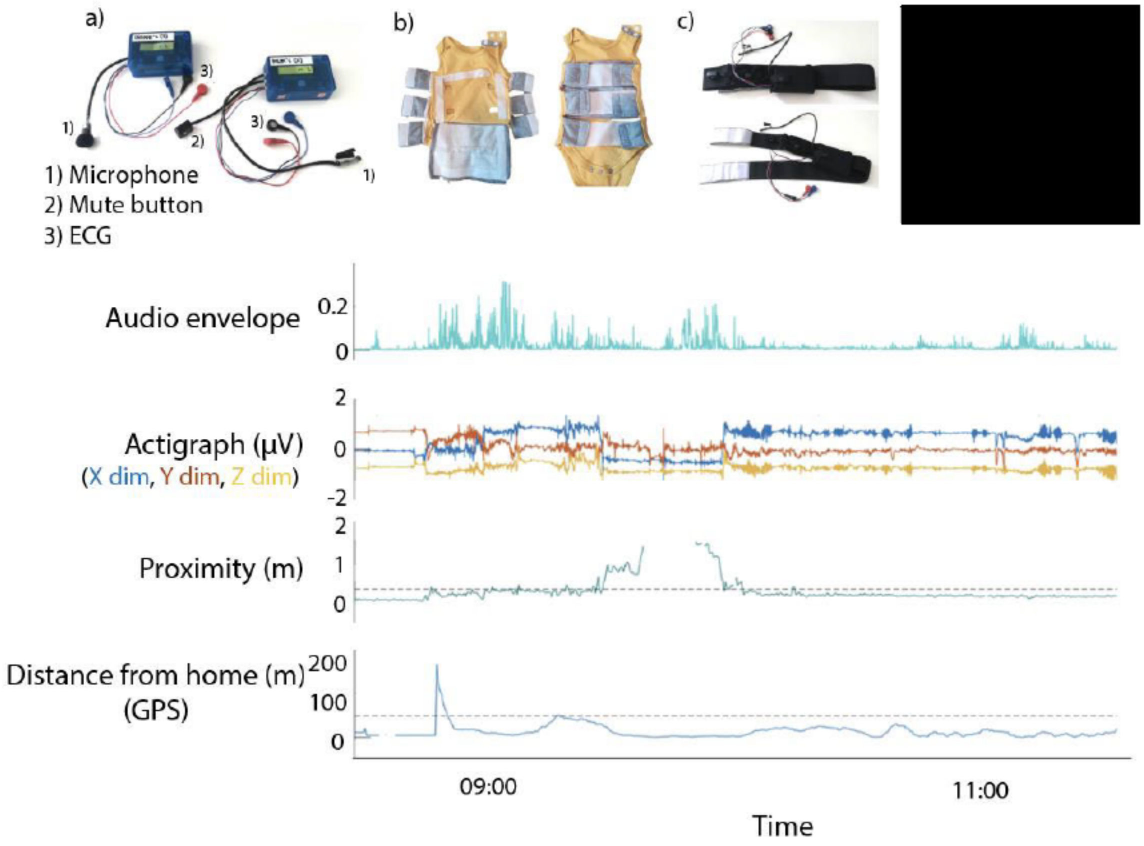
65 adult-infant dyad participants were recruited as part of a wider longitudinal project. Mean adult participant age was 35.86 years (standard error, 0.5), mean infant age 5.98 months (standard error, 0.09). All adult participants were female, infant participants included 35 male, 31 female. Infants had an average gestational age of 41.63 weeks (standard error, 0.38). The study involved day long wearable device recordings (mean, 8.41 hours, standard error, 0.26). A table showing participants details has been placed in the supplementary materials (SM6.1).

The current experiment was part of a longitudinal research project, which tests infant-caregiver dyads at 5-, 10-, 15- and 36-month timepoints. The University of East London ethics committee approved the study. All adult participants provided informed consent for both themselves and their children according to the Declaration of Helsinki. All participants were offered a £10 shopping voucher as a monetary reward for their time.

6.2.2 Procedure

Participants arranged a date and time to conduct the session with experimenters. Participants were asked to choose a day when they would be with their infant for the duration, and they wouldn’t be doing any water sports. An experimenter arrived in the morning, synced the wearable devices (WDs) to a GPS signal and to each other, then started the recording, before helping the participants to put on the WDs.

The WDs consist of a small box 8.3 x 5.6 x 2.7 cm with three ECG leads and a microphone protruding from the casing. Adult WDs also had a mute button protruding from the box, which stopped the audio recording for 90 seconds at a time. Adult participants were asked to wear the WD in an elastic belt around the centre of their torso, while infant participants wore a custom-made baby grow with a specialised pocket (see fig 1).



[redacted]

Figure 6.1. Example data collected in a 20 minute portion of the home recording for one dyad. Diagram showing a) concurrent data collected in a 20 minute period including audio envelope derived from raw audio signal, detection of cries using machine learning model, detection of cry oscillations using cycle-by-cycle analysis; actigraphy data for each of the 3 axes and actigraphy oscillation detection using a cycle-by-cycle analysis; proximity data show distance between participants in metres, and a dashed line showing the cut off and distance from the home derived from GPS data shown in metres with a dashed line showing the cut off. B) shows an image of the wearable devices along with labels for the ECG leads, microphone, mute button, on/off and start switches. C) shows the custom made clothes used for adult and infant participants.

Once setup, participants were instructed on how to mute the continuous audio of the recording for short periods and told that they could remove the devices at any point from themselves and their infants if they needed to. Participants were instructed not to stop the recording. At an arranged time in the afternoon or evening, the experimenter would return to collect the WDs, stop the recording and collect any feedback from the participants.

6.2.3 Wearable device data acquisition

The WDs are custom made continuous data recording devices with a battery life of up to 16 hours. Data collected includes: GPS signal (1Hz), which also doubled to timestamp all of the data; continuous audio (48kHz), this could be muted with a mute button, which zeroed the audio stream only; proximity (1Hz), measured in metres using longwave radio signals between WDs; ECG (250 Hz), measured using a three-lead configuration with leads placed on the right clavicle, lower left and right ribs; and three axis actigraphy, (30/40Hz) measured relative to gravity.

During device setup the WDs were automatically synced with GPS signals and then to each other. When the parent device was muted using the mute button this also muted the infant device if it was within range.

6.2.4 Preprocessing

The total number of hours recorded from each participant is listed in SM E.1, in total 546.58 hours were successfully recorded from the adult-infant dyads.

6.2.4.1 Actigraphy preprocessing

Raw actigraphy data show movement across three axes, which are preprocessed individually. Actigraphy sampling rate was first normalised to check that the number of samples taken per second was the same throughout the recording and to remove or add NaN values when this was not the case.

While attempting to isolate rhythmic periods of activity in the audio and actigraphy signals it was determined that a spectral decomposition gave too many false positives at different frequency ranges, especially during times when the audio was silent, or the devices weren't moving (see appendix E SM2).

Therefore, to determine where there were true oscillations in the signal, a cycle-by-cycle analysis (Cole & Voytek, 2018) was employed to verify when there were oscillations using signal characteristics in the temporal domain. The cycle-by-cycle analysis uses five parameters to determine where there are oscillations: minimum number of cycles, amplitude threshold, amplitude consistency, period consistency, and monotonicity, which looks at the symmetry of the gradients of the rise and fall of the oscillations.

A test sample of data with known oscillatory activity from both swinging and bouncing the wearable devices were used to calibrate the cycle-by-cycle parameters with a minimum number of 3 cycles, amplitude threshold of 0.5, amplitude consistency of 0.1, period consistency of 0.5 and monotonicity threshold of 0.6. See supplementary materials SM2. The cycle-by-cycle analysis searched for oscillations in the 0.1-5Hz range.

6.2.4.2 Audio preprocessing

There were two streams of audio preprocessing, first to detect cries using a machine learning algorithm and second to detect rhythmic oscillations in the audio envelope, which will be compared to rhythmic oscillations in the actigraphy signal. The audio files were originally recorded in an .mp3 format, which was first converted to .wav using ffmpeg in the custom-made Python script.

6.2.4.2.1 Detecting cries using machine learning

A machine learning cry detector using a support vector model was used to identify periods of crying and fussing in the audio recordings (Micheletti et al., 2022). This cry detector had previously been validated using 24-hour infant-adult LENA recordings. Cry detection is split into two stages, a preprocessing stage to identify likely periods of time where there may be a cry and a prediction stage where the support vector model is applied. The only change made was to the preprocessing stage, to highpass at 650 rather than 350Hz to reduce rustling and other ambient noise.

6.2.4.2.2 Detecting rhythmic oscillation in the audio envelope

Cries were found to be most distinct compared to other sounds in the 800-1000Hz spectral range, so an audio envelope of the 800-1000Hz range was created by first applying a Butterworth bandpass filter of 800-1000Hz with a filter order of 4 using *filter.m*. The audio was downsampled from 48000 to 2400Hz using *resample.m* to reduce the computation power needed to analyse the data. The audio envelope was calculated using *envelope.m* using a 600 tap Hilbert filter to compute the upper envelope of the signal. A moving mean using 100ms bins was then taken to smooth out the signal into just the vocalisation features. Figure 6.2 shows a graphical representation of these preprocessing steps with an example crying bout.

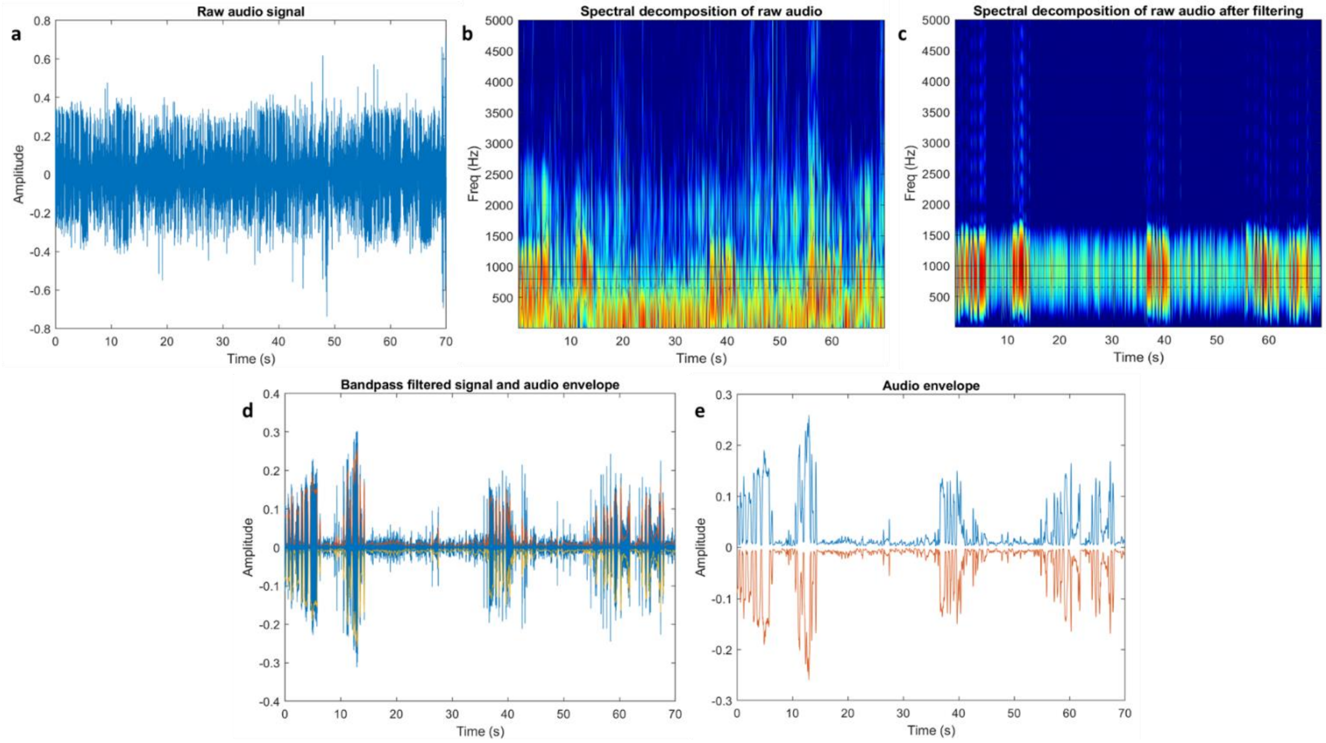


Figure 6.2 showing a) a section of raw audio data, b) the spectral decomposition, black lines at 800 and 1000Hz give an indication of where infant crying has been detected; c) filtered spectral decomposition between 800-1000Hz; d) filtered audio signal with overlaid amplitude envelope, d) amplitude envelope only of the filtered audio signal.

Using the audio envelope, a cycle-by-cycle analysis was performed to identify oscillatory time periods in the signal using features in the temporal domain in a similar way to the actigraphy oscillation detection (see 2.4.1). A test set of data with known oscillations (sourced using visual inspection of data), was used to calibrate the cycle-by-cycle analysis parameters, which were different from the actigraphy calibration due to the structure of the data, where more of the data was likely to be close to zero (see supplementary materials SM E.2). The parameters included were: minimum number of cycles 2, amplitude threshold of 0.5, amplitude consistency of 0.3, period consistency of 0.2 and monotonicity of 0.7.

6.2.4.3 Preprocessing conducted on crying bouts

Further preprocessing was conducted only on crying bout segments as identified by the cry detector. For each participant, the actigraphy, actigraphy cycle-by-cycle output, audio amplitude envelope, audio cycle-by-cycle output, GPS and proximity data were all loaded, synced and epoched into segments based on when the infant was crying as found by the cry detector.

GPS data was used to filter for when participants were inside their homes. Any data outside the participants' homes was excluded due to variability in environmental conditions which made it difficult to determine whether the infant participant or another infant were crying, or whether the infant was being rocked by the primary caregiver, another adult or had been placed inside a vehicle. To calculate when the participants were inside their homes, the first recorded GPS reading was taken, which was always inside the participants house, and the distance from the starting point was calculated for all subsequent timepoints. Any data past a 50m threshold for at least 20% of each crying bout was deemed as being outside the participants' homes.

Proximity data was used to determine whether the infants were within 30 cm of their caregivers and likely being held. As the wearable devices are worn on the torso of each of the participants, a distance of 30cm was chosen as it was thought to be unlikely that a caregiver would spend a sustained amount of time with their bodies within this range without the infants being held. Calibration tests on the proximity sensors were calculated (see supplementary materials SM3), this showed that a correction was required. A polynomial curve was calculated using *polyfit.m* and applied to the proximity data for each participant using *polyval.m*. Any data where the participants were further than 30cms apart for over 20% of the crying bout or where over 50% of the proximity data were missing were excluded from further analyses.

Finally, all data within the first and last 15 minutes of the recording were also excluded from further analyses to ensure that the experimenter would not be present during the crying bouts.

For each of the crying bouts, rocking activity was operationalised as there being at least two consecutive seconds of actigraphy oscillations, as determined by the cycle-by-cycle analysis, in the data. All crying bouts that met these criteria were put into a crying-rocking condition (N=1368, total=31,876 seconds, mean=23.3s, std=28.31s), all crying bouts that did not meet these criteria were put into a crying-not rocking condition (N=1242, total=31,683 seconds, mean=23.26, std=28.37). Matching non-crying conditions were also generated by randomly generating an onset with a matching period of time to the crying duration within the day long recordings. The cry detector output was checked to ensure that this episode had not been identified as a period that included any crying, and the level of rocking was matched as either passing the rocking criteria, as described above, or not. These conditions therefore created the not crying-rocking and not crying-not rocking conditions. While non-crying periods were not found to have included crying as detected by the cry detector, these segments may have included vocalisations, ambient noise or other sounds and so there may have also been auditory oscillations in these periods.

6.2.5 Data analysis plan

6.2.5.1 Descriptives

To ensure that the preprocessing steps had accurately separated the data into their designated groups, analyses were run to check the amount of crying and rocking, as determined using the methods in 2.4.3 (see SM4); as well as the amount of rocking per crying bout in the two rocking conditions (see SM4), the time to start rocking, the auditory oscillation duration after rocking had started.

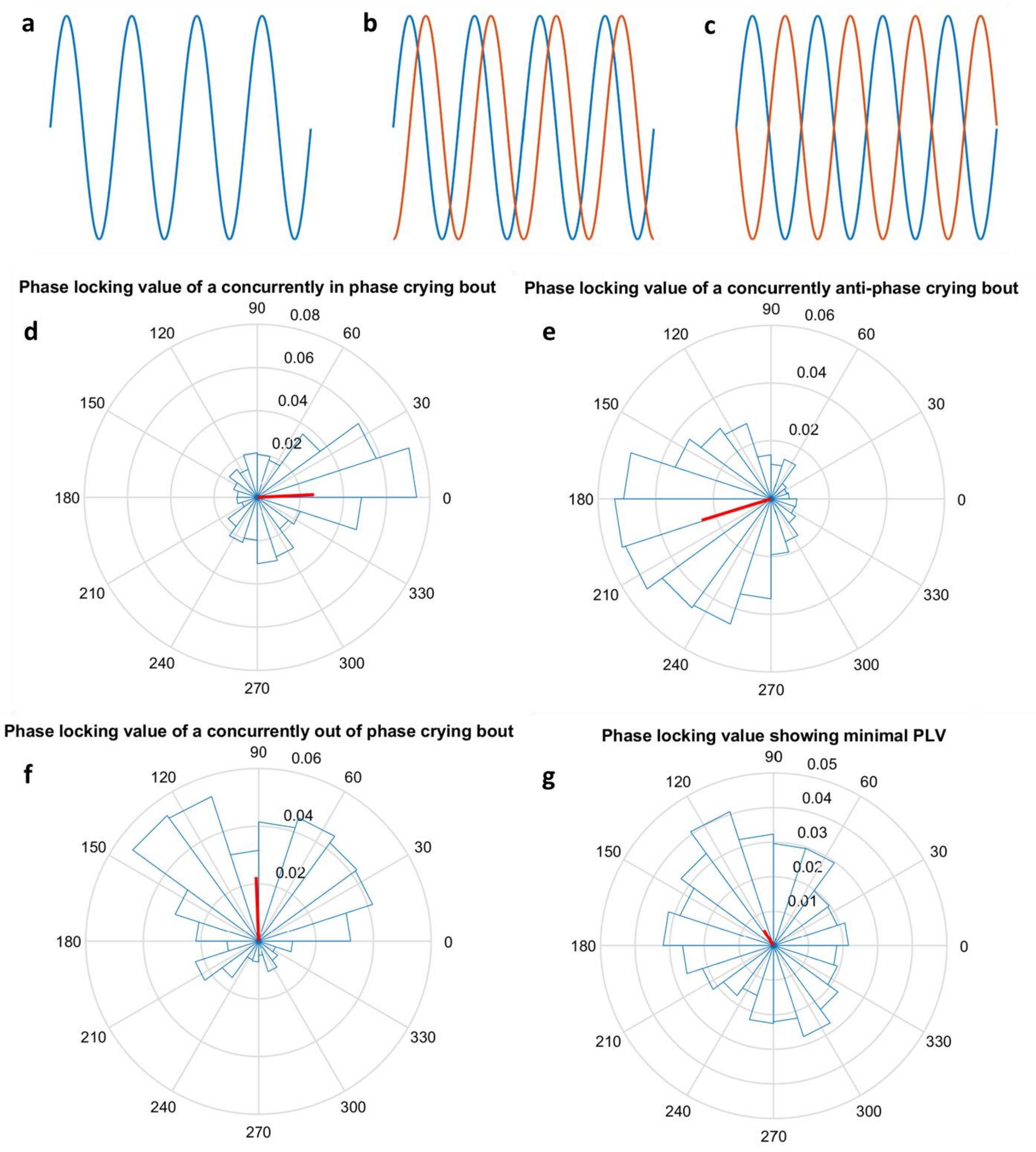
6.2.5.2 Crying-rocking bout behaviours vs bout length

Auditory oscillation frequency, rocking frequency across three axes, and the amount of rocking during the crying bout were chosen to see if these had an impact on the crying duration. To determine whether there was a change in these behaviours over time and to see if the behaviour at these points was greater than chance the data were split into deciles of the duration of the crying bout and ranked against a surrogate dataset produced with bootstrapping with replacement.

Next to determine if each of these behaviours had an impact on the bout length of the crying bout linear mixed effects models (LMEMs) were produced to model the level of the behaviour vs the bout length. To remove the potential confound that a longer cry gives more opportunity to reach the limits of the behaviours (e.g. a higher amount of rocking is possible with a longer cry), each crying bout was windowed into 10 second epochs up to 50 seconds. This allowed for investigation of whether auditory oscillation frequency in periods of time earlier or later in the crying bout had an effect on the auditory oscillation duration. Only crying bouts with a bout length longer than the maximum value for that window were included in that window. For example, a crying bout that was 35 seconds long would be included in the 1-10, 11-20 and 21-30 second windows but not the subsequent 30-40 and 40-50 second windows. For crying and rocking frequencies, the average frequency, range, minimum and maximum frequencies for each of these windows investigated and the value of these characteristics within each window was used in an LMEM as fixed effects variables vs the total bout length. Participant number was used as a grouping variable to check for individual differences that may skew the results. To preserve sample numbers, the rocking axes were collapsed into one. To correct for multiple comparisons a Benjamini-Yekutieli correction (Benjamini-Yekutieli, 2001) for multiple comparisons with dependency was used.

6.2.5.3 Phase characteristic analysis

To identify if there is an impact of phase characteristics on bout length, the phase locking value (PLV) of the actigraphy vs the audio envelope signals was calculated to investigate whether changes in frequency of the rocking behaviour were used to maintain a specific phase angle, e.g., to remain concurrently in phase, out of phase or antiphase, e.g. figure 6.3.



Phase locking value of a concurrently in phase crying bout

Phase locking value of a concurrently anti-phase crying bout

Phase locking value of a concurrently out of phase crying bout

Phase locking value showing minimal PLV

Figure 6.3 showing example diagrams of when signals are a) in phase, when one signal is high the other is high, when one is low the other is low; b) out of phase, where there is a phase shift along the time axis so that one signal is ahead of another; c) showing signals in antiphase, i.e. when one is high the other is low and vice versa. Example circular histograms of real data with mean PLV vector lengths in red are shown in d-g.

A longer mean vector length shows a stronger average phase angle and lower phase angle deviation, while spread out values cause a low vector length and a high phase angle deviation. Mean phase angles are d) concurrently in phase between signals, e) antiphase between signals, f) out of phase between signals and g) spread relatively evenly around the polar plot.

Due to the lower sampling rate for the actigraphy signals, for each axis the data were interpolated to the audio envelope frequency of 2400Hz using MATLAB's *interp1.m* function. Before calculating the PLV, both signals were z-scored.

PLV comparing actigraphy to audio envelope signals was calculated across each of the actigraphy axes for each crying bout. Similar to the analysis in 2.5.2 to remove the potential confound that PLV may be higher by chance with shorter crying bouts, the PLV was calculated in 10 second windows. PLV may be higher by chance with shorter crying bouts than longer ones simply because it is more likely that a small number of cycles align by chance than a large number of cycles aligning over a longer period to produce the same PLV result.

PLV was calculated in three steps, first create a phase angle timeseries for the sinusoid and the signal of interest by taking the Hilbert transform of the data and calculating the phase angle. Next the difference between the sinusoidal phase angle timeseries and signal of interest phase angle timeseries was taken, multiplied by the imaginary operator *i* to give a complex number and the exponential was determined for each value, this value was saved as *e*. PLV was calculated as the average exponential value in the timeseries.

The CircStat toolbox (Berens, 2009) was used to calculate the circular mean, PLV value as the average vector length and standard deviation. Then linear mixed effects models were used to assess the effect of these on bout length using a similar method as in section 2.5.2, using the subject number as a grouping variable to test for individual differences.

6.2.5.4 Cross correlations

To test if changes in an infant's crying frequency were met with similar changes in the caregivers rocking frequency a cross correlation was calculated.

For each crying bout in the crying-rocking condition the frequencies calculated in 6.2.4.1 and 6.2.4.3, resulting from the cycle-by-cycle analysis were used. The frequencies were found using the full sampling rate of each of the data streams, which were then binned into one second segments. The actigraphy axes were collapsed by taking the average rocking frequency for each second of activity. The average crying frequency for each second of audio activity was also calculated. The rocking and auditory oscillation frequency data were each normalised by removing the mean of that data. Where there were no oscillations found these were replaced with zeros. The Matlab *xcorr.m* function was used either with a max lag time of 10 seconds or if the crying bout was shorter than 2N-1 then with no max lag parameter.

To test for significance, 10,000 random permutations was calculated by shuffling the actigraphy and audio data relative to one another. Then the above cross correlations were completed. At each lag point the original data was ranked against the permutation data to determine the significance level.

6.2.5.5 Turn taking analysis

To test if any changes in frequency over time in the cross-correlation analysis (2.5.3) were due to a turn taking structure of crying bursts and rocking bursts, cries in the crying-rocking condition were also subject to a turn taking analysis similar to Yoo, Bowman & Oller’s**,** 2018 study, which investigated the lag between infant and caregiver vocalisations.

The onsets and offsets of rhythmic bouts in the infant audio and actigraphy signals were identified during cries in the crying-rocking condition using the cycle-by-cycle analysis (Cole & Voytek, 2019), similar to sections 2.4.1 and 2.4.2.2. The peaks identified as oscillations were used to identify the timepoints, with the start and end of each oscillation calculated as half of the measured period (given in samples) away from the peak sample. Any oscillations that were within 500ms of one another were grouped together into rocking or crying oscillatory bouts.

To calculate the lag times between cry oscillation offset and rocking onset, similar to Yoo, Bowman & Oller, (2018), who investigated infant cry offset to caregiver cry onset, the following rules were applied: the rocking oscillation must start after the start of the crying audio oscillation; the rocking oscillation must not start more than 5 seconds after the crying oscillation has ended; where there is more than one eligible rocking oscillation that could be associated with a crying oscillation then the first rocking oscillation is assigned; once an audio and rocking oscillation have been paired they cannot be paired with any other oscillation.

To generate surrogate data for each crying bouts, the onsets and offsets of the crying and rocking oscillations were shuffled to a random starting sample within the data. The durations of each crying and rocking oscillation were kept the same throughout. The random shuffled data were checked to ensure that within the crying and rocking signals the bursts of crying and rocking oscillations were not overlapping, and were at least 500ms apart, similar to the real data. The same lag time calculation as above was conducted as above. The real data were ranked against the surrogate data to generate a p value.

6.3 Results

6.3.1 Descriptive statistics

6.3.1.1 Amount of crying and rocking per condition

Analyses were run to check the amount of crying and rocking was correct for each condition, and to check that the matched control conditions had similar level of crying or rocking when compared to the recorded data. Further analyses investigated the time to start rocking, the auditory oscillation duration after rocking starts and the rocking bout duration to check for group wide differences.

6.3.1.1.1 Amount of crying

The level of crying as detected per bout by the cry detector was 100% in the crying conditions and 0% in the non-crying conditions (SM6.4). Bout length median and standard deviations per condition in seconds were crying-rocking M=14, std=28.31; not crying-rocking M= 14, std, 28.37; crying-not rocking M=10, std=12.59; and not crying-not rocking, M=10, std=12.62. A Kruskal-Wallis test showed that there were significant differences between conditions (Chi-square = 374.93, df = 3, $p < 0.001$), multiple comparisons confirmed that there were significant differences between rocking conditions (both $p < 0.001$), showing a longer bout length for rocking conditions, but no significant differences between crying conditions (both $p > 0.05$).

6.3.1.1.2 Amount of rocking

The amount of rocking as determined by the amount of oscillatory activity in the actigraphy data was between 0-100% of the crying bout (SM6.4). Number of seconds of rocking, median and standard deviations were: crying-rocking M=5, std=7.67 seconds;

crying-not rocking, M=0 (mean 0.115), std = 0.43 seconds; not crying-rocking, M=4, std=6.06 seconds; and not crying-not rocking, M=0 (mean 0.071), std =0.31 seconds. A Kruskal-Wallis test showed that there were significant differences between conditions (Chi square= 4236.31, df=3, $p < 0.001$), multiple comparisons confirmed that there were significant differences between the rocking and not rocking conditions (both $p < 0.001$), showing that there was more rocking in the rocking conditions, but not between the crying conditions (both $p > 0.05$).

6.3.1.1.3 Time to start rocking

The median and standard deviation time between the start of the crying bout to the start rocking in seconds for each condition was: crying-rocking M=4, std=10.41 seconds and not crying-rocking M=4 seconds std=13.56. A Wilcoxon rank sum test was used to test for significant differences. No significant different between the time to start rocking in the crying and not crying conditions $p > 0.05$.

6.3.1.1.4 Auditory oscillation duration after rocking starts

The median and standard deviations of the duration of the crying bout once rocking had started were: crying-rocking M=10, std=25.2 and not crying-rocking M=10, std=22.02. There was no significant difference between the crying and not-crying conditions $p > 0.05$.

6.3.1.1.5 Rocking bout duration

The median and standard deviations of the rocking bout duration in seconds per condition were: crying-rocking N=3078, M=2, std=2.89; not crying-rocking, N=2468, M=2, std=3.15. A Wilcoxon rank sum test showed that there was no significant difference ($p > 0.05$).

6.3.1.1.6 Descriptives summary

Sections 3.1.1.1 to 3.1.1.5 demonstrate that there are no differences between the experimental recorded crying conditions and the matched not-crying conditions suggesting that the matched not crying condition was a fair representation of the crying condition. No significant differences were found in the amount of rocking, time to start rocking, auditory oscillation duration after rocking had started and rocking bout duration.

Expected differences were found in the amount of rocking between the rocking and no rocking conditions. Similarly, expected differences between the crying and not crying conditions were found for the amount of crying. The only other significant difference that was was the amount of crying between crying and not crying conditions.

6.3.1.2 Dominant audio and actigraphy frequencies

6.3.1.2.1 Crying frequencies

This analysis aimed to investigate how the frequency (Hz) of infant crying oscillations changed across crying conditions. Median and standard deviation crying frequencies for each of the four conditions shows: crying-rocking, M=1.79Hz, std=0.61; crying-not rocking, M=1.76Hz, std=0.63; not crying-rocking, M=1.98Hz, std=0.7 not crying-not rocking, M=2.02Hz, std=0.79. A Kruskal-Wallis test showed that there were significant differences between the groups (Chi square=52.13, df=3, $p < 0.001$). Multiple comparisons showed that there were significant differences between the crying/not crying conditions (both $p < 0.001$) showing that the dominant audio frequency was significantly lower during crying conditions, while there were no significant differences between rocking/not rocking conditions (both $p > 0.05$).

6.3.1.2.2 Rocking frequencies

This analysis aimed to investigate how the frequency (Hz) of rocking oscillations changed across rocking conditions and across each axis. Median and standard deviations for each of the rocking conditions, for each of the axes showed: crying-rocking-up down, N=10,743, M=2.86Hz, std=1.08; crying-rocking-forward back, N=7798, M=2.67, std=1.13, cry-rocking-left right, N=5931, M=2.67, std=1.25; not crying-rocking-up down, N=9547, M=2.67, std=1.1; not crying-rocking-forward back, N=6748, M=2.5, std=1.14; not crying-rocking-left right, N=5317, M=2.5, std=1.19. A Kruskal-Wallis test showed that there were significant differences between conditions (Chi square=239.96, df=5, $p < 0.001$). Figure 6.4 shows the relative difference in means for each of the conditions and Table 6.1 shows the significance values between groups, with blue shading showing where a condition in the row is significantly higher than the condition in the column, purple shading showing where a condition in the row is significantly lower than the condition in the column.

Overall, the results showed that crying conditions had a higher rocking frequency than not crying conditions for all axes ($p < 0.001$). In both the crying and not crying conditions the up-down axis had a significantly higher rocking frequency than forward-back or left-right axes in the same condition ($p < 0.001$).

Other findings showed that the crying up-down axis had a higher rocking frequency than not crying forward-back and left-right axes ($p < 0.001$). Not crying up-down had a higher rocking frequency than crying left-right ($p < 0.01$). Crying forward-back axis had a higher rocking frequency than not crying left-right ($p < 0.001$) and crying left-right had a higher rocking frequency than not crying forward-back ($p < 0.001$). All other interactions were found to not be significant ($p > 0.05$).

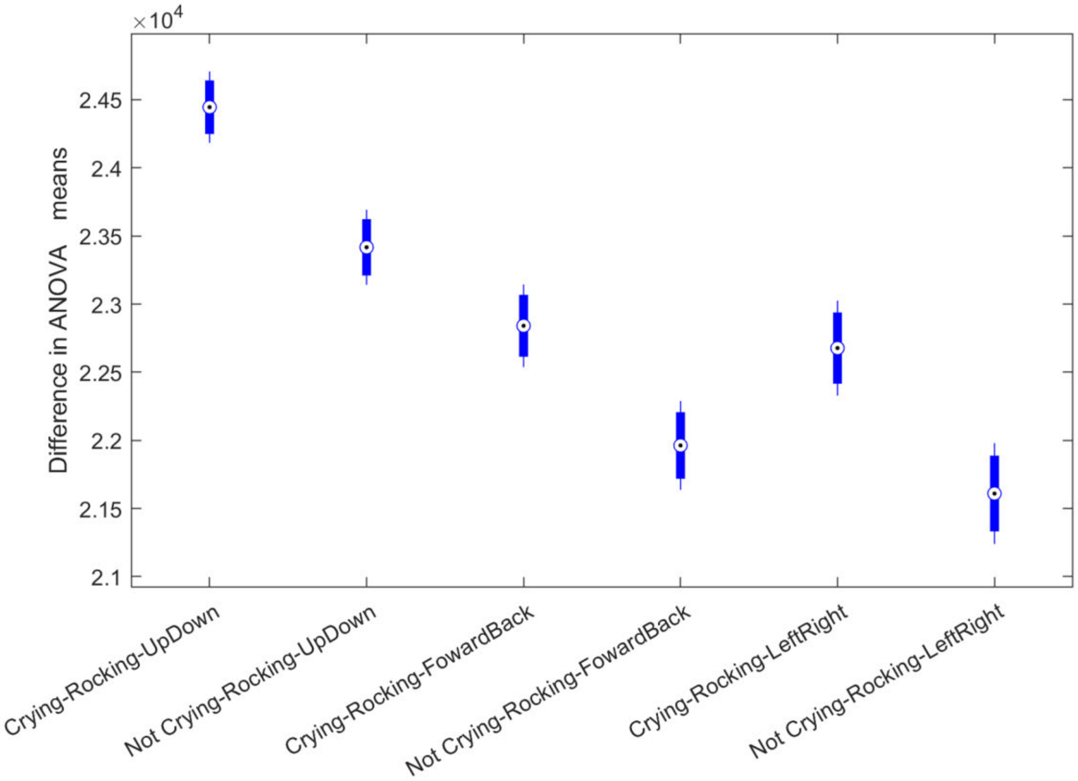
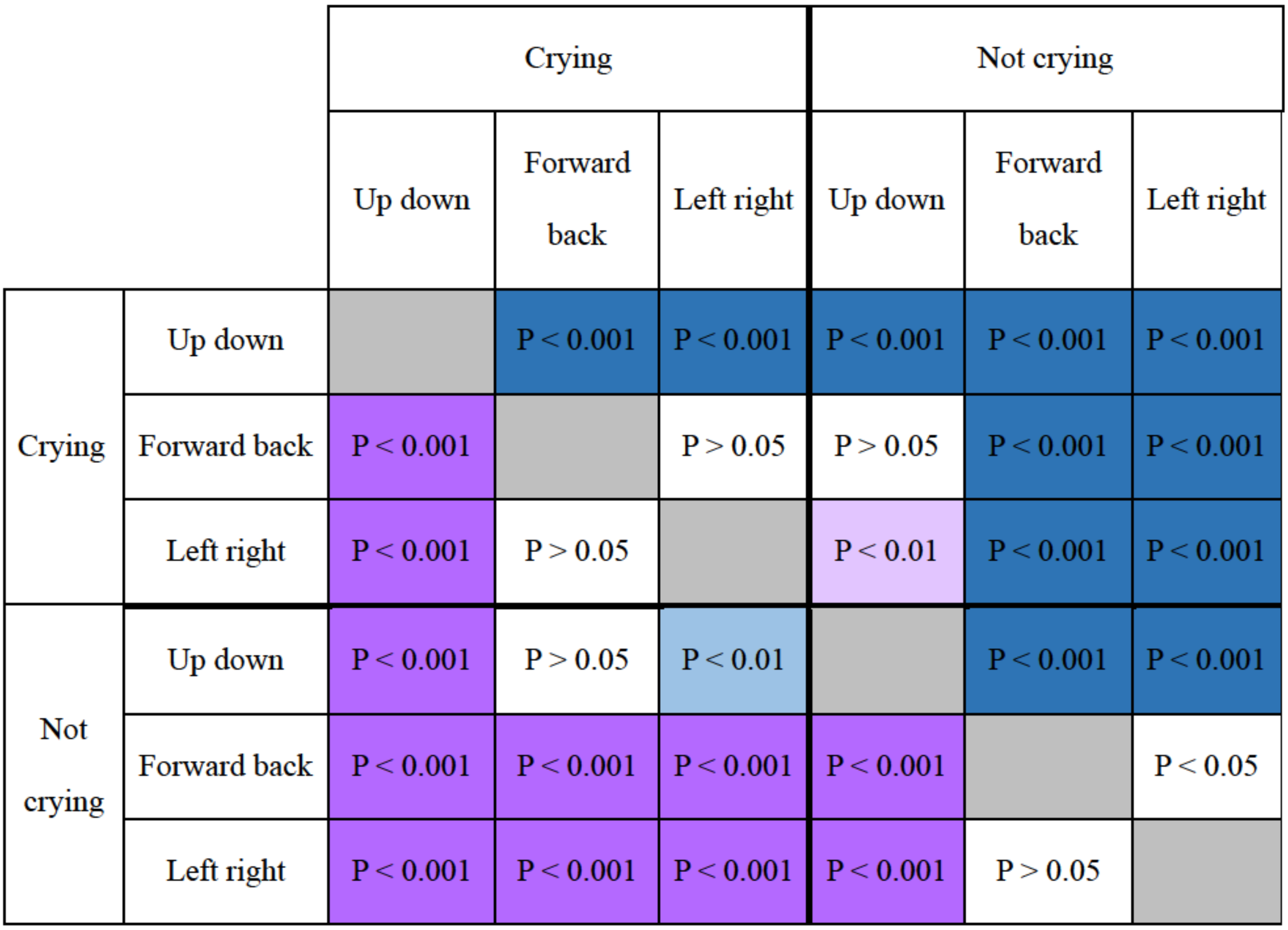


Figure 6.4 Showing the difference in ANOVA means for rocking frequencies for crying, not-crying conditions by axis.

Table 6.1. Showing the p values from multiple comparisons between crying and not crying conditions, by axis. Blue shading shows row conditions that have a significantly higher average rocking frequency than column conditions, while purple shading shows row conditions that have a significantly lower average rocking frequency,



		Crying			Not crying		
		Up down	Forward back	Left right	Up down	Forward back	Left right
Crying	Up down		P < 0.001	P < 0.001	P < 0.001	P < 0.001	P < 0.001
	Forward back	P < 0.001		P > 0.05	P > 0.05	P < 0.001	P < 0.001
	Left right	P < 0.001	P > 0.05		P < 0.01	P < 0.001	P < 0.001
Not crying	Up down	P < 0.001	P > 0.05	P < 0.01		P < 0.001	P < 0.001
	Forward back	P < 0.001	P < 0.001	P < 0.001	P < 0.001		P < 0.05
	Left right	P < 0.001	P < 0.001	P < 0.001	P < 0.001	P > 0.05	

6.3.2 Crying-rocking bout behaviours vs bout length

6.3.2.1 Changes in behaviour over time

These analyses aimed to investigate whether the crying and rocking frequencies change over the crying bout by investigating the trend across deciles of data and by comparing each decile to a bootstrapped surrogate dataset. Figure 6.5 shows the change over time for rocking (across all three axes) and crying frequencies found during the cycle-by-cycle analysis. The results for the auditory oscillation frequency analysis show a clear decreasing trend in auditory oscillation frequency over time (trendline y = -0.013x +

2.1, $R^2=0.67$). Data for each decile were ranked in a two-tailed bootstrapped ranking, which showed that the auditory oscillation frequency in the 9th decile was significantly lower than chance ($p < 0.025$).

For the rocking frequency analyses, a limited trend was observed for each of the axes over time ($R^2 < 0.25$). In the bootstrapping analysis rocking frequencies were found to peak first in the up-down axis in the 3rd decile ($p < 0.025$), followed by the 5th decile in the forward-back axis and 7th decile left-right axis ($p < 0.025$). Forward back rocking frequency was found to be significantly lower than chance in the 8th and 9th deciles ($p < 0.0005$ and $p < 0.025$).

Overall, this shows that auditory oscillation frequency decreases over the crying bout, and that rocking frequency across different axes are significantly higher than chance at different times during the crying bout, starting with up-down rocking, followed by forward-back rocking and finally left-right rocking. The rocking for each axis then decreases back to within chance or in the case of forward-back rocking, less than chance rocking frequencies. For clarity, this does not suggest that rocking across all three axes are employed in sequence during a single crying bout, just that the type of rocking peaks on average at different times during the crying bout.

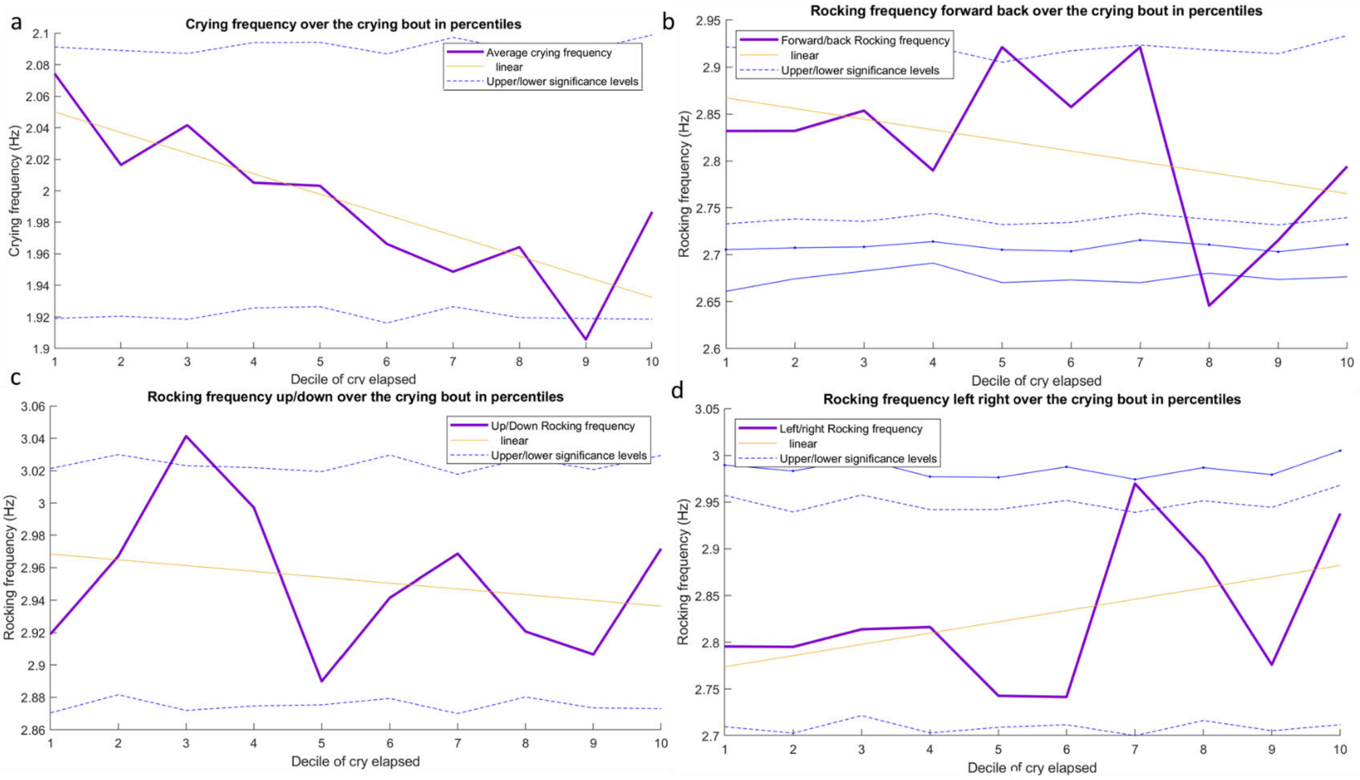


Figure 6.5. Average crying (a) and rocking (b-d) frequencies of each decile of a crying bout. Blue dashed, dotted-dashed and solid lines show upper and lower significance levels at the $p < 0.025$, $p < 0.005$ and $p < 0.0005$ levels. Yellow lines show the linear lines of best fit for each graph.

This analysis investigated whether the amount of rocking changes over the crying bout by investigating the trend across deciles of data and by comparing each decile to a bootstrapped surrogate dataset. Figure 6.5 shows the amount of rocking in deciles over the crying bout, there was a positive trend shown ($y = 0.00358x + 0.27$, $R^2 = 0.42$). In the two-tailed bootstrapping analysis there was significantly less rocking in the 1st ($p < 0.0005$), 3rd ($p < 0.025$), 4th ($p < 0.025$) and 6th ($p < 0.005$) deciles.

These results suggest that despite the proximity between infants and their caregivers being consistently below 30cm during the entire crying bout, the caregivers are not likely to begin rocking immediately, and rocking is more likely to occur in the second half of the crying bout.

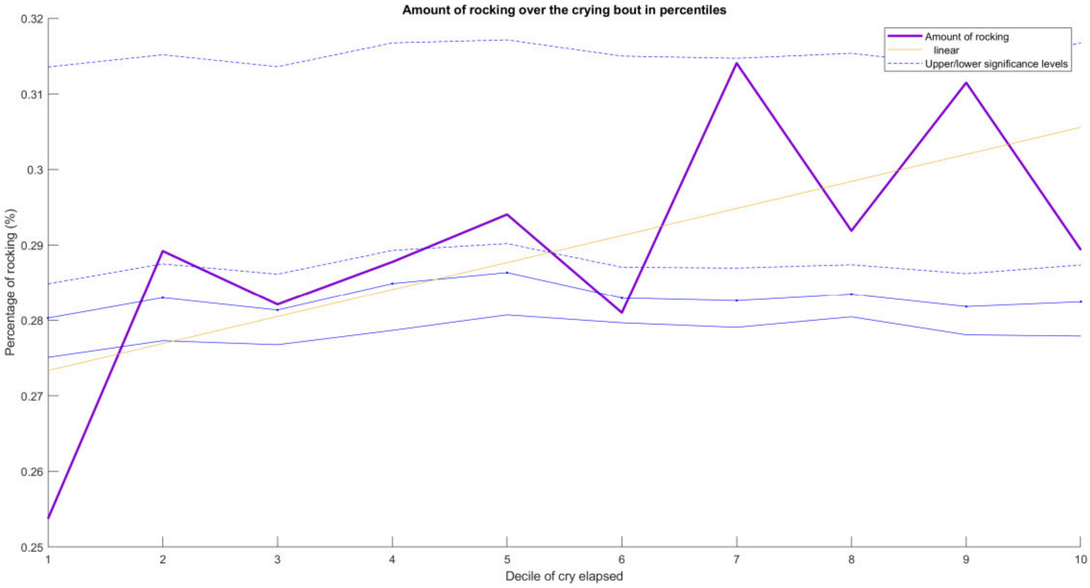


Figure 6.6. Average amount of rocking of each decile of a crying bout. Blue dashed, dotted-dashed and solid lines show upper and lower significance levels at the p < 0.025, p < 0.005 and p < 0.0005 levels. Yellow lines show the linear lines of best fit for each graph

6.3.2.2 Windowed rocking behaviour vs bout length

To investigate whether infant crying behaviour was a major factor in bout length, the data were split into 10 second windows, between 1-50 seconds (for further detail see section 2.5.2). The analysis aimed to examine whether features of auditory oscillation frequency: average frequency, frequency range, minimum and maximum frequencies, in each of the 5 windows had an impact on bout length, for example does a higher auditory oscillation frequency early in the crying bout relate to a longer auditory oscillation duration?

LMEMs were used to examine each feature of auditory oscillation frequency, across each time window with bout length. For each LMEM the frequency feature was a fixed effects variable, and participant number was used as a grouping variable to check for individual differences.

After correction for multiple comparisons using the Benjamini-Yekutieli adjustment, the results showed no significant effect for each of the features of auditory oscillation frequency listed above on crying length.

Similar LMEMS were conducted using the same frequency features of rocking frequency and the same procedure as above. After correction for multiple comparisons using the Benjamini-Yekutieli adjustment, the results showed no significant differences for each of the features, over each time window.

Finally, a similar LMEM investigating the amount of rocking in each 10 second window also showed no significant effects on the total auditory oscillation duration.

Overall, the results suggest that a particular auditory oscillation frequency at different timepoints during the crying bout is not related to the crying length. Likewise, rocking infants in a specific way at a specific time during the crying bout has not been shown to affect crying length. Finally, changing the amount of rocking at specific times during the crying bout has not been shown to affect the crying length.

6.3.3 Phase characteristics analysis

6.3.3.1 Changes in phase characteristics over time

These analyses investigated the phase characteristics between crying amplitude envelope and rocking actigraphic envelope to examine how concurrent dynamic changes in one signal impact responses in another. PLV between two signals can remain consistently high even when the frequency of one signal increases, providing there is a corresponding increase in frequency in the second signal. Phase angle deviation is a measure of circular variance, which shows the spread of phase angle differences across time, a higher phase angle deviation suggests a wider variance between the PLV of the two signals over time.

This analysis investigated changes in PLV vector length and angular deviation averaged over deciles of the crying bout and each decile was ranked against a bootstrapped surrogate dataset to examine whether the decile was greater than or lower than chance. Across each of the three axes, figure 6.7 shows the change in PLV (a, c, e) and angular deviation (b, d, f), over each of the crying bout deciles. Data for each decile were ranked in a two-tailed bootstrapped ranking, which showed that for all deciles, PLV was significantly lower than chance ($p < 0.0005$) and that angular deviation was significantly higher than change ($p < 0.0005$).

A significantly lower PLV vector length over each decile, for each axis, suggests that there was not concurrent tracking between infant auditory oscillation frequency and caregiver rocking frequency. Similarly, a significantly higher phase angular deviation suggests a significantly higher than chance variance in phase angle differences between the two signals. Overall, these two findings suggest there is not evidence for concurrent tracking between the crying and rocking frequencies.

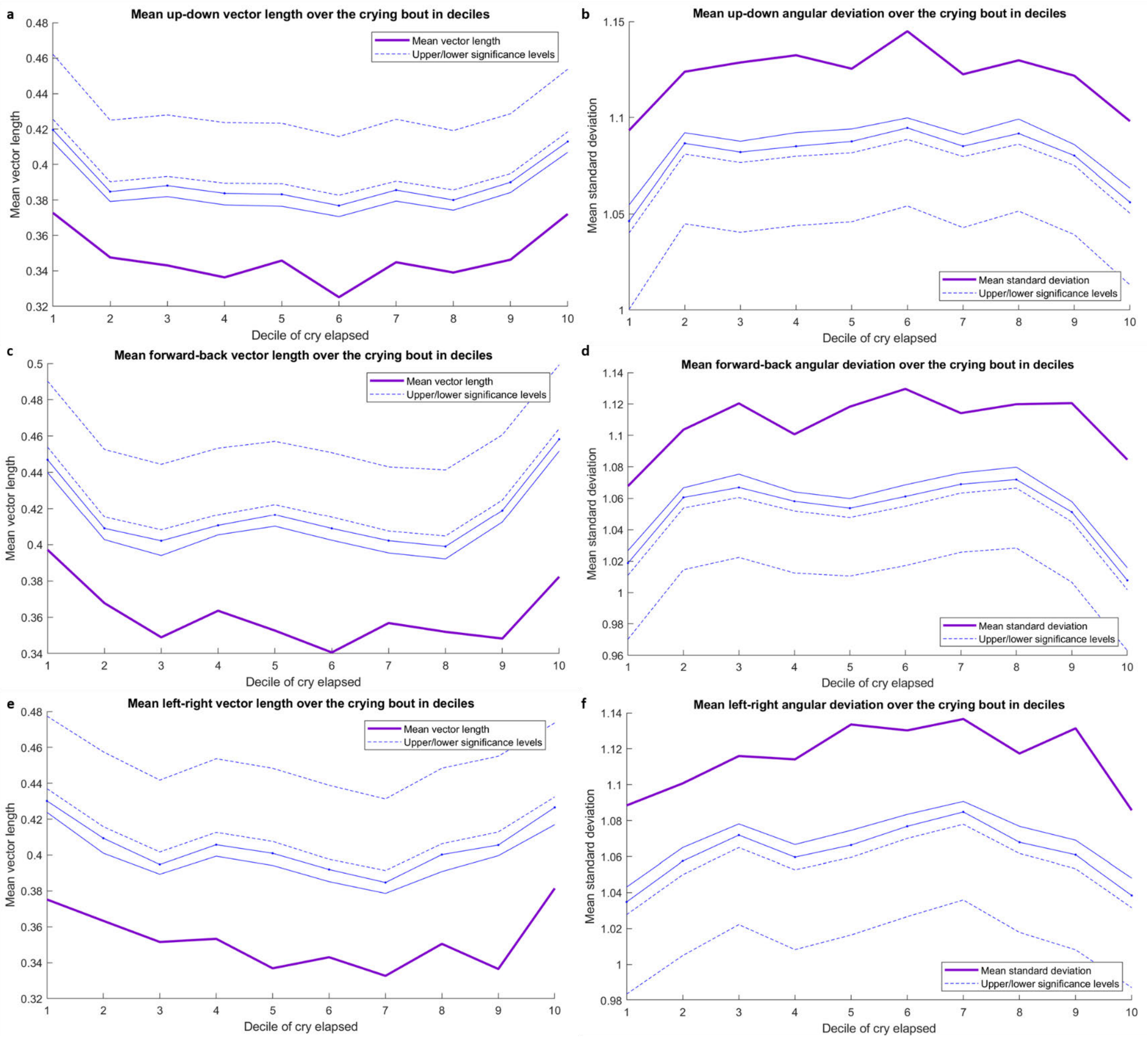


Figure 6.7. Showing the average phase locking value of each decile of each axis a) up-down, c) forward-back and showing the angular deviation in axes b) up-down, c) forward-back and e) left-right. Blue dashed, dotted-dashed and solid lines show upper and lower significance levels at the $p < 0.025$, $p < 0.005$ *and* $p < 0.0005$ *levels.*

6.3.3.2 Windowed PLV characteristics vs bout length

As a high PLV score over the entire crying bout may be difficult, this analysis investigated whether periods of dynamic concurrent tracking of crying and rocking frequencies for discrete periods of time affects the overall crying length, LMEMs were used to test whether the phase relationship between infant crying oscillations and caregiver rocking oscillations impacted the length of infant cries, over 5 windows, 10 seconds longwith the participants grouped to test for individual differences. For each of the 10 second windows PLV was shown to have a significant negative effect on crying length: 1-10 seconds (*$\beta=-0.35, t=-8.02, p < 0.001$); 11-20 seconds ($\beta=-0.44, t=-10.24, p < 0.001$); 21-30 seconds ($\beta=-0.41, t=-9.93, p < 0.001$); 31-40 seconds ($\beta=-0.44, t=-8.88, p < 0.001$); 41-50 seconds ($\beta=-0.19, t=-3.81, p < 0.001$).*

The other phase differences: mean phase angle, phase angle skewness, mean rocking frequency and number of rocking oscillations per bout were found to not have a significant effect ($p > 0.05$).

This suggests that some features of rocking have no discernable impact on crying length, including the position of the mean phase angle, i.e. whether a caregiver chooses to be in phase, out of phase or antiphase relative to the infant crying oscillations (see figure 6.3); how skewed the phase angles are, i.e. how symmetrical the phase angles are around the mean phase angle; the mean rocking frequency; and the number of rocking oscillations in a single window were all found to not have a significant impact on the crying length.

However, having a high PLV vector length, i.e. showing concurrent tracking of phase angles through dynamic changes in rocking frequency, was found to have a significant impact, reducing crying length compared to crying bouts with low PLV. This was found across all time windows tested, suggesting that tracking phase angles between crying and rocking oscillations at any time during a crying bout is likely to reduce the crying length.

6.3.4 Cross correlation of crying and rocking frequencies

As previous analyses (section 3.3.1) had demonstrated that caregivers in general do not concurrently track rocking frequencies with crying frequencies, a cross correlation of the crying and rocking frequencies with a maximum lag of 10 seconds was used to investigate if changes in an infant's auditory oscillation frequency were met with changes in the caregiver's rocking frequency after a lagged period of time.

This was ranked against 10,000 permutations of shuffled data. The median duration and standard deviations of oscillations were: real crying oscillations (M=2.17, std=1.57), real rocking oscillations (M=2.21, std=1.91), shuffled crying oscillations (M=1.64, std=1.49), shuffled rocking oscillations (M=1.63, std=2.64).

The results are shown in figure 6.8 and demonstrate that there was a significant cross correlation between audio and actigraphy signals at 3 ($p < 0.025$), 6 ($p < 0.025$) and 7 ($p < 0.025$) seconds. Showing that changes in rocking frequency are lagged by 3, 6 and 7 seconds after changes in auditory oscillation frequency.

This suggests that changes in rocking frequency lag behind changes in auditory oscillation frequency by a number of seconds. That there are two peaks at 3 and 6-7 seconds suggests that there may be a turn taking structure to crying and rocking oscillations, with changes in auditory oscillation frequency being met with changes in rocking frequency 3 seconds later that are in turn met with further changes in crying and rocking frequency a further 3 seconds later.

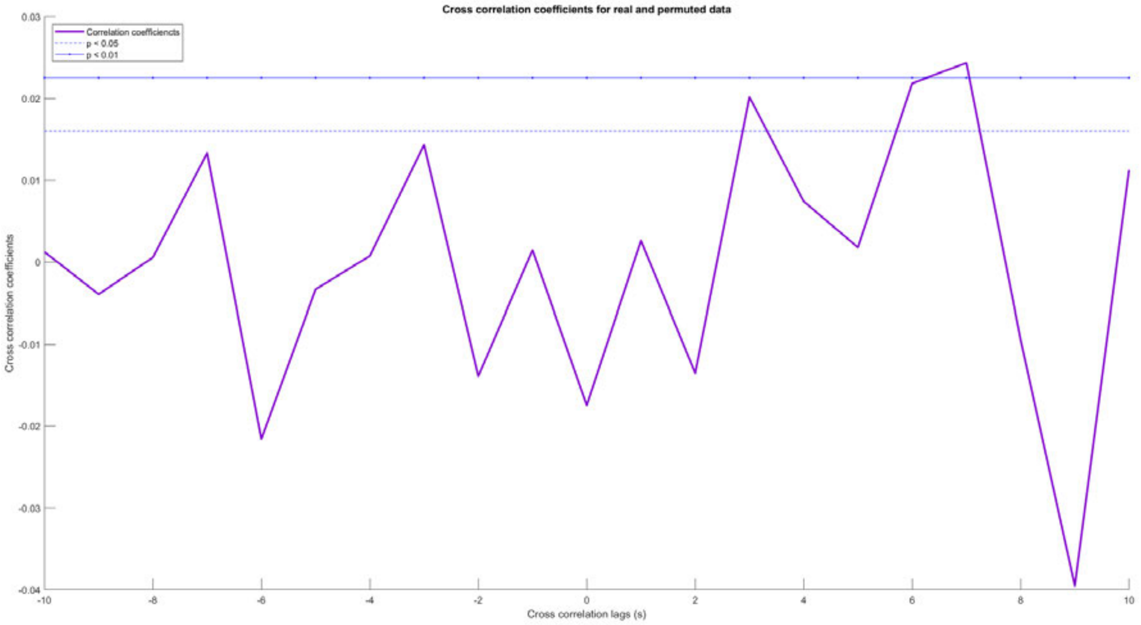


Figure 6.8. Cross correlation coefficients for each of the cross correlation lag times between -10 and 10 seconds. The purple line shows the test data, the dashed and dotted-dashed blue lines show the threshold of significance at $p < 0.05$ and $p < 0.01$ level.

6.3.5 Turn taking analyses

To investigate whether there was a turn taking or overlapping structure to the crying and rocking oscillations during the crying bout a turn taking analysis similar to Yoo et al., (2018) was performed. The crying-rocking data showed a positive skewed lag with a mean of 0.97 seconds between crying oscillatory offset and rocking onset. This was shown to be significantly less ($p < 0.05$) than the average shuffled crying offset to crying onset of 1.39 seconds. Figure 6.9 and 6.10 shows the distributions of lag times as well as the cry and rocking oscillation durations for both real and shuffled data respectively.

Overall, this data shows that there is a turn taking structure to the crying and rocking oscillations within a crying bout, when added to with median cry oscillation shown in section 3.4 of 2.17 seconds this shows a close to 3 second duration between crying onset and rocking onset, which supports the results shown in section 3.4 that showed a cross correlational change in rocking frequency with a lag of 3 seconds.

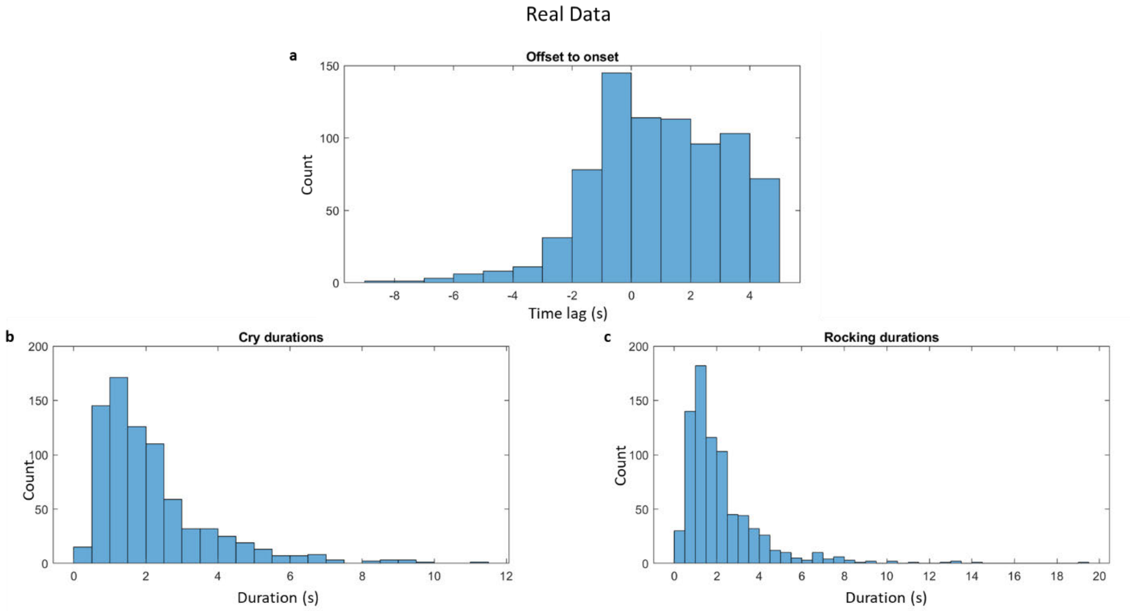


Figure 6.9 a histogram of real data points showing a) the lag time in seconds between the offset of crying oscillations and the onset of rocking oscillations, b) the durations of crying oscillations and c) the durations of rocking oscillations.

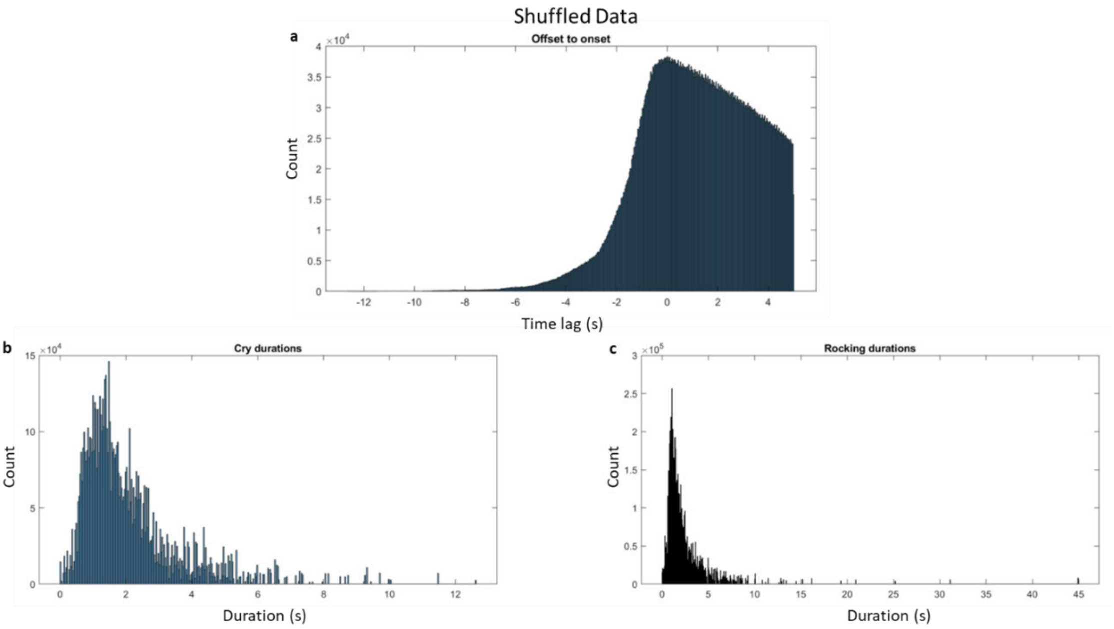


Figure 6.10 a histogram of shuffled data points showing a) the lag time in seconds between the offset of crying oscillations and the onset of rocking oscillations, b) the durations of crying oscillations and c) the durations of rocking oscillations.

6.4. Discussion

6.4.1 Overview

In this study we examined naturalistic crying and rocking behaviour in day long home recordings with mothers and 5-7 month old infants. Recordings were made with custom wearable devices that measured continuous audio, actigraphy, proximity and GPS. This study has reported on natural crying and soothing episodes through a typical day; typical crying and rocking frequency details and the impacts on bout length; phase difference characteristics of crying and rocking behaviour; cross correlations showing changes in crying and rocking frequency during crying bouts; and a turn taking analysis. To the best of our knowledge no previous studies have investigated these topics in a natural setting.

6.4.2 Descriptives of crying and rocking bouts

A significantly lower auditory oscillation frequency was found between crying and not crying conditions, which suggests that infants were changing their vocalisation rate to exhale and produce longer vocalisations. Rocking frequencies were also shown to be higher for all crying vs not crying conditions, suggesting that caregivers were more likely to rock their infants faster to help soothe them. Interestingly, all conditions where rocking occurred happened at a faster rate than previous mechanical rocking studies.

Rocking in the up-down axis was shown to be significantly faster than in the forward-back and left-right conditions, potentially supporting previous research (Bynre & Horowitz, 1981; Nishimura, Kanakogi & Myowa-Yamakoshi, 2016) which has shown that parents tend to rock their infants in an up-down motion when attempting to elicit a “bright alert” state or during uncomfortable periods. However, the infants’ orientation was not taken into account (see limitations), so infants could have been rocked in a supine position, which would have been different to these studies.

There were no significant differences seen between the forward-back and left-right conditions.

6.4.3 Crying and rocking behaviours vs bout length

Data first split into deciles showed that there was a clear negative trend in crying oscillation frequencies over the crying bout, with a significantly lower than chance response in the 9th decile. A lower auditory oscillation frequency is likely caused by an increased amount in the inhalation-exhalation cycle, similar to previous studies (Zeskind et al., 2013) it is likely that infants are elongating their exhalation during the later part of their cries, maximising sound production. By better coordinating their respiratory anatomy, infants would be able to exhale for longer, which would decrease the frequency on their crying rhythm. This trend is also supported by the sharp uptick in the 10th decile, as at this point in time the crying bout would be being resolved.

Changes in rocking frequency over time did not show any consistent trends in the frequencies of rocking used for each axis. However, these results do point to caregivers trying a range of strategies with significantly higher than chance rocking frequencies seen for each of the axes at different points during the crying bout.

Results also showed that the amount of rocking increased over time, with a positive trend. Rocking was shown to be lower than chance during the first half of the crying bout, and while there were no deciles that were significantly higher than chance, there were two that showed a positive trend. This makes sense, as we would expect that caregivers would respond to their infants more as the crying bout progressed. This also supports previous lab based claims that caregivers are likely to respond to infant crying after some time has passed (Zeifman, 2004).

Finally, there were no significant results shown with the windowed analyses, which may reflect a lack of statistical power (see limitations).

6.4.4 Phase characteristic analyses

PLV vector length showing the strength of the mean phase angle between crying bout audio and actigraphy was significantly lower than chance for all real vs shuffled deciles, while variance as measured by angular deviation was higher in all deciles. This suggests that caregivers may specifically try to keep rocking rhythms from forming consistent phase differences relative to infant crying oscillations. This is supported by the larger than chance variance in phase differences seen, which may suggest that infant crying oscillations and caregiver rocking oscillations are continually changing on a second by second level. Further evidence to support these changes in rocking frequency being driven by the caregivers in this study could be drawn from the results seen in section 3.2.1 where there was a clear trend for crying rhythms over the course of the crying bout, but the adult rocking rhythms were much more varied over the three rocking axes, despite the amount of rocking increasing over the crying bout.

However, results from the windowed PLV analysis showed that this technique of changing rocking frequency may have a detrimental effect on bout length. Caregivers who were able to keep a more constant phase difference between the infant crying oscillations and rocking rhythms were consistently shown to be associated with shortened bout lengths. This opens the door to further research in the lab and home settings investigating how different rocking techniques could be used to match infant rhythms either in phase, out of phase or anti phase, plus research into how multiple rhythm frequencies can be best used.

6.4.5 Cross correlation and turn taking analyses

The results of the cross correlation between crying and rocking frequencies, as identified by the cycle-by-cycle analyses demonstrated that there was a significant cross correlation seen at plus three, six and seven seconds. This showed that caregiver rocking frequency was likely to match changes in auditory oscillation frequency three, six and seven seconds after changes in the auditory oscillation frequency. This may demonstrate a dynamic response from caregivers to increase or decrease the rate at which they rock their infants in response to the changes in crying behaviour, with a short delay.

The results of the turn taking analysis showed that there was a positive lag between crying oscillation offset and rocking onset. This is potentially contrary to previous findings in studies such as Yoo et al., (2018), which showed that caregivers responded to negative vocalisations by overlapping with their infants vocalisations, causing a negative lag between infant vocal offset and caregiver vocal onset. However, the current results support the cross-correlation results, as the average crying oscillation plus the lag time was around three seconds. Suggesting that the cross-correlation results may be due to caregivers taking turns to soothe their infants with rocking behaviour.

6.4.6 Limitations

One fundamental limitation of this study is that, as discussed in the introduction, there are a wide variety of cry types, reasons for infant crying, differences in parent responses and other factors (e.g. Truby & Lind, 1965; Zeskind, 2013). Despite collecting hundreds of hours of home recordings, similar to previous studies there are a comparatively low amounts of crying (e.g. James-Roberts & Halil, 1991; Micheletti et al., 2022; Vermillet et al., 2022). Therefore, we decided not to split the dataset further than necessary so we could investigate the core questions for the present study. Future research using day long home recordings investigating responses to intensity of cries, crying intentionality and splitting parents by parenting style would further our understanding of infant crying and caregiver soothing behaviour.

Similarly, while this study did investigate the direction of rocking relative to the infant, we did not separate rocking episodes by the infant's orientation. This was primarily because splitting the conditions further would have led to a small number of crying episodes in each condition. It would also be interesting to see whether parents used a change in infant orientation during the crying bout to elicit a soothing response. As seen in previous studies (see section 1.2.1), the direction and orientation that infants are rocked has been shown in the lab to have an impact on infant response. It wasn't feasible to get enough crying bouts to split these conditions further, especially as orientation may change throughout the crying bouts, however it would be valuable data to collect and report on in future.

Given the natural setting of this study, every precaution has been taken to try to cleanly identify periods of crying and rocking. This study filtered data based on GPS as outdoor activity was found to have too much variation, data were filtered based on proximity to try to normalise the likely positions of caregiver and infant and to ensure that other caregivers weren’t providing the soothing behaviour, which would have been a confound. To ensure that audio included crying, a machine learning model was used. To try to ensure that actigraphy oscillations weren’t spurious, two consecutive seconds of rocking were required, which would have meant a minimum of four to six rocking cycles with the average rocking frequency. Despite this, audio that has been identified as crying may also include covocalisations from the caregiver(s), other infants or other sources (e.g. TV or radio) which would have affected the audio envelope. The actigraphy data similarly could have other artefacts for example the infant may have moved sufficiently to cause their own rocking rhythm, however this is less likely. Finally, other soothing methods were not ruled out, including feeding, screens, or other methods. This study was not able to place cameras in every room, so it would have been difficult to conclusively rule out other soothing methods and visually verify rocking. Similarly, we were not able to listen to all raw audio recordings given the large amount of data that was produced.

As this study was primarily focused on behaviour that impacted bout length, it was difficult to avoid the confound that longer cries are more likely to include a wider range of the behaviour being investigated (e.g., there is more opportunity for rocking in a longer cry). This led to a number of difficult decisions, as we couldn’t directly investigate behaviour during the entire cry vs bout length.

For analyses that used data split into deciles, due to the nature of how shorter cries were stretched and longer cries were compacted to fit the deciles, there was an impact on the representation of shorter cries vs longer cries on the analyses. For example, a cry that was 5 seconds long would be included in 10 decile bins. The alternative was to have a "lumpy" distribution of the data, which would have meant that some bins would receive more data than others (e.g. at the 33%, 50%, 66% points) misrepresenting the results with peaks and troughs simply because of the split of the data. We believe that producing the decile analyses in using this methodology was the fairest way to calculate the change over time.

By windowing into ten second chunks, we necessarily removed a lot of data that didn't meet the minimum length for even the first window. This is necessary as it would have unfairly biased the first window for all features that were investigated. However, this came with the limitation that it reduced the potential power for statistical analyses. It is also an arbitrary cut off to choose 10 seconds. The decision was made to window at this length and no other options were investigated to remove the possibility of tuning the results.

6.4.7 Conclusion

To our knowledge this is study is the first to investigate and report on a range of crying and rocking features in a natural setting. The results show that caregivers do not have a static rhythmic rocking response to their infants' cries, they do however show a dynamic response to help soothe their infants. Over a crying bout, caregivers were shown to increase the amount of rocking, change their rocking frequency relative to changes in infant crying oscillations, and employ a range of phase differences. For the first time, caregivers were also shown to produce a turn taking behaviour when soothing negative vocalisations with rocking. There is evidence to show that more consistency in rocking behaviour may be best rocking method for soothing infants. Further investigations using day long home recordings are needed to further study this area.

Chapter 7 – General Discussion

7.1 Summary and integration of findings

This thesis aimed to investigate the capabilities, mechanisms and applications of infant neural and physiological oscillatory entrainment using a multi-method approach. Focussing on these areas of infant vs adult oscillatory entrainment gives researchers a valuable insight into development that would not be possible with adult participants alone. To do this, adult and infants' neural activity was recorded during sleep and social play interactions in a laboratory setting, and on separate occasions infant and adult actigraphy and audio were recorded in day-long home recordings. Using time-locked and continuous analyses, this thesis advances our current understanding of how infants entrain to different rhythmic stimuli, quasi-rhythmic stimuli, infant directed speech and song and how external influences can affect infant vocalisations.

In Chapter 2, the processes and capabilities that drive infant entrainment to simulated and naturalistic stimuli were outlined. Studies have highlighted that infants are able to show a rhythmic entrainment response to simple isochronous auditory stimuli from the third trimester of pregnancy (Niepel et al., 2020) and in premature infants of a similar gestational age (Daneshvarfard et al., 2019; Edalati et al., 2023). After birth, newborns continue to show contingent responding to rhythmic auditory components (Háden et al., 2015), have been shown to categorise complex stimuli based on rhythmic structures such as languages (Nazzi, Bertoncini & Mehler, 1998) and model acoustic regularity based on rhythmic properties (Stefanics et al., 2007). Newborns also demonstrate a mismatch negativity to missed beats (Winkler et al., 2009) which may suggest innate oscillatory entrainment or a linear dampened contingent response. Infants up to 6 months were shown to be able to respond to a variety of rhythms and can group multitone rhythms with a "Gestalt law of proximity" (Demany, McKenzie & Vurpillot, 1977; Hsing-Wu & Trehub, 1977). Comparing between 6- and 12-month-olds, shows a sensitive period where younger infants showed a greater proficiency for metrical violations across cultural boundaries (Hannon & Trehub, 2005). Also, between 6 and 12 months, continuous neural mapping of complex stimuli research begins, which has shown that infants can successfully track infant directed speech, song and cartoons (Attaheri et al., 2022; Jessen et al., 2019; Kalashnikova et al., 2018; Menn et al., 2022; Nguyen et al., under review; Ortiz Barajas, Guerva & Gervain, 2021; Phillips et al., in prep). In this vein it is clear that there is progression in development either driven by the infant entrainment progression or by the choices made by researchers about which stimuli to show infants at each age.

Chapter 2 also discusses evidence of mechanisms of entrainment between contingent responding to external stimuli or oscillatory entrainment of internal oscillators with external rhythms. It is noted by Bánki et al., (2022) and Wass, Perapoch Amadó & Ives (2022) that while evidence is building for oscillatory entrainment mechanisms being an adult neural entrainment mechanism with further evidence of benefits to memory performance and temporal predictions (see Bánki et al., 2022 for review), there is limited research that has examined whether infant neural entrainment uses the same mechanisms. Infant entrainment could represent one of four possibilities: a) infants passively contingently respond and learn oscillatory entrainment throughout development, b) infants employ “dumb” oscillators, internal oscillators that are cajoled into alignment with external oscillators through activity generated while processing of external oscillators to help them achieve oscillatory entrainment without top-down processing, c) infants use “smart” top-down oscillatory entrainment from early infancy, where internal oscillators are brought into alignment with external oscillators through conscious or subconscious processing or, d) infants display a sensitive plastic entrainment period that shows oscillatory entrainment early on, but that this diminishes at the same time that other sensitivities diminish (for example sensitivity to culturally familiar and unfamiliar metrical structures).

In Chapter 3, we investigated whether infants and adults entrained differently across a range (2-12Hz) of isochronous auditory frequencies when all other acoustic components were held constant. Sensitivity of power and phase analyses, spatial mapping of neural responses and time taken to entrain to the stimuli were also examined. A frequency tagging paradigm was used while infants slept, and adults rested or slept. After a signal to noise calculation to normalise the data and avoid the 1/f brain component, the neural response at the stimulation frequency was measured against surrounding frequency bins and then between stimulation frequencies.

The results showed that for both adult and infant participants that there was above chance entrainment to all stimulation frequencies tested, but also that some frequencies showed significantly stronger entrainment than others. For infant participants 2Hz and 4Hz, and in adults 2Hz, 4Hz, 6Hz and 12Hz stimuli showed stronger neural entrainment. This to our knowledge is the first time that this entrainment response has been demonstrated and contradicts the hypothesis of previous studies (for review see Kabdebon et al., 2022) that suggested neural entrainment would be strongest at the dominant resting frequencies for each participant group (infants 6-9Hz, adults 9-12Hz). This was hypothesised as a result of marshalling the activity at these dominant frequency bands to entrain to the stimulus frequency in line with work by Notbohm, Kurths and Sack (2016). One possible explanation for the pattern of strong frequencies seen is that even by 6 months, infants have been exposed to a variety of low amplitude modulated stimuli, for example infant directed language, which has been shown to have a 2Hz stress rate (Goswami and Leong, 2013; Leong, 2014; Leong and Goswami, 2015), while adults have learned to entrain to a wider range of frequencies. However, other studies (Hannon & Trehub, 2005) have suggested that 6-month-old infants show a sensitivity to a wider range of meters, which disappears with age, which contradicts this hypothesis. Further study is needed to understand the mechanisms behind these entrainment responses.

Further results investigating the sensitivity of power and phase analysis showed no group differences but did show individual differences. Interestingly, similar to previous studies (e.g. Assaneo et al., 2019) that have shown there is an individual variability in whether adults entrain to a stimulus, there was an individual variability across infants and adults in this study. Not all infants or adults were shown to entrain to the stimulus, and while the majority of participants that did entrain showed a response in both the power and phase domains there were some who showed a response singularly in the power or phase domain. One alternative explanation for the individual differences could be varying levels of noise between participants. Drastically increased noise would make it difficult to identify signal in the signal to noise calculation, which could give a false negative entrainment result. However, the preprocessing steps too great care to remove as much noise as possible and where participants, channels or segments (1 second windows) were too noisy these were removed evenly across participants.

Temporal analyses showed that longer stimulus durations associated with more significant entrainment responses. Consistent frequency tagging responses were seen to the 2- and 4-minute blocks, but not to each of the 30-second- and 1-minute blocks. This suggests that longer stimulus durations promote neural entrainment through prolonged exposure. It also suggests frequency tagging stimuli should be at least 2 minutes long. To our knowledge this is the first time this has been investigated in this context.

Spatial analyses showed that there were clusters of increased activity in adult data in a slightly left lateralised fronto-central region and in a temporo-occipital cluster spread across both hemispheres. In infant data a right lateralised fronto-central-parietal cluster was shown along with a left lateralised occipital-temporal cluster. Previous studies had restricted the spatial regions being investigated, and it was noted that almost all of the common clusters (see section 3.1.6) found overlapped with regions of significant and non-significant results. This may suggest that previous results had been underpowered by including these regions and that it is recommended to either redraw clusters being generated or collect whole head data.

Chapter 4 built on these results by using a similar frequency tagging paradigm but with a range of jittered audio stimuli (0%, 5%, 10% and 20%). These are stimuli that have a consciously perceived isochronous rhythm, but the individual stimulus beats are presented ahead or behind this rhythm. The positive and negative lags are set across a flat distribution and randomised, so that all stimuli are counterbalanced but not in sequential order. Audio analysis showed that as stimulus beats overlapped the isochronous rhythm there is more power at the rhythm than surrounding frequencies but that as the level of jitter increased the beats became more spaced out and this decreased in a linear manner. The hypothesis tested was that if infants and adults showed a purely contingent response to the jittered beats then the neural response at the rhythm frequency should decrease at the same rate as the power in the audio stimuli. However, a sustained or uneven decrease would suggest that other mechanisms are being used.

The results showed that both adults and infants produced an above chance frequency tagging response at the rhythm frequency, which is contradictory to previous findings in adults with a similar paradigm that did not show a response (Lehmann, Arias & Schönwiesner, 2016; Tavano, et al., 2022). Differences may have been due to the stimulus design, with stimuli in the present study being both longer (when compared to 50ms clicks) and more mildly jittered (previous studies used a 55-65% jitter rate). Between conditions, infants did not show a significant decrease in neural response as jitter increased from isochronous beats to 20% jittered beats, while adults showed an uneven decrease in both power and phase alignment as jitter level increased. These results suggest that infants were able to sustain the same level of neural entrainment to the beats despite the increase in jitter, which would be contradictory to a model that relies solely on contingent responding and points towards an oscillatory entrainment mechanism as hypothesised above. While adult results are not clearly described by either a purely contingent responding mechanism or oscillatory entrainment, suggesting that a mixture of mechanisms was being employed. Alternatively, infants could be displaying an increased sensitivity to a wider range of rhythms similar to those reported by Hannon and Trehub (2005) and Cantiani et al., (2022). This may mean that infants are temporarily better at entraining to the regular overall structure of irregular rhythms and that this diminishes over time. An alternate hypothesis is that there are fundamental differences in how infants and adults neurally processed the individual beats, and that infant neural processing took longer, capping all frequency tagging responses at a much lower rate. However, as a typical neural stimulus response is likely to be much longer than the 125ms beat interval used this could not be tested without interference from multiple stimuli. Future research should attempt to determine whether infant neural processing is the primary driver of this finding.

Further results from Chapter 4 investigated inter trial power and phase coherence segmented relative to either the stimulus beat or where the isochronous rhythm would have been. The results showed a significantly larger phase alignment to the rhythm for adults in the 5% jitter condition and to the stimulus beat for infants in the 20% stimulus condition, but not other significant differences. These results showed limited differences between stimulus beat and rhythm conditions with no clear trend which makes it difficult to make further generalisations.

Chapter 5 assessed infant speech brain tracking in an interactive caregiver-infant context which naturalistically encouraged caregivers to use infant directed speech and song similar to previous studies by (Attaheri et al., 2022; Kalashnikova et al., 2018; Menn et al., 2022; Nguyen et al., under review; Ortiz Barajas, Guerva & Gervain, 2021). The present study however focussed on the impact of different rhythms within the infant directed language conditions. Infants in this study were 5, 10 and 15 months old so that differences in age could also be assessed. A forwards multivariate temporal response function was used to continuously map auditory stimuli onto neural responses in an attempt to recreate the expected neural response. The results showed that there was significant delta band speech-brain tracking to the infant directed speech condition, which was subsequently shown to be driven by the 15-month-old infants. The current results replicated a previous finding by Philips et al., (in prep) in a similar free-flowing tabletop play context, which showed significantly above chance delta tracking in older infants over 11 months old. Theta and alpha band neural tracking were shown in the infant directed song condition, which was subsequently shown to be driven by the 6-month-old infants (theta) and 6- and 15-month-old infants (alpha). In some ways this is similar to work by Attaheri et al., (2022), who demonstrated delta and theta tracking to sung nursery rhymes at 4, 7 and 11 months but no alpha band tracking. However, the stimuli in Attaheri, et al.'s study were much more tightly controlled recordings where the actor listened to a 2Hz click track and timed their vocalisations accordingly, which may account for the differences in frequency band response seen. Similar to Attaheri et al., the youngest age group showed the most neural tracking.

To examine the amount and frequency of rhythms present in each of the play conditions, a cycle-by-cycle analysis (Cole & Voytek, 2018) was conducted to search for oscillatory activity in the temporal domain. Correlation between the amount and dominant frequency of each trial against the amount of speech brain tracking was conducted to investigate whether features of the rhythmicity had an impact on infant neural tracking. As expected, there was more time identified as oscillatory in the song vs speech condition, and there was a lower dominant frequency in the song compared to the speech condition. However, there was also a much broader spread of rhythmic frequencies identified in the song condition compared to the speech condition. The speech and song conditions were pooled to increase statistical power, but there was no significant correlation between either the amount of oscillatory audio or dominant frequency of the trial compared to the amount of speech brain tracking. This is interesting as previous research has suggested that more rhythmic intonation of speech elicits a stronger rhythmic neural response in adults and introducing a more rhythmic structure helps language processing in older children who have language deficits (for review see Fiveash et al., 2021). However, it's important to note that mTRF analysis investigates neural mapping not only to rhythmic features but takes a holistic approach to neural encoding relative to the input stimuli, which may suggest that other environmental features had evoked a stronger response.

Finally, in Chapter 6 infant crying rhythms, caregiver rocking rhythms and interactions between these oscillators was investigated in the context of caregiver bids to soothe their infants. Day-long home recordings of infants and their caregivers were made with custom wearable devices that measured continuous audio, actigraphy, GPS and proximity. Crying periods were detected using a machine learning algorithm (Micheletti et al., 2022), and the audio envelope and actigraphic activity for these segments were assessed for rhythmicity using a cycle-by-cycle analysis (Cole & Voytek, 2018).

In brief, for analyses 1 and 2, frequency and phase characteristics were assessed vs bout length to determine whether these characteristics significantly shortened or lengthened bout length. Frequency characteristics included the amount and dominant frequency of the infant's audio envelope (representing the crying audio) and actigraphic activity across each of the three directional axes, which was taken from the cycle-by-cycle output. Phase characteristics included the circular mean phase, phase locking value and circular variance of phase of the infant's audio envelope and actigraphy data. For both frequency and phase characteristics, the data were split into deciles and ranked against a bootstrapped surrogate dataset. Next to check for an impact on bout length, the data were split into segments and linear mixed effects models were run for each characteristic and segment with individual participants as a grouping variable. Analysis 3 investigated whether there was tracking of changes between the crying and rocking frequencies using a cross correlation analysis of the identified oscillation frequencies from the cycle-by-cycle analysis. Finally in analysis 4 a turn taking analysis similar to Yoo, Bowman and Oller (2018) was conducted on audio and actigraphy oscillations to investigate whether caregivers used competing rocking rhythms to attempt to overlap and override infant crying rhythms.

The results from analysis 1 showed that over the crying period auditory oscillation frequency decreased, amount of rocking increased, and that rocking frequency peaked significantly at different points during the crying period based on the axis of motion. There were no significant effects of frequency characteristics on bout length. Results from analysis 2 showed that over the crying period there was no significance compared to chance for mean phase angle, phase locking value was significantly lower than chance and phase variance was significantly higher. No significant effects on bout length were found for mean phase angle or variance but there was a significant negative effect of increased PLV. Taken together this suggests that as PLV is lower than a bootstrapped dataset, caregivers are intentionally attempting to change the phase angle difference between their rocking and infant crying, which is likely a result changes in rocking technique. However, when caregivers keep a consistent phase angle between rocking and crying rhythm so that PLV increases, this in fact soothes infants. Results of analysis 3 showed that there was a significant cross correlation of rocking frequency at 3, 6 and 7 seconds after changes in auditory oscillation frequency. This is supported by results from analysis 4 that showed that cry oscillations were around 2 seconds and there was a positive turn taking lag of around 1 second, which combined matches the 3 second cross correlation lag seen in analysis 3. This suggests that caregivers are employing a turn taking structure of oscillatory behaviour and that when they take their turn, they match increases or decreases in auditory oscillation frequency.

7.2 Implications for theories of early infancy entrainment

7.2.1 Infant entrainment capabilities

Infants showed above chance entrainment responses to all computer-generated auditory rhythms in Chapters 3 and 4, while infants showed limited entrainment through speech-brain tracking across age ranges. Evidence from within controlled setting shows that infants are capable of entraining to a wide variety of low amplitude modulated rhythms and across a range of irregularities. While further research is needed, the strong entrainment from infants at 2Hz and 4Hz may reflect experience with rhythmic structures in their environment such as infant directed speech and song. However, while rhythms were detected across infant directed speech and song in Chapter 5, infants at the lowest age point showed no above chance speech brain tracking and song brain tracking in the theta and alpha frequency bands which were both above the 2/4Hz strong entrainment rhythms shown in Chapter 3. This may represent the impact of increasing confounding variables in a more naturalistic study, such as variability in rhythms found in infant directed song, visual cues, changes in other acoustic components (e.g. pitch).

Neural speech-brain tracking in 15-month-olds showed an interesting difference between frequency bands with significantly more delta band speech-brain tracking while also showing a significant decrease in theta and alpha frequency bands. This may represent a marshalling of activity from theta and alpha bands towards delta band to facilitate speech-brain tracking. Cautious interpretation of the pattern of song-brain tracking may suggest an early sensitive period as seen in the theta and alpha frequency bands at 5 months, which decreases with age by 10 months. This could also be met with an increase in learned experience at faster frequencies as seen in the alpha band at 15 months. An alternative explanation that was not explored in the current study due to time and resource limitations were differences in the type of infant directed song used across age points. If caregivers systematically preferred lullabies or playsongs at particular ages, this could also bias neural tracking responses as differences have been seen between the song types (Nguyen et al., under review). Further research would be needed to investigate this more thoroughly across timepoints.

Infants did not show physiological entrainment of respiration cycles while crying as caregivers employed competing rocking frequencies to soothe the infants. Direct 1:1 and 2:1 entrainment between respiration and rocking had been shown previously in calm settings (Sammon & Darnall, 1994), which suggests that either infants allowed themselves to entrain to the rocking rhythms in this previous study or perhaps they showed a resistance to physiological entrainment while crying in the present study. It could be argued that as the setting has a much higher arousal, that the crying activity drowned out rhythmic attempts to soothe. However, rhythmic rocking has been used as an effective soothing mechanism across the world, so it would be expected that the natural rocking recorded would be representative of typically effective rocking behaviours.

Taken together the results from this thesis expand on previous knowledge of infant entrainment capabilities by showing differences in entrainment to steady and jittered rhythms, by investigating across infant directed language types to focus on rhythmic elements and by showing potential evidence of entrainment resistance.

7.2.2 Differences between adult and infant entrainment

Limited studies have investigated differences between infant and adult entrainment. In Chapters 3 and 4, a direct comparison of infant and adult entrainment was possible. For the first-time entrainment preferences based on rhythmic frequency were shown between the two age groups, as well as spatial differences in stimulus processing. These differences may reflect differing processing pathways between infant and adult entrainment. However, they may also reflect differences in attentional states between the participant groups, as infants were generally asleep and adults were generally awake. This may have allowed adult participants to consciously attend to a wider range of frequencies. It is unclear though why there would be a systematic decrease in adult entrainment specifically to the 8 and 10Hz stimulus frequencies if the driving factor behind strong entrainment to particular frequency rates was attention based. Similar individual rates of entrainment were demonstrated across power and phase analyses (see section 3.3.2.2). Interestingly, there was no strong association between an adult that showed neural entrainment and their infant or vice versa. This may point to evidence that entrainment responses are not driven by hereditary factors or infant-caregiver shared environment since birth. Further research would be needed to test these hypotheses. Adults and infants also demonstrated a similar time taken to show consistent significant stimulus entrainment (see section 3.3.3). It is unclear why this is given the other differences in infant and adult entrainment, but it may represent the time taken to shift internal oscillators for oscillatory entrainment.

Infant-adult differences were also seen in Chapter 4 with infants showing no differences between conditions as jitter increased, while adults showed significant decreases in neural responses as jitter increased. This may represent the first evidence of differences in entrainment mechanisms across adult-infant populations. Differences may be due to adult and infant populations processing the auditory stimuli using different neural pathways or may be due increased experience in adult participants who may have recognised that there was less information at the isochronous rhythm.

One factor to consider which may explain the potential more sensitive response seen from infants but not adults may be the idea that particular sensory scaffolding may be in place in early stages of development (e.g. Vogelsang et al., 2024). Here preferential processing of low-frequency auditory stimuli that had been trained *in utero* could be used as a prerequisite for language acquisition. Once the required skills have been learned, this extended ability may no longer be needed and could diminish over time.

However, as stimulus processing time and high-resolution spatial imaging could not be tested the results are considered cautiously and future study should test for stimulus processing times.

7.2.3 Infant mechanisms of entrainment

Infants showed capacity to entrain to all sinusoidal auditory stimuli in Chapter 3 and 4 when compared to chance However, evidence points away from this being driven by a purely contingent responding mechanism of entrainment, as it would be expected that infants would show similar entrainment across non-jittered frequency stimuli, while showing a decreasing entrainment response in jittered conditions as jitter increases. However, the opposite was seen with preferences for some non-jittered stimulation frequencies over others and a plateaued response across jittered conditions. Under a contingent responding mechanism as neural responses were segmented into shorter time periods (section 3.3.3) it would also have been expected that entrainment would have remained constant, as there would have been the same number of stimulus beats in each period. However, infant frequency tagging responses showed an initial significant response, which diminished and returned later in the trial. This may represent an initial contingent response, habituation to the stimulus and growing oscillatory entrainment. Further study would be needed to investigate this mechanism more thoroughly. Further evidence against a purely contingent responding mechanism of entrainment is seen in Chapters 5 and 6. In Chapter 5, the dominant auditory frequencies identified in the speech and song conditions did not match the frequency bands where significant neural tracking was demonstrated. If infants demonstrated a purely contingent response to the auditory rhythms, it would be expected that these would have matched. Similarly, if infants responded purely contingently to rocking rhythms, infant crying rhythms would have been expected to rise to match the dominant rocking rhythms and infants would be expected to show greater entrainment to these rhythms and faster disruption of infant crying, however this was not seen.

There is some evidence from this thesis that suggests that infants show oscillatory entrainment from Chapter 2 with the plateaued response across jittered conditions. However, questions remain regarding how quickly infants can process these stimuli and whether an increased time to process the stimulus audio further smears the entrainment response, lowering the responses for all conditions. There is also no evidence regarding whether infants showed top-down “smart” oscillatory entrainment, entrainment driven by “dumb oscillators” with exogenous stimuli driving endogenous oscillatory entrainment or whether this result shows a developmental sensitivity to irregular rhythm that decreases with age. Further research is needed to investigate this more broadly. Studying older infants would allow for testing of a sensitive period. While more precise analysis would be needed to investigate whether infants show a gradual entrainment to the stimuli, which may potentially suggest dumb oscillatory entrainment, or whether there is a sharp change in entrainment which would suggest top-down oscillatory entrainment.

The evidence from this thesis points away from infants as passive entrainers to their external environment and instead points to a more involved entrainment profile. Differences between entrainment responses as seen above suggests further infant development of entrainment responses with time. Future study should investigate further entrainment responses throughout development and investigate dumb vs smart infant oscillatory entrainment.

7.3 Limitations and future directions

The experimental chapters in this thesis show the trade-off between controlled lab settings and more naturalistic experiments. Chapters 3 and 4 replicate previous results showing strong neural entrainment to auditory stimuli but are not ecologically valid and a range of more naturalistic rhythms shown in Chapter 5 did not elicit strong frequency tagging within wider frequency bands. Whereas Chapter 5 and 6, which used more naturalistic designs, suffered from a range of potentially confounding variables, such as types of speech and song, level of engagement with objects and affective expressions used by adult caregivers in Chapter 5. Infant rocking position, time of day, underlying reason for the infant cry and unseen environmental factors may have confounded results in Chapter 6 despite efforts to filter out noisy time points. While lab-based studies are important to establish foundations of infant capabilities it is important that future studies attempt to be as natural as possible to capture the quasi-chaotic nature of infant development.

Further research investigating neural capabilities of infant and adult entrainment preferences across frequencies should focus on lower amplitude modulated rhythms between 0.1 and 4Hz. This would show whether there are fine-grained frequency differences or specific entrainment preferences at the 2Hz and 4Hz and give a clearer picture of entrainment profiles. Capabilities to entrain to further irregular temporal rhythms could also be explored with higher levels of jitter, but it would also be important to measure temporal differences in neural responses.

Further research investigating whether speech brain tracking is impacted by the amount and frequency of rhythm should explore much larger participant populations so that further categorisations could be made based on naturally occurring speech and song types, levels of affect displayed, and number of visual cues to infants. Alternatively, future studies could sacrifice the spontaneous nature of the content and give clear instructions to narrow the type of infant directed vocalisations.

Similarly, investigation into rhythmic infant soothing should utilise larger population pools so that distinctions can be made between orientation of rocking, amount and content of caregiver vocalisations and broad categorisation of the type of situation (e.g. nap time, feeding time, play etc).

7.4 Overall conclusions

The contribution of this thesis is in taking a multi-method approach combining neural and physiological techniques to examine infant entrainment to simulated and natural rhythmic stimuli. In all 4 empirical chapters, this work has provided evidence that expands on current knowledge of infant capabilities, entrainment mechanisms and applications of infant neural and physiological entrainment. In particular, the findings of this thesis point to developmental differences between infants and adults. Infants were also shown not to be passive entrainers to their environment as evidence points against infant using a purely contingent responding mechanism of entrainment and towards forms of infant oscillatory entrainment.

However, much future research is needed to confidently assess infant entrainment mechanisms and fully illuminate infant rhythmic entrainment capabilities. Development of well-designed future experiments will be needed to investigate differences in the types of oscillatory entrainment and the developmental trajectory of infant rhythmic entrainment. Finally, future work will be needed to marry tightly controlled lab-based responses and naturalistic experimental designs.

References

Abboub, N., Nazzi, T., & Gervain, J. (2016). Prosodic grouping at birth. *Brain and Language, 162,* 46–59. https://doi.org/10.1016/j.bandl.2016.08.002

Adibpour, P., Lebenberg, J., Kabdebon, C., Dehaene-Lambertz, G., & Dubois, J. (2020). Anatomo-functional correlates of auditory development in infancy. *Developmental Cognitive Neuroscience*, *42*, 100752. https://doi.org/10.1016/j.dcn.2019.100752

Aharoni, M., Breska, A., Müller, M. M., & Schröger, E. (2023). Mechanisms of sustained perceptual entrainment after stimulus offset. *European Journal of Neuroscience*. https://doi.org/10.1111/ejn.16032

Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., Merzenich, M. (2001). Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *PNAS*, *98*(23), 13367–13372.

Ahlo, K., Sainio, K., Sajaniemi, N., Reinikainen, K. Näätänen, R. (1990). Event-related brain potential of human newborns to pitch change of an acoustic stimulus. *Electroencephalography and Clinical Neurophysiology1*, *77*, 151–155.

Ahtola, E., Stjerna, S., Tokariev, A., & Vanhatalo, S. (2020). Use of complex visual stimuli allows controlled recruitment of cortical networks in infants. *Clinical Neurophysiology*, *131*(8), 2032–2040. https://doi.org/10.1016/j.clinph.2020.03.034

Aiken, S. J., & Picton, T. W. (2008). Human Cortical Responses to the Speech Envelope. *Ear & Hearing*, *29*(2), 139–157.

Alaerts, J., Heleen, L., & Hofmann, M. Wouters, J. (2009). Cortical auditory steady-state responses to low modulation rates. *Internation Journal of Audiology*, *48*, 582–593. https://doi.org/10.1080/14992020902894558

Allen, S. L., Howlett, M. D., Coulombe, J. A., & Corkum, P. V. (2016). ABCs of SLEEPING: A review of the evidence behind pediatric sleep practice recommendations. *Sleep Medicine Reviews*, *29*, 1–14. https://doi.org/10.1016/j.smrv.2015.08.006

Alschuler, D. M., Tenke, C. E., Bruder, G. E., & Kayser, J. (2014). Identifying electrode bridging from electrical distance distributions: A survey of publicly-available EEG data using a new method. *Clinical Neurophysiology*, *125*(3), 484–490. https://doi.org/10.1016/j.clinph.2013.08.024

Ambrose, A. (1969). Stimulation in early infancy. In *Stimulation in early infancy.* Academic Press.

Ames, R. W., Khazaie, S., Gavel, S., & Farrell, T. (1987). Mothers' reports of the effectiveness of soothing techniques. *Cry Newsletter*, *3*, 5–8.

Aoyagi, M., Kiren, T., Furuse, H., Fuse, T., Suzuki, Y., Yokota, M., & Koike, Y. (1994). Effects of Aging on Amplitude-modulation Following Response Effects of Aging on Amplitude-modulation Following Response. *Acta Oto-Laryngologica*, *511*, 15–22. https://doi.org/10.3109/00016489409128295

Arnal, L. H., & Giraud, A. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, *16*(7), 390–398. https://doi.org/10.1016/j.tics.2012.05.003

Arnal, L. H., Doelling, K. B., & Poeppel, D. (2015). *Delta – Beta Coupled Oscillations Underlie Temporal Prediction Accuracy*. *September*, 3077–3085. https://doi.org/10.1093/cercor/bhu103

Assaneo, M. F., Ripollés, P., Orpella, J., Lin, W. M., de Diego-Balaguer, R., & Poeppel, D. (2019). Spontaneous synchronization to speech reveals neural mechanisms facilitating language learning. *Nature Neuroscience*, *22*(4), 627–632. https://doi.org/10.1038/s41593-019-0353-z

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403–450. https://doi.org/10.1146/annurev.neuro.28.061604.135709

Attaheri, A., Choisdealbha, Á. N., Di Liberto, G. M., Rocha, S., Brusini, P., Mead, N., Olawole-Scott, H., Boutris, P., Gibbon, S., Williams, I., Grey, C., Flanagan, S., & Goswami, U. (2022). Delta- and theta-band cortical tracking and phase-amplitude coupling to sung speech by infants. *NeuroImage*, *247*(October 2021), 118698. https://doi.org/10.1016/j.neuroimage.2021.118698

Ayala, Y. A., Lehmann, A., & Merchant, H. (2017). Monkeys share the neurophysiological basis for encoding sound periodicities captured by the frequency-following response with humans. *Scientific Reports*, *7*(1), 1–11. https://doi.org/10.1038/s41598-017-16774-8

Babiloni, F., & Astolfi, L. (2014). Neuroscience and Biobehavioral Reviews Social neuroscience and hyperscanning techniques : Past , present and future.

Neuroscience and Biobehavioral Reviews, *44*, 76–93. https://doi.org/10.1016/j.neubiorev.2012.07.006

Backer, K. C., Kessler, A. S., Lawyer, L. A., Corina, D. P., & Miller, L. M. (2019). A novel EEG paradigm to simultaneously and rapidly assess the functioning of auditory and visual pathways. *Journal of Neurophysiology*, *122*(4), 1312–1329. https://doi.org/10.1152/jn.00868.2018

Bánki, A., Brzozowska, A., Hoehl, S., & Köster, M. (2022). Neural Entrainment vs. Stimulus-Tracking: A Conceptual Challenge for Rhythmic Perceptual Stimulation in Developmental Neuroscience. *Frontiers in Psychology*, *13*(May), 1–5. https://doi.org/10.3389/fpsyg.2022.878984

Barnard, K. E., & Bee, H. L. (1983). The Impact of Temporally Patterned Stimulation on the Development of Preterm Infants. *Child Development*, *54*(5), 1156. https://doi.org/10.2307/1129671

Barnet, Ann; Ohlrich, Elizabeth; Weiss, Ira; Shanks, B. (1975). AUDITORY EVOKED POTENTIALS DURING SLEEP IN NORMAL CHILDREN FROM TEN DAYS TO THREE YEARS OF AGE. *Electroencephalography and Clinical Neurophysiology*, *39*, 29–41.

Baruch, C., & Drake, C. (1997). Tempo discrimination in infants. *Infant Behavior and Development*, *20*(4), 573–577. https://doi.org/10.1016/S0163-6383(97)90049-7

Bednar, A., & Lalor, E. C. (2020). Where is the cocktail party? Decoding locations of attended and unattended moving sound sources using EEG. *NeuroImage*, *205*(August 2019), 116283. https://doi.org/10.1016/j.neuroimage.2019.116283

Begum Ali, J., Charman, T., Johnson, M. H., Jones, E. J. H., Agyapong, M., Bazelmans, T., Dafner, L., Ersoy, M., Gliga, T., Goodwin, A., Haartsen, R., Hendry, A., Holman, R., Kalwarowsky, S., Kolesnik, A., Lloyd-Fox, S., Mason, L., Pasco, G., Pickles, A., … Taylor, C. (2020). Early Motor Differences in Infants at Elevated Likelihood of Autism Spectrum Disorder and/or Attention Deficit Hyperactivity Disorder. *Journal of Autism and Developmental Disorders*, *50*(12), 4367–4384. https://doi.org/10.1007/s10803-020-04489-1

Bell, S. M., & Ainsworth, M. D. S. (1972). Infant Crying and Maternal Responsiveness Author (s): Silvia M . Bell and Mary D . Salter Ainsworth Published by : Wiley on behalf of the Society for Research in Child Development Stable URL : http://www.jstor.org/stable/1127506 REFERENCES Linked refer. *Child Development*, *43*(4), 1171–1190.

Benjamini, Y., & Yekutieli, D. (2001). The Control of the False Discovery Rate in Multiple Testing Under Dependency. *The Annals of Statistics*, *29*(4), 1165–1188.

Bergeson, T. R., & Trehub, S. E. (2002). Absolute pitch and tempo in mothers' songs to infants. *Psychological Science*, *13*(1), 72–75. https://doi.org/10.1111/1467-9280.00413

Bernstein-Ratner, N. (1985). Dissociations between Vowel Durations and Formant Frequency Characteristics. *Journal of Speech, Language, and Hearing Research*, *28*(2), 255–264. https://doi.org/10.1044/jshr.2802.255

Bertoncini, J., Nazzi, T., Cabrera, L., & Lorenzi, C. (2011). Six-month-old infants discriminate voicing on the basis of temporal envelope cues (L). The Journal of the Acoustical Society of America, 129(5), 2761–2764. https://doi.org/10.1121/1.3571424

Besle, J., Schevon, C. A., Mehta, A. D., Lakatos, P., Goodman, R. R., Mckhann, G. M., Emerson, R. G., & Schroeder, C. E. (2011). Tuning of the Human Neocortex to the Temporal Dynamics of Attended Events. *The Journal of Neuroscience*, *31*(9), 3176–3185. https://doi.org/10.1523/JNEUROSCI.4518-10.2011

Birca, A., Carmant, L., Lortie, A., & Lassonde, M. (2006). Interaction between the flash evoked SSVEPs and the spontaneous EEG activity in children and adults. *Clinical Neurophysiology*, *117(2)*, 279-288. https://doi.org/10.1016/j.clinph.2005.10.001

Blume-Peytavi, U., Hauser, M., Stamatas, G. N., Pathirana, D., & Bartels, N. G. (2012). Skin care practices for newborns and infants: Review of the clinical evidence for best practices. *Pediatric Dermatology*, *29*(1), 1–14. https://doi.org/10.1111/j.1525-1470.2011.01594.x

Bobin-Bègue, A. (2019). Rhythms in Early Development. In G. Apter, E. Devouche, & M. Gratier (Eds.), *Early Interaction and Developmental Psychopathology, Volume 1: Infancy: Vol. I* (1st ed., pp. 89–105). Springer-Nature. https://doi.org/10.1007/978-3-030-04769-6_5

Bouwer, F. L., Werner, C. M., Knetemann, M., & Honing, H. (2016). Disentangling beat perception from sequential learning and examining the influence of attention and musical abilities on ERP responses to rhythm. *Neuropsychologia*, *85*, 80–90. https://doi.org/10.1016/j.neuropsychologia.2016.02.018

Boyer, T. W., Harding, S. M., & Bertenthal, B. I. (2020). The temporal dynamics of infants' joint attention: Effects of others' gaze cues and manual actions. *Cognition*, *197*(December). https://doi.org/10.1016/j.cognition.2019.104151

Broesch, T. L., & Bryant, G. A. (2015). Prosody in Infant-Directed Speech Is Similar Across Western and Traditional Cultures. *Journal of Cognition and Development*, *16*(1), 31–43. https://doi.org/10.1080/15248372.2013.833923

Bronkhorst, A. W. (2015). The cocktail-party problem revisited: early processing and selection of multi-talker speech. *Attention, Perception, and Psychophysics*, *77*(5), 1465–1487. https://doi.org/10.3758/s13414-015-0882-9

Burgess, A. (2013). On the interpretation of synchronization in EEG hyperscanning studies : a cautionary note. *Frontiers in Human Neuroscience*, *7*(December), 1–17. https://doi.org/10.3389/fnhum.2013.00881

Burgess, A. P. (2012). Towards a Unified Understanding of Event-Related Changes in the EEG: The Firefly Model of Synchronization through Cross-Frequency Phase Modulation. *PLoS ONE*, *7*(9). https://doi.org/10.1371/journal.pone.0045630

Burnham, D., & Dodd, B. (2004). Auditory-visual speech integration by prelinguistic infants: Perception of an emergent consonant in the McGurk effect. *Developmental Psychobiology*, *45*(4), 204–220. https://doi.org/10.1002/dev.20032

Buzsáki, G. (2006). *Rhythms of the Brain*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195301069.001.0001

Buzsáki, G., & Draguhn, A. (2004). Neuronal Oscillations in Cortical Networks. *Science*, *304*(June), 1926–1929. http://science.sciencemag.org/

Byrne, J. M., & Horowitz, F. D. (1981). Rocking as a soothing intervention: The influence of direction and type of movement. *Infant Behavior and Development*, *4*(1), 207–218. https://doi.org/10.1016/S0163-6383(81)80024-0

Cantiani, C., Dondena, C., Molteni, M., Riva, V., & Piazza, C. (2022). Synchronizing with the rhythm: Infant neural entrainment to complex musical and speech stimuli. *Frontiers in Psychology* (Vol. 13). https://www.frontiersin.org/articles/10.3389/fpsyg.2022.944670

Capilla, A., Pazo-Alvarez, P., Darriba, A., Campo, P., & Gross, J. (2011). Steady-state visual evoked potentials can be explained by temporal superposition of transient event-related responses. *PLoS ONE*, *6*(1). https://doi.org/10.1371/journal.pone.0014543

Carozza, S., & Leong, V. (2021). The Role of Affectionate Caregiver Touch in Early Neurodevelopment and Parent–Infant Interactional Synchrony. *Frontiers in Neuroscience*, *14*(January), 1–11. https://doi.org/10.3389/fnins.2020.613378

Carral, V., Huotilainen, M., Ruusuvirta, T., Fellman, V., Näätänen, R., & Escera, C. (2005). A kind of auditory “primitive intelligence” already present at birth. *European Journal of Neuroscience*, *21*(11), 3201–3204. https://doi.org/10.1111/j.1460-9568.2005.04144.x

Cellier, D., Riddle, J., Petersen, I., & Hwang, K. (2021). The development of theta and alpha neural oscillations from ages 3 to 24 years. *Developmental Cognitive Neuroscience*, *50*, 100969. https://doi.org/10.1016/j.dcn.2021.100969

Chen-Haftek, L. (2021). Music and language development in early childhood: Integrating past research in the two domains. In *Music in the Lives of Young Children*.

Choi, D., Batterink, L. J., Black, A. K., Paller, K. A., & Werker, J. F. (2020). Preverbal Infants Discover Statistical Word Patterns at Similar Rates as Adults : Evidence From Neural Entrainment. *Association for Psychological Science*, 1–13. https://doi.org/10.1177/0956797620933237

Choisdealbha, Á. N., Attaheri, A., Rocha, S., Mead, N., Brusini, P., Gibbon, S., Boutris, P., Grey, C., Flanagan, S., & Goswami, U. (2022). *Cortical Oscillations in Pre-verbal Infants Track Rhythmic Speech and Non-speech Stimuli*.

Christodoulou, J., Leland, D. S., & Moore, D. S. (2018). Overt and Covert Attention in Infants Revealed Using Steady-State Visually Evoked Potetials. *Developmental Psychology, 54(5)*, 803-815. https://doi.org/10.1037/dev0000486

Ciccone, A. (2013). Rhythmicity in infants' experiences and their development. *Journal of Physiology Paris*, *107*(4), 286–290. https://doi.org/10.1016/j.jphysparis.2013.03.003

Cirelli L. K, Jurewicz Z. B, Trehub S. E. Effects of Maternal Singing Style on Mother-Infant Arousal and Behavior. J Cogn Neurosci. (2020) ;32(7):1213-1220. doi: 10.1162/jocn_a_01402. Epub 2019 Mar 26. PMID: 30912725.

Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014). Interpersonal synchrony increases prosocial behavior in infants. *Developmental Science*, *17*(6), 1003–1011. https://doi.org/10.1111/desc.12193

Cirelli, L. K., Spinelli, C., Nozaradan, S., & Trainor, L. J. (2016a). Measuring Neural Entrainment to Beat and Meter in Infants : Effects of Music Background. *Frontiers in Neuroscience*, *10*(May), 1–11. https://doi.org/10.3389/fnins.2016.00229

Cirelli, L. K., Trehub, S. E., & Trainor, L. J. (2018). Rhythm and melody as social signals for infants. *Annals of the New York Academy of Sciences*, *1423*(1), 66–72. https://doi.org/10.1111/nyas.13580

Cirelli, L. K., Wan, S. J., & Trainor, L. J. (2014). Fourteen-month-old infants use interpersonal synchrony as a cue to direct helpfulness. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1658). https://doi.org/10.1098/rstb.2013.0400

Cirelli, L. K., Wan, S. J., & Trainor, L. J. (2016b). Social Effects of Movement Synchrony: Increased Infant Helpfulness only Transfers to Affiliates of Synchronously Moving Partners. *Infancy*, *21*(6), 807–821. https://doi.org/10.1111/infa.12140

Cirelli, L., Wan, S., Spinelli, C., & Trainor, L. (2017). Effects of Interpersonal Movement Synchrony on Infant Helping behaviours: Is Music Necessary? *Music, Synchrony, and Prosocial Behaviour in Infants*, *34*(3), 319–326.

Clark, D. L., Cordero, L., Goss, K. C., & Manos, D. (1989). Effects of Rocking on Neuromuscular Development in the Premature. *Biol Neonate*, *56*, 306–314.

Clerkin, E. M., Hart, E., Rehg, J. M., Yu, C., Smith, L. B., & Smith, L. B. (2017). *Real-world visual statistics and infants ' first-learned object names*.

Cohen, L. T., Rickards, F. W., & Clark, G. M. (2014). A comparison of steady-state evoked potentials to modulated tones in awake and sleeping humans. *Journal of the Acoustical Society of America*, *90*(5), 2467–2479.

Cohen, M. A., & Taylor, J. A. (2002). Short-term cardiovascular oscillations in man: Measuring and modelling the physiologies. *Journal of Physiology*, *542*(3), 669–683. https://doi.org/10.1113/jphysiol.2002.017483

Cohn, J. F., & Tronick, E. Z. (1988). Mother-Infant Face-to-Face Interaction: Influence is Bidirectional and Unrelated to Periodic Cycles in Either Partner's Behavior. *Developmental Psychology*, *24*(3), 386–392. https://doi.org/10.1037/0012-1649.24.3.386

Cole, J. (2015). Prosody in context: a review. *Language, Cognition and Neuroscience, 30(1-2), 1-31.* https://doi.org/10.1080/23273798.2014.963130

Cole, S., & Voytek, B. (2019). Cycle-by-cycle analysis of neural oscillations. *Journal of Neurophysiology*, *122*(2), 849–861. https://doi.org/10.1152/JN.00273.2019

Colombo, J., & Cheatham, C. L. (2006). The emergence and basis of endogenous attention in infancy and early childhood. In *Advances in Child Development and Behavior* (Vol. 34). https://doi.org/10.1016/S0065-2407(06)80010-8

Cooper, R. P., & Aslin, R. N. (1990). Preference for Infant-directed Speech in the First Month after Birth. *Child Development*, *61*(5), 1584–1595. https://doi.org/10.1111/j.1467-8624.1990.tb02885.x

Cooper, R. P., & Aslin, R. N. (1994). Developmental Differences in Infant Attention to the Spectral Properties of Infant-directed Speech. *Child Development*, *65*(6), 1663–1677. https://doi.org/10.1111/j.1467-8624.1994.tb00841.x

Cox, C., Bergmann, C., Fowler, E., Keren-Portnoy, T., Roepstorff, A., Bryant, G., & Fusaroli, R. (2023). A systematic review and Bayesian meta-analysis of the acoustic features of infant-directed speech. *Nature Human Behavior*, *7*, 114–133.

Cravo, A. M., Rohenkohl, G., Wyart, V., & Nobre, A. C. (2013). Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. *Journal of Neuroscience*, *33*(9), 4002–4010. https://doi.org/10.1523/JNEUROSCI.4675-12.2013

Cristia, A. (2013). Input to Language: The Phonetics and Perception of Infant-Directed Speech. *Linguistics and Language Compass*, *7*(3), 157–170. https://doi.org/10.1111/lnc3.12015

Crosse, M. J., Butler, J. S., & Lalor, E. C. (2015). Congruent visual speech enhances cortical entrainment to continuous auditory speech in noise-free conditions. *Journal of Neuroscience*, *35*(42), 14195–14204. https://doi.org/10.1523/JNEUROSCI.1829-15.2015

Crosse, M. J., Di Liberto, G. M., Bednar, A., & Lalor, E. C. (2016). The multivariate temporal response function (mTRF) toolbox: A MATLAB toolbox for relating neural signals to continuous stimuli. *Frontiers in Human Neuroscience*, *10*(NOV2016), 1–14. https://doi.org/10.3389/fnhum.2016.00604

Csibra, G., Davis, G., Spratling, M. W., & Johnson, M. H. (2000). *Gamma Oscillations and Object Processing in the Infant Brain*. *290*(November), 5–8.

Cuevas, K., Cannon, E. N., Yoo, K., & Fox, N. A. (2014). The infant EEG mu rhythm : Methodological considerations and best practices. *Developmental Review*, *34*(1), 26–43. https://doi.org/10.1016/j.dr.2013.12.001

Cullen, J., Fargo, N., Chase, R., & Baker, P. (1968). The Development of Ausitory Feedback Monitoring: I. Delayed Auditory Feedback Studies on Infant Cry. *Journal of Speech and Hearing Research*, *11*(1), 85–93.

Daneshvarfard, F., Moghaddam, H. A., & Dehaene-lambertz, G. (2019). Neurodevelopment and asymmetry of auditory-related responses to repetitive syllabic stimuli in preterm neonates based on frequency-domain analysis. *Scientific Reports*, *9*(October 2018), 10654. https://doi.org/10.1038/s41598-019-47064-0

Darnley, F. (1981). Periodicity in the Family: What Is It and How Does It Work? *Family Relations*, *30*(1), 31. https://doi.org/10.2307/584232

Daume, J., Wang, P., Maye, A., Zhang, D., & Engel, A. K. (2021). NeuroImage Non-rhythmic temporal prediction involves phase resets of low-frequency delta oscillations. *NeuroImage*, *224*(September 2020), 117376. https://doi.org/10.1016/j.neuroimage.2020.117376

David, O., Harrison, L., & Friston, K. J. (2005). Modelling event-related responses in the brain. *NeuroImage*, *25*(3), 756–770. https://doi.org/10.1016/j.neuroimage.2004.12.030

David, O., Kilner, J. M., & Friston, K. J. (2006). Mechanisms of evoked and induced responses in MEG/EEG. *NeuroImage*, *31*(4), 1580–1591. https://doi.org/10.1016/j.neuroimage.2006.02.034

de Barbaro, K., & Fausey, C. M. (2022). Ten Lessons About Infants' Everyday Experiences. *Current Directions in Psychological Science*, *31*(1), 28–33. https://doi.org/10.1177/09637214211059536

de Haan, M. (Ed.). (2013). *Infant EEG and Event-Related Potentials*. Taylor & Francis.

De Lucia, L. D. (1969). To Rock or not to Rock, that is the Question. *Bulletin of Brown University*, *3*(1), 1–2.

Debener, S., Minow, F., Emkes, R., Gandras, K., & Vos, M. D. E. (2012). *How about taking a low-cost, small, and wireless EEG for a walk?* 1–5. https://doi.org/10.1111/j.1469-8986.2012.01471.x

Decasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, *208*(4448), 1174–1176. https://doi.org/10.1126/science.7375928

Delorme, A. (2001a). *pop_runICA EEGLAB function*. The Mathworks, Inc.

Delorme, A. (2001b). *pop_subcomp EEGLAB function*. The Mathworks, Inc.

Delorme, A. (2006). *eeg_interp EEGLAB function*. The Mathworks, Inc.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Demany, L. (1982). Auditory stream segregation in infancy. *Infant Behavior and Development*, *5*(2–4), 261–276. https://doi.org/10.1016/S0163-6383(82)80036-2

Demany, L., McKenzie, B., & Vurpillot, E. (1977). Rhythm perception in early infancy. *Nature*, *266*(21), 718–719.

Demos, A. P., Chaffin, R., & Marsh, K. L. (2007). *Spontaneous Vs Intentional Entrainment To a Musical Beat*. *January*, 1–4.

Dikker, S., Wan, L., Davidesco, I., Bavel, J. J. Van, Ding, M., Poeppel, D., Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., Mcclintock, J., & Rowland, J. (2017). Brain-to-Brain Synchrony Tracks Real-World Report Brain-to-Brain Synchrony Tracks Real-World Dynamic Group Interactions in the Classroom. *Current Biology*, 1–6. https://doi.org/10.1016/j.cub.2017.04.002

Ding, J., Sperling, G., & Srinivasan, R. (2006). Attentional modulation of SSVEP power depends on the network tagged by the flicker frequency. *Cerebral Cortex*, *16*(7), 1016–1029. https://doi.org/10.1093/cercor/bhj044

Ding, N., & Simon, J. Z. (2012). Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. *Journal of Neurophysiology*, *107*(1), 78–89. https://doi.org/10.1152/jn.00297.2011

Ding, N., & Simon, J. Z. (2013). Power and phase properties of oscillatory neural responses in the presence of background activity. *Journal of Computational Neuroscience*, *34*(2), 337–343. https://doi.org/10.1007/s10827-012-0424-6

Ding, N., & Simon, J. Z. (2014). Cortical entrainment to continuous speech: Functional roles and interpretations. *Frontiers in Human Neuroscience*, *8*(MAY), 1–7. https://doi.org/10.3389/fnhum.2014.00311

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2015). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, 1–10. https://doi.org/10.1038/nn.4186

Doelling, K. B., Arnal, L. H., Ghitza, O., & Poeppel, D. (2014). NeuroImage Acoustic landmarks drive delta – theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage*, *85*, 761–768. https://doi.org/10.1016/j.neuroimage.2013.06.035

Doelling, K. B., Florencia Assaneo, M., Bevilacqua, D., Pesaran, B., & Poeppel, D. (2019). An oscillator model better predicts cortical entrainment to music. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(20), 10113–10121. https://doi.org/10.1073/pnas.1816414116

Doelling, K., Poeppel, D. (2015). Cortical entrainment to music and its modulation by expertise. *PNAS*, 6233–6242.

Dominguez, S., Devouche, E., Apter, G., & Gratier, M. (2016). The Roots of Turn-Taking in the Neonatal Period. *Infant and Child Development*, *25*, 240–254. https://doi.org/10.1002/icd

Drullman, R., Festen, J. M., & Plomp, R. (1994). Effect of reducing slow temporal modulations on speech reception. *Journal of the Acoustical Society of America*, *95*(5), 2670–2680. https://doi.org/10.1121/1.409836

Duprez, J., Stokkermans, M., Drijvers, L., & Cohen, M. X. (2021). Synchronization between Keyboard Typing and Neural Oscillations. *Journal of Cognitive Neuroscience 1:33(5)* 887-901. DOI: 10.1162/jocn_a_01692

Edalati, M., Wallois, F., Safaie, J., Ghostine, G., Kongolo, G., Trainor, L. J., & Moghimi, S. (2023). Rhythm in the Premature Neonate Brain: Very Early Processing of Auditory Beat and Meter. *Journal of Neuroscience*, *43*(15), 2794–2802. https://doi.org/10.1523/JNEUROSCI.1100-22.2023

Efron, B., & Tibshirani, R. (1993). *Bootstrp.m MATLAB function*. The Mathworks, Inc.

Einarson, K., & Trainor, L. (2016). Hearing the Beat: Young Children's Perceptual Sensitivity to Beat Alignment Varies According to Metric Structure. *Music Perception*, *34*(1), 56–70.

Eisermann, M., Kaminska, A., Moutard, M. L., Soufflet, C., & Plouin, P. (2013). Normal EEG in childhood: From neonates to adolescents. In *Neurophysiologie Clinique* (Vol. 43, Issue 1, pp. 35–65). https://doi.org/10.1016/j.neucli.2012.09.091

Elliott, M. R., Fisher, K., & Ames, E. W. (1988). The effects of rocking on the state and respiration of normal and excessive cryers. *Canadian Journal of Psychology*, *42*(2), 163–172. https://doi.org/10.1037/h0084186

Engel, A., Fries, P., & Singer, W. (2001). Dynamic Predictions: Oscillations and Synchrony in Top-Down Processing. *Nature Reviews Neuroscience*, *2*, 704–716.

Esposito, G., Yoshida, S., Ohnishi, R., Tsuneoka, Y., Del Carmen Rostagno, M., Yokota, S., Okabe, S., Kamiya, K., Hoshino, M., Shimizu, M., Venuti, P., Kikusui, T., Kato, T., & Kuroda, K. O. (2013). Infant calming responses during maternal carrying in humans and mice. *Current Biology*, *23*(9), 739–745. https://doi.org/10.1016/j.cub.2013.03.041

Fairhurst, M. T., & Dumas, G. (2019). *Reciprocity and alignment: quantifying coupling in dynamic interactions*. https://doi.org/10.31234/osf.io/nmg4x

Falk, S., & Kello, C. T. (2017). Hierarchical organization in the temporal structure of infant-direct speech and song. *Cognition*, *163*, 80–86. https://doi.org/10.1016/j.cognition.2017.02.017

Fallon, A., Van der Putten, D., Dring, C., Moylett, E. H., Fealy, G., & Devane, D. (2016). Baby-led compared with scheduled (or mixed) breastfeeding for successful breastfeeding. *Cochrane Database of Systematic Reviews*, *2016*(9). https://doi.org/10.1002/14651858.CD009067.pub3

Farhud, D., & Aryan, Z. (2018). Circadian rhythm, lifestyle and health: A narrative review. *Iranian Journal of Public Health*, *47*(8), 1068–1076.

Fawcett, C., Arslan, M., Falck-Ytter, T., Roeyers, H., & Gredebäck, G. (2017). Human eyes with dilated pupils induce pupillary contagion in infants. *Scientific Reports*, *7*(1), 2–8. https://doi.org/10.1038/s41598-017-08223-3

Feldman, R. (2003). Infant-mother and infant-father synchrony: The coregulation of positive arousal. *Infant Mental Health Journal*, *24*(1), 1–23. https://doi.org/10.1002/imhj.10041

Feldman, R. (2006). From biological rhythms to social rhythms: Physiological precursors of mother-infant synchrony. *Developmental Psychology*, *42*(1), 175–188. https://doi.org/10.1037/0012-1649.42.1.175

Feldman, R. (2007a). Parent-infant synchrony and the construction of shared timing; physiological precursors, developmental outcomes, and risk conditions. *Journal of Child Psychology and Psychiatry2*, *48*(3/4), 329–354.

Feldman, R. (2007b). Parent-infant synchrony: Biological foundations and developmental outcomes. *Current Directions in Psychological Science*, *16*(6), 340–345. https://doi.org/10.1111/j.1467-8721.2007.00532.x

Feldman, R. (2012). Bio-behavioral Synchrony: A Model for Integrating Biological and Microsocial Behavioral Processes in the Study of Parenting. *Parenting*, *12*(2–3), 154–164. https://doi.org/10.1080/15295192.2012.683342

Feldman, R. (2016). The neurobiology of mammalian parenting and the biosocial context of human caregiving. *Hormones and Behavior*, *77*, 3–17. https://doi.org/10.1016/j.yhbeh.2015.10.001

Feldman, R., Magori-Cohen, R., Galili, G., Singer, M., & Louzoun, Y. (2011). Mother and infant coordinate heart rhythms through episodes of interaction synchrony. *Infant Behavior and Development*, *34*(4), 569–577. https://doi.org/10.1016/j.infbeh.2011.06.008

Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults.: EBSCOhost. *American Psychological Association, Inc.*, *27*(2), 209–221. http://web.b.ebscohost.com.revproxy.brown.edu/ehost/pdfviewer/pdfviewer?sid=a7d41e1d-1263-4ba9-9b94-bbaf722e1f4e%40sessionmgr110&vid=1&hid=107

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, *20*(1), 104–113. https://doi.org/10.1037/0012-1649.20.1.104

Fiveash, A., Bedoin, N., Lalitte, P., & Tillmann, B. (2020). Rhythmic priming of grammaticality judgments in children: Duration matters. *Journal of Experimental Child Psychology, 197,* Article 104885. https://doi.org/10.1016/j.jecp.2020.104885

Flaten, E., Marshall, S. A., Dittrich, A., & Trainor, L. J. (2022). Evidence for top-down metre perception in infancy as shown by primed neural responses to an ambiguous rhythm. *European Journal of Neuroscience, 55*(8), 2003–2023. https://doi.org/10.1111/ejn.15671

Fló, A., Benjamin, L., Palu, M., & Dehaene-Lambertz, G. (2022). Sleeping neonates track transitional probabilities in speech but only retain the first syllable of *ntific Reports*, *12*(1), 1–13. https://doi.org/10.1038/s41598-022-08411-w

Fraisse, P. (1982). Rhythm and Tempo *in The Psychology of Music*. Academic Press Inc. https://doi.org/10.1016/C2009-0-03009-8

Fries, P. (2015). Rhythms for Cognition: Communication through Coherence. *Neuron*, *88*(1), 220–235. https://doi.org/10.1016/j.neuron.2015.09.034

Frischen, U., Degé, F., & Schwarzer, G. (2022). The relation between rhythm processing and cognitive abilities during child development: The role of prediction. *Frontiers in Psychology*, *13*(September), 1–11. https://doi.org/10.3389/fpsyg.2022.920513

Galambos, R., Makeig, S., & Talmachoff, P. J. (1981). A 40-Hz auditory potential recorded from the human scalp. *Proceedings of the National Academy of Sciences of the United States of America*, *78*(4 II), 2643–2647. https://doi.org/10.1073/pnas.78.4.2643

Gasparini, L., Langus, A., Tsuji, S., Boll-Avetisyan, N. (2021). Quantifying the role of rhythm in infants' language discrimination abilities: A meta-analysis. *Cognition, 213,* 104757

Ghinst, M. Vander, Bourguignon, M., Niesen, M., Wens, V., Hassid, S., Choufani, G., Jousmäki, V., Hari, R., Goldman, S., & De Tiège, X. (2019). Cortical tracking of speech-in-noise develops from childhood to adulthood. *Journal of Neuroscience*, *39*(15), 2938–2950. https://doi.org/10.1523/JNEUROSCI.1732-18.2019

Gibbon, S., Attaheri, A., Ní Choisdealbha, Á., Rocha, S., Brusini, P., Mead, N., Boutris, P., Olawole-Scott, H., Ahmed, H., Flanagan, S., Mandke, K., Keshavarzi, M., & Goswami, U. (2021). Machine learning accurately classifies neural responses to rhythmic speech vs. non-speech from 8-week-old infant EEG. *Brain and Language*, *220*(July 2020), 0–6. https://doi.org/10.1016/j.bandl.2021.104968

Gingras, J. L., Mitchell, E. A., & Grattan, K. E. (2005). Fetal homologue of infant crying. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, *90*(5), 415–418. https://doi.org/10.1136/adc.2004.062257

Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511–517. https://doi.org/10.1038/nn.3063

Glass, L. (2001). Synchronization and rhythmic processes in physiology. *Nature*, *410*(March), 277–284.

Gliga, T., Farroni, T., & Cascio, C. J. (2019). Social touch: A new vista for developmental cognitive neuroscience? *Developmental Cognitive Neuroscience*, *35*(May 2018), 1–4. https://doi.org/10.1016/j.dcn.2018.05.006

Golinkoff, R. M., Can, D. D., Soderstrom, M., & Hirsh-Pasek, K. (2015). (Baby)Talk to Me: The Social Context of Infant-Directed Speech and Its Effects on Early Language Acquisition. *Current Directions in Psychological Science*, *24*(5), 339–344. https://doi.org/10.1177/0963721415595345

Golub, H. L. (1979). A physioacoustic model of the infant cry and its use for medical diagnosis and prognosis. *The Journal of the Acoustical Society of America*, *65*(S1), S25–S26. https://doi.org/10.1121/1.2017179

Golumbic, E. M. Z., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., Mckhann, G. M., Goodman, R. R., Emerson, R., Mehta, A. D., Simon, J. Z., & Poeppel, D. (2012). Article Mechanisms Underlying Selective Neuronal Tracking of Attended Speech at a "" Cocktail Party ."' *Neuron*, *77*(5), 980–991. https://doi.org/10.1016/j.neuron.2012.12.037

Gomez-Ramirez, M., Kelly, S. P., Molholm, S., Sehatpour, P., Schwartz, T. H., & Foxe, J. J. (2011). Oscillatory sensory selection mechanisms during intersensory attention to rhythmic auditory and visual inputs: A human electrocorticographic investigation. *Journal of Neuroscience*, *31*(50), 18556–18567. https://doi.org/10.1523/JNEUROSCI.2164-11.2011

Gordon, R. L., Shivers, S. M., Wieland, E. A., Kotz, S. A., Yoder, P. J., & McAuley, J, D. (20154). Musical rhythm discrimination explains individual differences in grammar skills in children. *Developmental Science, 18(4), 635-644.* https://doi.org/10.1111/desc.12230

Goswami, U. (2018). A Neural Basis for Phonological Awareness? An Oscillatory Temporal-Sampling Perspective. *Current Directions in Psychological Science*, *27*(1), 56–63. https://doi.org/10.1177/0963721417727520

Goswami, U., & Leong, V. (2013). Speech rhythm and temporal structure: Converging perspectives? *Laboratory Phonology*, *4*(1), 67–92. https://doi.org/10.1515/lp-2013-0004

Greenberg, S., Carvey, H., Hitchcock, L., & Chang, S. (2003). Temporal properties of spontaneous speech - A syllable-centric perspective. *Journal of Phonetics*, *31*(3–4), 465–485. https://doi.org/10.1016/j.wocn.2003.09.005

Gross, J., Kluger, D. S., Abbasi, O., Chalas, N., Steingräber, N., Daube, C., & Schoffelen, J. M. (2021). Comparison of undirected frequency-domain connectivity measures for cerebro-peripheral analysis. *NeuroImage*, *245*(September), 118660. https://doi.org/10.1016/j.neuroimage.2021.118660

Gross, J., Kluger, D. S., Abbasi, O., Chalas, N., Steingräber, N., Daube, C., & Schoffelen, J. M. (2021). Comparison of undirected frequency-domain connectivity measures for cerebro-peripheral analysis. *NeuroImage*, *245*(September), 118660. https://doi.org/10.1016/j.neuroimage.2021.118660

Gulbinaite, R., Ilhan, B., & Vanrullen, R. (2016). Something out of nothing: The role of alpha-frequency reverberation in the triple-flash illusion. *Journal of Vision*, *16*(12), 1128.

Gulbinaite, R., İlhan, B., & Vanrullen, R. (2017). The triple-flash illusion reveals a driving role of alpha-band reverberations in visual perception. *Journal of*

Neuroscience, *37*(30), 7219–7230. https://doi.org/10.1523/JNEUROSCI.3929-16.2017

Guo, Y., Bufacchi, R. J., Novembre, G., Kilintari, M., Moayedi, M., Hu, L., & Iannetti, G. D. (2020). Ultralow-frequency neural entrainment to pain. *PLoS Biology*, *18*(4), 1–27. https://doi.org/10.1371/journal.pbio.3000491

Háden, G. P., Bouwer, F. L., Honing, H., Winkler, I., & Bouwer, F. (2022). Beat processing in newborn infants cannot be explained by statistical learning based on transition probabilities. In *bioRxiv*. https://doi.org/10.17605/OSF.IO/EFQBK

Háden, G. P., Honing, H., Török, M., & Winkler, I. (2015). Detecting the temporal structure of sound sequences in newborn infants. *International Journal of Psychophysiology*, *96*(1), 23–28. https://doi.org/10.1016/j.ijpsycho.2015.02.024

Haegens, S., & Zion Golumbic, E. (2018). Rhythmic facilitation of sensory processing: A critical review. *Neuroscience and Biobehavioral Reviews*, *86*(November 2017), 150–165. https://doi.org/10.1016/j.neubiorev.2017.12.002

Hamilton, A. (2020, June 22). Hype, hyperscanning and embodied social neuroscience. https://doi.org/10.31234/osf.io/rc9wp

Hannon, E. E., & Johnson, S. P. (2005). Infants use meter to categorize rhythms and melodies: Implications for musical structure learning. *Cognitive Psychology*, *50*(4), 354–377. https://doi.org/10.1016/j.cogpsych.2004.09.003

Hannon, E. E., & Trehub, S. E. (2005). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences of* *the United States of America*, *102*(35), 12639–12643. https://doi.org/10.1073/pnas.0504254102

Hannon, E. E., Vanden Bosch der Nederlanden, C. M., & Tichko, P. (2012). Effects of perceptual experience on children's and adults' perception of unfamiliar rhythms. *Annals of the New York Academy of Sciences*, *1252*(1), 92–99. https://doi.org/10.1111/j.1749-6632.2012.06466.x

Hari, R., Hämäläinen, M., & Joutsiniemi, S. L. (1989). Neuromagnetic steady-state responses to auditory stimuli. *Journal of the Acoustical Society of America*, *86*(3), 1033–1039. https://doi.org/10.1121/1.398093

Harris, K. D. (2020). Nonsense correlations in neuroscience. *BioRxiv*, 2020.11.29.402719. https://doi.org/10.1101/2020.11.29.402719

Haslbeck, F. B., & Bassler, D. (2018). Music from the very beginning—A neuroscience-based framework for music as therapy for preterm infants and their parents. *Frontiers in Behavioral Neuroscience*, *12*(June), 1–7. https://doi.org/10.3389/fnbeh.2018.00112

Hasson, U., & Frith, C. D. (2016). Mirroring and beyond: Coupled dynamics as a generalized framework for modelling social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1693). https://doi.org/10.1098/rstb.2015.0366

Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J. D., Blankertz, B., & Bießmann, F. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage*, *87*, 96–110. https://doi.org/10.1016/j.neuroimage.2013.10.067

He, W., Donoghue, T., Sowman, P. F., Seymour, R. A., Brock, J., Crain, S., Voytek, B., & Hillebrand, A. (2019). Co-increasing neuronal noise and beta power in the developing brain. *BioRxiv*, *December*. https://doi.org/10.1101/839258

Henao, D., Navarrete, M., Valderrama, M., & Le Van Quyen, M. (2020). Entrainment and synchronization of brain oscillations to auditory stimulations. *Neuroscience Research*, *156*, 271–278. https://doi.org/10.1016/j.neures.2020.03.004

Henry, M. J., & Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(49), 20095–20100. https://doi.org/10.1073/pnas.1213390109

Henry, M. J., & Hermann, B. (2014). Low-Frequency Neural Oscillations Support Dynamic Attending in Temporal Context. *Timing & Time Perception 2, 2(1),* 62-86. https://doi.org/10.1163/22134468-00002011

Henry, M. J., Herrmann, B., & Grahn, J. A. (2017). What can we learn about beat perception by comparing brain signals and stimulus envelopes? *PLoS ONE*, *12*(2), 1–17. https://doi.org/10.1371/journal.pone.0172454

Herdman, A. T., Lins, O., Van Roon, P., Stapells, D. R., Scherg, M., & Picton, T. W. (2002). Intracerebral sources of human auditory steady-state responses. *Brain Topography*, *15*(2), 69–86. https://doi.org/10.1023/A:1021470822922

Hessels, R. S. (2020). How does gaze to faces support face-to-face interaction? A review and perspective. *Psychonomic Bulletin and Review*, *27*(5), 856–881. https://doi.org/10.3758/s13423-020-01715-w

Hickey, P., Merseal, H., Patel, A. D., & Race, E. (2020). NeuroImage Memory in time: Neural tracking of low-frequency rhythm dynamically modulates memory formation. *NeuroImage*, *213*(February), 116693. https://doi.org/10.1016/j.neuroimage.2020.116693

Hoehl, S., Fairhurst, M., & Schirmer, A. (2021). Interactional synchrony: Signals, mechanisms and benefits. *Social Cognitive and Affective Neuroscience*, *16*(1–2), 5–18. https://doi.org/10.1093/scan/nsaa024

Hofstra, W. A., & de Weerd, A. W. (2008). How to assess circadian rhythm in humans: A review of literature. *Epilepsy and Behavior*, *13*(3), 438–444. https://doi.org/10.1016/j.yebeh.2008.06.002

Höhle, B., Pauen, S., Hesse, V., & Weissenborn, J. (2014). Discrimination of rhythmic pattern at 4 months and language performance at 5 years: A longitudinal analysis of data from German-learning children. *Language Learning, 64*(Suppl 2), 141–164. https://doi.org/10.1111/lang.12075

Holler, J., Kendrick, K. H., Casillas, M., & Levinson, S. C. (2015). Editorial: Turn-Taking in Human Communicative Interaction. In *Frontiers in Psychology* (Vol. 6, Issue DEC). https://doi.org/10.3389/fpsyg.2015.01919

Hunziker, U. A., & Barr, R. G. (1986). Increased carrying reduces infant crying: A randomized controlled trial. *Pediatrics*, *77*(5), 641–648. https://doi.org/10.1542/peds.77.5.641

Hyafil, A., Fontolan, L., Kabdebon, C., Gutkin, B., & Giraud, A. L. (2015). Speech encoding by coupled cortical theta and gamma oscillations. *Elife*, *4*(MAY), 1–45. https://doi.org/10.7554/eLife.06213

Ilari, B., (2005). On musical parenting of young children:,musical beliefs and behaviors of mothers and infants. *Early Child Development and Care, 175(7-8),* 647-660. https://doi.org/10.1080/0300443042000302573

Inui, N. (2018). An Overview of the Study on Interpersonal Coordination. In *Interpersonal Coordination* (Issue Clark 1996). https://doi.org/10.1007/978-981-13-1765-1_3

Ives, J., Labendzki, P., Perapoch Amadó, M., Greenwood, E., Viswanathan, N., & Wass, S. (preprint). At Which Low Amplitude Modulated Frequency Do Infants Best Entrain? A Frequency Tagging Study. *BioRxiv*. https://doi.org/10.1101/2022.12.08.519576

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. D. (2001). Rhythms of dialogue in infancy: coordinated timing in development. *Monographs of the Society for Research in Child Development*, *66*(2), i–viii, 1–132. http://www.ncbi.nlm.nih.gov/pubmed/11428150

James-Roberts, I. S., & Halil, T. (1991). Infant Crying Patterns in the First Year: Normal Community and Clinical Findings. *Journal of Child Psychology and Psychiatry*, *32*(6), 951–968. https://doi.org/10.1111/j.1469-7610.1991.tb01922.x

Jerger, J., Chmiel, R., Frost, J. D., & Coker, N. (1986). Effect of sleep on the auditory steady state evoked potential. *Ear and Hearing*, *7*(4), 240–245. https://doi.org/10.1097/00003446-198608000-00004

Jervis, B. W., Nichols, M. J., Johnson, T. E., Allen, E., & Hudson, N. R. (1983). A Fundamental Investigation of the Composition of Auditory Evoked Potentials.

IEEE Transactions on Biomedical Engineering, *BME-30*(1), 43–50. https://doi.org/10.1109/TBME.1983.325165

Jessen, S., Fiedler, L., Münte, T. F., & Obleser, J. (2019). Quantifying the individual auditory and visual brain response in 7-month-old infants watching a brief cartoon movie. *NeuroImage*, *202*(April), 116060. https://doi.org/10.1016/j.neuroimage.2019.116060

Jessen, S., Obleser, J., & Tune, S. (2021). Neural tracking in infants – An analytical tool for multisensory social processing in development. *Developmental Cognitive Neuroscience*, *52*, 101034. https://doi.org/10.1016/j.dcn.2021.101034

Jiang, J., Dai, B., Peng, D., Zhu, C., Liu, L., & Lu, C. (2012). Neural Synchronization during Face-to-Face Communication. *The Journal of Neuroscience*, *32*(45), 16064–16069. https://doi.org/10.1523/JNEUROSCI.2926-12.2012

Jones, E. J. H., Venema, K., Lowy, R., Earl, R. K., & Webb, S. J. (2015). Developmental changes in infant brain activity during naturalistic social experiences. *Developmental Psychobiology*, *57*(7), 842–853. https://doi.org/10.1002/dev.21336

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review, 83*(5), 323–355. https://doi.org/10.1037/0033-295X.83.5.323

Kabdebon, C., Fló, A., Heering, A. De, & Aslin, R. (2022). The power of rhythms: how steady-state evoked responses reveal early neurocognitive development. *NeuroImage*, *254*(August 2021), 119150. https://doi.org/10.1016/j.neuroimage.2022.119150

Kabdebon, C., Pena, M., Buiatti, M., & Dehaene-lambertz, G. (2015). Brain & Language Electrophysiological evidence of statistical learning of long-distance dependencies in 8-month-old preterm and full-term infants. *Brain and Language*, *148*, 25–36. https://doi.org/10.1016/j.bandl.2015.03.005

Kalashnikova, M., & Kember, H. (2020). Prosodic cues in infant-directed speech facilitate young children's conversational turn predictions. *Journal of Experimental Child Psychology, 199,* Article 104916. https://doi.org/10.1016/j.jecp.2020.104916

Kalashnikova, M., Peter, V., Liberto, G. M. Di, & Lalor, E. C. (2018). Infant-directed speech facilitates seven-month-old infants ' cortical tracking of speech. *Scientific Reports*, *April*, 1–8. https://doi.org/10.1038/s41598-018-32150-6

Kandylaki, K. D., & Kotz, S. A. (2020). Distinct cortical rhythms in speech and language processing and some more : a commentary on. *Language, Cognition and Neuroscience*, *0*(0), 1–5. https://doi.org/10.1080/23273798.2020.1757729

Kaneshiro, B., Nguyen, D. T., Norcia, A. M., Dmochowski, J. P., & Berger, J. (2020). Natural music evokes correlated EEG responses reflecting temporal structure and beat. *NeuroImage*, *214*(January), 116559. https://doi.org/10.1016/j.neuroimage.2020.116559

Kawasaki, M., Yamada, Y., Ushiku, Y., Miyauchi, E., & Yamaguchi, Y. (2013). *Inter-brain synchronization during coordination of speech rhythm in human-to-human social interaction*. 1–8. https://doi.org/10.1038/srep01692

Kayhan, E., Nguyen, T., Matthes, D., Langeloh, M., Michel, C., & Jiang, J. (2022). Interpersonal neural synchrony when predicting others' actions during a game of rock - paper - scissors. *Scientific Reports*, 1–11. https://doi.org/10.1038/s41598-022-16956-z

Kayser, C., Wilson, C., Safaai, H., Sakata, S., & Panzeri, S. (2015). Rhythmic auditory cortex activity at multiple timescales shapes stimulus–response gain and background firing. *Journal of Neuroscience*, *35*(20), 7750–7762. https://doi.org/10.1523/JNEUROSCI.0268-15.2015

Kayser, S. J., Ince, R. A. A., Gross, J., & Kayser, C. (2015). Irregular speech rate dissociates auditory cortical entrainment, evoked responses, and frontal alpha. *Journal of Neuroscience*, *35*(44), 14691–14701. https://doi.org/10.1523/JNEUROSCI.2243-15.2015

Keitel, C., Quigley, C., & Ruhnau, P. (2014). Stimulus-driven brain oscillations in the alpha range: Entrainment of intrinsic rhythms or frequency-following response? *Journal of Neuroscience*, *34*(31), 10137–10140. https://doi.org/10.1523/JNEUROSCI.1904-14.2014

Keitel, C., Thut, G., & Gross, J. (2017). Visual cortex responses reflect temporal structure of continuous quasi-rhythmic sensory stimulation. *NeuroImage*, *146*(November 2016), 58–70. https://doi.org/10.1016/j.neuroimage.2016.11.043

Keller, H., Schölmerich, A., & Eibl-Eibesfeldt, I. (1988). Communication Patterns in Adult-Infant Interactions in Western and Non-Western Cultures. *Journal of Cross-Cultural Psychology1*, *19*(4), 427–445.

Kimbrough Oller, D. (2000). *The Emergence of the Speech Capability* (1st ed.). Taylor & Francis.

Kingsbury, L., Huang, S., Wang, J., Gu, K., Golshani, P., Wu, Y. E., & Hong, W. (2019). Correlated Neural Activity and Encoding of Behavior across Brains of Socially Interacting Animals. *Cell*, *178*(2), 429-446.e16. https://doi.org/10.1016/j.cell.2019.05.022

Kinsbourne, M. (2002). The role of imitation in body ownership and mental growth. In A. N. Meltzoff & W. Prinz (Eds.), *The imitative mind: development, evolution, and brain bases* (pp. 311–330). Cambridge University Press.

Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology*, *102*(3), 299–314. https://doi.org/10.1016/j.jecp.2008.07.005

Korner, A. F., Guilleminault, C., Van den Hoed, J., & Baldwin, R. B. (1978). Reduction of sleep apnea and bradycardia in preterm infants on oscillating water beds: a controlled polygraphic study. *Pediatrics*, *61*(4), 528–533.

Kösem, A., Gramfort, A., & Wassenhove, V. Van. (2014). NeuroImage Encoding of event timing in the phase of neural oscillations. *NeuroImage*, *92*, 274–284. https://doi.org/10.1016/j.neuroimage.2014.02.010

Köster, M., Brzozowska, A., Bánki, A., Tünte, M., Ward, E. M., & Hoehl, S. (2023). Rhythmic visual stimulation as a window into early brain development: A systematic review. *Developmental Cognitive Neurosciece, 64, 101315*. https://doi.org/10.1016/j.dcn.2023.101315

Koskinen, M., & Seppä, M. (2014). NeuroImage Uncovering cortical MEG responses to listened audiobook stories. *NeuroImage*, *100*, 263–270. https://doi.org/10.1016/j.neuroimage.2014.06.018

Kothe, C. (2010). *clean_windows EEGLAB plugin*. The Mathworks, Inc.

Kothe, C. (2014). *Clean_channels EEGLAb plugin*. The Mathworks, Inc.

Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The Evolution of Rhythm Processing. *Trends in Cognitive Sciences*, *22*(10), 896–910. https://doi.org/10.1016/j.tics.2018.08.002

Koukkari, W. L., & Sothern, R. B. (2007). Introducing Biological Rhythms. *Springer* Dordrecht. https://doi.org/10.1007/978-1-4020-4701-5

Kozulin, P., Almarza, G., Gobius, I., & Richards, L. J. (2016). Prenatal and Postnatal Determinants of Development. *Neuromethods*, *109*, 3–20. https://doi.org/10.1007/978-1-4939-3014-2

Kumar, J., Meena, J., Debata, P., MJeeva, S., Kumar, P., & Arvind, S. (2022). Three-hourly versus two-hourly feeding interval in stable preterm infants: an updated systematic review and meta-analysis of randomized controlled trials. *European Journal of Pediatrics*, *181*, 2075–2086.

Kurth, E., Kennedy, H. P., Spichiger, E., Hösli, I., & Zemp Stutz, E. (2011). Crying babies, tired mothers: What do we know? A systematic review. *Midwifery*, *27*(2), 187–194. https://doi.org/10.1016/j.midw.2009.05.012

Kushnerenko, E., Eponiene, R., Balan, P., Fellman, V., Huotilainen, M., & Näätänen, R. (2002). Maturation of the auditory event-related potentials during the first year of life. *NeuroReport*, *13*(1), 47–51. https://doi.org/10.1097/00001756-200201210-00014

Kuwada, S., Batra, R., & Maher, V. L. (1986). Scalp potentials of normal and hearing-impaired subjects in response to sinusoidally amplitude-modulated tones. *Hearing Research*, *21*(2), 179–192. https://doi.org/10.1016/0378-5955(86)90038-9

Lachaux, J. P., Rodriguez, E., Martinerie, J., & Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Human Brain Mapping*, *8*(4), 194–208. https://doi.org/10.1002/(SICI)1097-0193(1999)8:4<194::AID-HBM4>3.0.CO;2-C

Ladányi, E., Persici, V., Fiveash, A., Tillmann, B., & Gordon, R. L. (2020). Is atypical rhythm a risk factor for developmental speech and language disorders? *WIREs Cognitive Science, 11*(5), Article e1528. https://doi.org/10.1002/wcs.1528

Lahmiri, S., Tadj, C., Gargour, C., & Bekiros, S. (2021). Characterization of infant healthy and pathological cry signals in cepstrum domain based on approximate entropy and correlation dimension. *Chaos, Solitons and Fractals*, *143*, 110639. https://doi.org/10.1016/j.chaos.2020.110639

Lakatos, P., Gross, J., & Thut, G. (2019). A New Unifying Account of the Roles of Neuronal Entrainment. *Current Biology*, *29*(18), R890–R905. https://doi.org/10.1016/j.cub.2019.07.075

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*, *320*(5872), 110–113. https://doi.org/10.1126/science.1154735

Lakatos, P., Musacchia, G., O'Connel, M. N., Falchier, A. Y., Javitt, D. C., & Schroeder, C. E. (2013). The Spectrotemporal Filter Mechanism of Auditory Selective Attention. *Neuron*, *77*(4), 750–761. https://doi.org/10.1016/j.neuron.2012.11.034

Langus, A., Mehler, J., & Nespor, M. (2017). Rhythm in language acquisition. *Neuroscience and Biobehavioral Reviews, 81*(Part B), 158–166. https://doi.org/10.1016/j.neubiorev.2016.12.012

Lankinen, K., Saari, J., Hari, R., & Koskinen, M. (2014). Intersubject consistency of cortical MEG signals during movie viewing. *NeuroImage*, *92*, 217–224. https://doi.org/10.1016/j.neuroimage.2014.02.004

Lau, J. C. Y., Fyshe, A., & Waxman, S. R. (2022). Rhythm May Be Key to Linking Language and Cognition in Young Infants: Evidence From Machine Learning. *Frontiers in Psychology*, *13*(May), 1–10. https://doi.org/10.3389/fpsyg.2022.894405

Lecanuet, J. P., & Jacquet, A. Y. (2002). Fetal responsiveness to maternal passive swinging in low heart rate variability state: Effects of stimulation direction and duration. *Developmental Psychobiology*, *40*(1), 57–67. https://doi.org/10.1002/dev.10013

Lehmann, A., Arias, D. J., & Schönwiesner, M. (2016). Tracing the neural basis of auditory entrainment. *Neuroscience*, *337*(September), 306–314. https://doi.org/10.1016/j.neuroscience.2016.09.011

Lense, M. D., Shultz, S., Astésano, C., & Jones, W. (2022). Music of infant-directed singing entrains infants' social visual behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(45), 1–9. https://doi.org/10.1073/pnas.2116967119

Leong, V., & Goswami, U. (2015). Acoustic-emergent phonology in the amplitude envelope of child-directed speech. *PLoS ONE*, *10*(12), 1–37. https://doi.org/10.1371/journal.pone.0144411

Leong, V., Byrne, E., Clackson, K., Georgieva, S., Lam, S., & Wass, S. (2017). Speaker gaze increases information coupling between infant and adult brains. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(50), 13290–13295. https://doi.org/10.1073/pnas.1702493114

Leong, V., Kalashnikova, M., Burnham, D., & Goswami, U. (2014). Infant-directed speech enhances temporal rhythmic structure in the envelope. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, *September*, 2563–2567. https://doi.org/10.21437/interspeech.2014-549

Leong, V., Kalashnikova, M., Burnham, D., & Goswami, U. (2017). The Temporal Modulation Structure of Infant-Directed Speech. *Open Mind*, *1*(2), 78–90. https://doi.org/10.1162/opmi_a_00008

Leong, V., Noreika, V., Clackson, K., Georgieva, S., Brightman, L., Nutbrown, R., Fujita, S., Neale, D., & Wass, S. (2019). *Mother-infant interpersonal neural connectivity predicts infants' social learning*. https://doi.org/10.31234/osf.io/gueaq

Leong, V., Stone, M. A., Turner, R. E., & Goswami, U. (2014). A role for amplitude modulation phase relationships in speech rhythm perception. *The Journal of the Acoustical Society of America*, *136*(1), 366–381. https://doi.org/10.1121/1.4883366

Lerousseau, J. P., Trébuchon, A., Morillon, B., & Schön, D. (2021). Frequency selectivity of persistent cortical oscillatory responses to auditory rhythmic stimulation. *Journal of Neuroscience*, *41*(38), 7991–8006. https://doi.org/10.1523/JNEUROSCI.0213-21.2021

Levi, E. C., Folsom, R. C., & Dobie, R. A. (1993). Amplitude-modulation following response (AMFR): Effects of modulation rate, carrier frequency, age, and state. *Hearing Research*, *68*(1), 42–52. https://doi.org/10.1016/0378-5955(93)90063-7

Levitan, R., & Hirschberg, J. (2011). Measuring acoustic-prosodic entrainment with respect to multiple levels and dimensions. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 3081–3084. https://doi.org/10.21437/interspeech.2011-771

Liégeois-Chauvel, C., Lorenzi, C., Trébuchon, A., Régis, J., & Chauvel, P. (2004). Temporal envelope processing in the human left and right auditory cortices. *Cerebral Cortex*, *14*(7), 731–740. https://doi.org/10.1093/cercor/bhh033

Llinas, R. R. (1988). Intrinsic Electrophysiological Properties Central Nervous System Function. *Science*, *242*, 1654–1664.

Lo, C.-W., Tung, T.-Y., Ke, A. H., & Brennan, J. R. (2022). Hierarchy, Not Lexical Regularity, Modulates Low-Frequency Neural Synchrony During Language Comprehension. *Neurobiology of Language*, *3*(4), 538–555. https://doi.org/10.1162/nol_a_00077

Lorenzini, I., Labendzki, P., Hababou, M., & Basire, C. (2022). Human neural processing of auditory temporal modulations during the first year of life. *PsyArXiv*, *(Preprint)*, 1–21.

Ludington-Hoe, S. M., Cong, X., & Hashemi, F. (2002). Infant crying: nature, physiologic consequences, and select interventions. *Neonatal Network : NN*, *21*(2), 29–36. https://doi.org/10.1891/0730-0832.21.2.29

Luo, H., Wang, Y., Poeppel, D., & Simon, J. Z. (2006). Concurrent encoding of frequency and amplitude modulation in human auditory cortex: MEG evidence. *Journal of Neurophysiology*, *96*(5), 2712–2723. https://doi.org/10.1152/jn.01256.2005

MacKinnon, J. L., Ames, E. W., Davis, C. M., Carfanlan, M., Pierlot, L. K., & Lemke, K. (1986). Effects of rocking on infant respiration. *Infant Behaviour & Development*, *9*, 234.

Maitha, C., Goode, J. C., Maulucci, D. P., Lasassmeh, S. M. S., Yu, C., Smith, L. B., & Borjon, J. I. (2020). *An open-source , wireless vest for measuring autonomic function in infants*.

Makeig, S., Westerfield, M., Jung, T. P., Enghoff, S., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2002). Dynamic brain sources of visual evoked responses. *Science*, *295*(5555), 690–694. https://doi.org/10.1126/science.1066168

Mampe, B., Friederici, A. D., Christophe, A., & Wermke, K. (2009). Newborns' Cry Melody Is Shaped by Their Native Language. *Current Biology*, *19*(23), 1994–1997. https://doi.org/10.1016/j.cub.2009.09.064

Mancini, F., Pepe, A., Bernacchia, A., Di Stefano, G., Mouraux, A., & Iannetti, G. D. (2018). Characterizing the short-term habituation of event-related evoked potentials. *ENeuro*, *5*(5), 1–14. https://doi.org/10.1523/ENEURO.0014-18.2018

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake –but only for skilled producers. Journal of Experimental Psychology: Human Perception and Performance 38 (4). 843–847. DOI: 10.1037/a0029284

Mannix, P., Inwald, D., hathorn, M. Coseloe, K. (1997). Tehrman Entrainment of Heart Rate in the Preterm Infant. *Pediatric Research*, *42*, 282–286.

Marie, C., Magne, C., & Besson, M. (2011). Musicians and the metric structure of words. *Journal of Cognitive Neuroscience, 23*(2), 294–305. https://doi.org/10.1162/jocn.2010.21413

Markova, G., Nguyen, T., & Hoehl, S. (2019). Neurobehavioral Interpersonal Synchrony in Early Development: The Role of Interactional Rhythms. *Frontiers in Psychology*, *10*(September). https://doi.org/10.3389/fpsyg.2019.02078

Marriott Haresign, I., Phillips, E., Whitehorn, M., Noreika, V., Jones, E. J. H., Leong, V., & Wass, S. V. (2021). Automatic classification of ICA components from infant EEG using MARA. *Developmental Cognitive Neuroscience*, *52*(October), 101024. https://doi.org/10.1016/j.dcn.2021.101024

Marshall, P. J., Bar-Haim, Y., & Fox, N. A. (2002). Development of the EEG from 5 months to 4 years of age. *Clinical Neurophysiology*, *113*(8), 1199–1208. https://doi.org/10.1016/S1388-2457(02)00163-3

Masataka, N. (1999). Preference for infant-directed singing in 2-day-old hearing infants of deaf parents. *Developmental Psychology*, *35*(4), 1001–1005. https://doi.org/10.1037/0012-1649.35.4.1001

Matsumoto, H., & Ugawa, Y. (2017). Adverse events of tDCS and tACS: A review. *Clinical Neurophysiology Practice*, *2*, 19–25. https://doi.org/10.1016/j.cnp.2016.12.003

Maurizi, M., Almadori, G., Paludetti, G., Ottavani, F., Rosignoli, M., Luciano, R. (1990). 40-Hz Steady-State Responses in Newborns and in Children. *Audiology*, *29*, 322–328.

Mazaheri, A., & Jensen, O. (2010). Rhythmic pulsing: Linking ongoing brain activity with evoked responses. *Frontiers in Human Neuroscience*, *4*(October), 1–13. https://doi.org/10.3389/fnhum.2010.00177

McAdams, S., & Bertoncini, J. (1997). Organization and discrimination of repeating sound sequences by newborn infants. *The Journal of the Acoustical Society of America*, *102*(5), 2945–2953. https://doi.org/10.1121/1.420349

McCulloch, D. L., Orbach, H., & Skarf, B. (1999). Maturation of the pattern-reversal VEP in human infants: A theoretical framework. *Vision Research*, *39*(22), 3673–3680. https://doi.org/10.1016/S0042-6989(99)00091-7

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, *29*(2), 143–178. https://doi.org/10.1016/0010-0277(88)90035-2

Mendoza, J. K., & Fausey, C. M. (2021). Everyday music in infancy. *Developmental Science*, *24*(6), 1–15. https://doi.org/10.1111/desc.13122

Menn, K. H., Michel, C., Meyer, L., Hoehl, S., & Männel, C. (2022). Natural infant-directed speech facilitates neural tracking of prosody. *NeuroImage*, *251* (2021), 118991. https://doi.org/10.1016/j.neuroimage.2022.118991

Meyer, L., Sun, Y., & Martin, A. E. (2019). Synchronous, but not entrained : exogenous and endogenous cortical rhythms of speech and language processing. *Language, Cognition and Neuroscience 35*(9), 1089–1099. https://doi.org/10.1080/23273798.2019.1693050

Meyer, L., Sun, Y., & Martin, A. E. (2020). “Entraining” to speech, generating language? *Language, Cognition and Neuroscience*, *35*(9), 1138–1148. https://doi.org/10.1080/23273798.2020.1827155

Michel, C. M., & He, B. (2019). EEG source localization. In *Handbook of Clinical Neurology* (1st ed., Vol. 160). Elsevier B.V. https://doi.org/10.1016/B978-0-444-64032-1.00006-0

Micheletti, M., Yao, X., Johnson, M., & de Barbaro, K. (2022). Validating a model to detect infant crying from naturalistic audio. *Behavior Research Methods*, *August*. https://doi.org/10.3758/s13428-022-01961-x

Miller, G. (1951). *Language and communication*. McGraw-Hill.

Millman, R. E., Prendergast, G., Kitterick, P. T., Woods, W. P., & Green, G. G. R. (2010). Spatiotemporal reconstruction of the auditory steady-state response to frequency modulation using magnetoencephalography. *NeuroImage*, *49*(1), 745–758. https://doi.org/10.1016/j.neuroimage.2009.08.029

Mirmiran, M., & Kok, J. H. (1991). Circadian rhythms in early human development. *Early Human Development*, *26*(2), 121–128. https://doi.org/10.1016/0378-3782(91)90016-V

Mogan, R., Fischer, R., & Bulbulia, J. A. (2017). To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect. *Journal of Experimental Social Psychology*, *72*(September 2016), 13–20. https://doi.org/10.1016/j.jesp.2017.03.009

Möller, S., & Schönweiler, R. (1997). Analysis of Infant Cries for the Early Detection of Hearing Impairment. *5th European Conference on Speech Communication and Technology, EUROSPEECH 1997*, *28*, 1759–1762. https://doi.org/10.21437/eurospeech.1997-301

Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, *16*(4), 495–500. https://doi.org/10.1016/0163-6383(93)80007-U

Moritz, C., Yampolsky, S., Papadelis, G., Thomson, J., & Wolf, M. (2013). Links between early rhythm skills, musical training, and phonological awareness. *Reading and Writing: An Interdisciplinary Journal, 26*(5), 739–769. https://doi.org/10.1007/s11145-012-9389-0

Mühler, R., Rahne, T., & Verhey, J. L. (2013). Auditory brainstem responses to broad-band chirps: Amplitude growth functions in sedated and anaesthetised infants. *International Journal of Pediatric Otorhinolaryngology*, *77*(1), 49–53. https://doi.org/10.1016/j.ijporl.2012.09.028

Mullen, T. (2012). *CleanLine EEGLAB plugin*. The Mathworks, Inc.

Murray, L., & Cooper, P. (2002). The Impact of Irritable Infant Behavior on Maternal Mental State: A Longitudinal Study and a Treatment Trial. In R. Barr, I. St. James-Roberts, & M. Keefe (Eds.), *New Evidence on Unexplained Early Infant Crying: Its Origins, Nature and Management.* (Vol. 34, Issue 1, pp. 96–97). Johnson & Johnson. https://doi.org/10.1097/00005176-200201000-00026

Murray, L., & Trevarthen, C. (1986). The infant's role in mother-infant communications. *Journal of Child Language*, *13*(1), 15–29. https://doi.org/10.1017/S0305000900000271

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). 'Primitive intelligence' in the auditory.pdf. *Trends in Neurosciences*, *24*(5), 283–288.

Nakata, T., & Trehub, S. E. (2011). Expressive timing and dynamics in infant-directed and non-infant-directed singing. *Psychomusicology: Music, Mind and Brain*, *21*(1–2), 45–53. https://doi.org/10.1037/h0094003

Narayan, C. R., & McDermott, L. C. (2016). Speech rate and pitch characteristics of infant-directed speech: Longitudinal and cross-linguistic observations. *The Journal of the Acoustical Society of America*, *139*(3), 1272–1281. https://doi.org/10.1121/1.4944634

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 756–766. https://doi.org/10.1037//0096-1523.24.3.756

Neale, D., Georgieva, S., Wass, S., & Leong, V. (2017). *Towards a neuroscientific understanding of play : A neuropsychological coding framework for analysing infant-adult play patterns*. *October*. https://doi.org/10.1101/202648

Newman, R. S. (2005). The cocktail party effect in infants revisited: Listening to one's name in noise. *Developmental Psychology*, *41*(2), 352–362. https://doi.org/10.1037/0012-1649.41.2.352

Newman, R. S., & Jusczyk, P. W. (1996). The cocktail party effect in infants. *Perception and Psychophysics*, *58*(8), 1145–1156. https://doi.org/10.3758/BF03207548

Nguyen, T., Bánki, A., Markova, G., & Hoehl, S. (2020). Studying parent-child interaction with hyperscanning. *Progress in Brain Research*, *254*, 1–24. https://doi.org/10.1016/bs.pbr.2020.05.003

Nguyen, T., Reisner, S., Lueger, A., Wass, S., Hoehl, S., & Markova, G. (2023). *Sing to me, baby: Infants show neural tracking and rhythmic movements to live and dynamic maternal singing*.

Nguyen, T., Schleihauf, H., Kayhan, E., Matthes, D., Vrtička, P., & Hoehl, S. (2020). The effects of interaction quality on neural synchrony during mother-child problem solving. *Cortex*, *124*, 235–249. https://doi.org/10.1016/j.cortex.2019.11.020

Nguyen, T., Zimmer, L., & Hoehl, S. (2023). Your turn, my turn. Neural synchrony in mother-infant proto-conversation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *378*(1875). https://doi.org/10.1098/rstb.2021.0488

Niepel, D., Krishna, B., Siegel, E. R., Draganova, R., Preissl, H., Govindan, R. B., & Eswaran, H. (2020). A pilot study: Auditory steady-state responses (ASSR) can be measured in human fetuses using fetal magnetoencephalography (fMEG). *PLoS ONE*, *15*(7 July), 1–19. https://doi.org/10.1371/journal.pone.0235310

Nishihara, K., Horiuchi, S., Eto, H., & Uchida, S. (2002). The development of infants' circadian rest-activity rhythm and mothers' rhythm. *Physiology and Behavior*, *77*(1), 91–98. https://doi.org/10.1016/S0031-9384(02)00846-6

Nishimura, Y., Kanakogi, Y., & Myowa-Yamakoshi, M. (2016). Infants' emotional states influence maternal behaviors during holding. *Infant Behavior and Development*, *43*, 66–74. https://doi.org/10.1016/j.infbeh.2016.02.003

Notbohm, A., Kurths, J., & Herrmann, C. S. (2016). Modification of brain oscillations via rhythmic light stimulation provides evidence for entrainment but not for superposition of event-related responses. *Frontiers in Human Neuroscience*, *10*(FEB2016). https://doi.org/10.3389/fnhum.2016.00010

Novak, G. P., Kurtzberg, D., Kreuzer, J. A., & Vaughan, H. G. (1989). Cortical responses to speech sounds and their formants in normal infants: maturational sequence and spatiotemporal analysis. *Electroencephalography and Clinical Neurophysiology*, *73*(4), 295–305. https://doi.org/10.1016/0013-4694(89)90108-9

Novembre, G., & Iannetti, G. D. (2018). Tagging the musical beat: Neural entrainment or event-related potentials? *Proceedings of the National Academy of Sciences of the United States of America*, *115*(47), E11002–E11003. https://doi.org/10.1073/pnas.1815311115

Nozaradan, S., Keller, P. E., Rossion, B., & Mouraux, A. (2018). EEG Frequency-Tagging and Input–Output Comparison in Rhythm Perception. *Brain Topography*, *31*(2), 153–160. https://doi.org/10.1007/s10548-017-0605-8

Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *Journal of Neuroscience*, *32*(49), 17572–17581. https://doi.org/10.1523/JNEUROSCI.3203-12.2012

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *Journal of Neuroscience*, *31*(28), 10234–10240. https://doi.org/10.1523/JNEUROSCI.0411-11.2011

Obleser, J., & Kayser, C. (2019). Neural Entrainment and Attentional Selection in the Listening Brain. *Trends in Cognitive Sciences*, *23*(11), 913–926. https://doi.org/10.1016/j.tics.2019.08.004

Olsen, W. O. (1977). Average Speech Levels and Spectra in Various Speaking/Listening Conditions: A Summary of the Pearson, Bennett & Fidell (1977) Report. *American Journal of Audiology*, *7*.

Ortiz Barajas, M. C., Guevara, R., & Gervain, J. (2021). The origins and development of speech envelope tracking during the first months of life. *Developmental Cognitive Neuroscience*, *48*. https://doi.org/10.1016/j.dcn.2021.100915

Ostchega, Y. Porter, K., Hughes, J., Dillon, C., Nwankwo, T. (2011). *Resting Pulse Rate Reference Data for Children, Adolescents , and Adults: United States, 1999-2008* (pp. 1–16).

Park, H., Ince, R. A. A., Schyns, P. G., Thut, G., & Gross, J. (2015). Frontal Top-Down Signals Increase Coupling of Auditory Low-Frequency Oscillations to Continuous Speech in Human Listeners. *Current Biology*, *25*(12), 1649–1653. https://doi.org/10.1016/j.cub.2015.04.049

Paulus, M., Hunnius, S., & Bekkering, H. (2013). *Neurocognitive mechanisms underlying social learning in infancy: infants ' neural processing of the effects of others ' actions*. 774–779. https://doi.org/10.1093/scan/nss065

Pederson, D. R. (1975). The soothing effect of rocking as determined by the direction and frequency of movement. *Canadian Journal of Behavioural Science*, *7*(3), 237–243. https://doi.org/10.1037/h0081911

Pederson, D. R., Champagne, L., & Pederson, L. L. (1969). *Relative Soothing Effects of Vertical and Horizontal Rocking*. 16. https://files.eric.ed.gov/fulltext/ED046504.pdf

Pederson, David, R. (1973). *The Soothing Effets of Vestibular Stimulation acs determined by Frequency and Direction of Rocking*. 15p.; Not available in hard copy due to marginal l.

Peelle, J. E., Gross, J., & Davis, M. H. (2013). Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cerebral Cortex*, *23*(6), 1378–1387. https://doi.org/10.1093/cercor/bhs118

Pegg, J. E., Werker, J. F., & Mcleod, P. J. (1992). *Preference for Infant-Directed Over Speech : Evidence From 7-Week-Old Infants*. 325–345.

Pérez, A., Carreiras, M., & Duñabeitia, J. A. (2017). Brain-To-brain entrainment: EEG interbrain synchronization while speaking and listening. *Scientific Reports*, *7*(1), 1–12. https://doi.org/10.1038/s41598-017-04464-4

Perone, S., & Gartstein, M. A. (2019). Relations between dynamics of parent-infant interactions and baseline EEG functional connectivity. *Infant Behavior and Development*, *57*(March), 101344. https://doi.org/10.1016/j.infbeh.2019.101344

Peter, V., Goswami, U., Burnham, D., & Kalashnikova, M. (2023). Impaired neural entrainment to low frequency amplitude modulations in english-speaking children with dyslexia or dyslexia and DLD. *Brain and Language, 236,* 1–10. https://doi.org/10.1016/j.bandl.2022.105217

Pethe, J., Mühler, R., Siewert, K., & Von Specht, H. (2004). Near-threshold recordings of amplitude modulation following responses (AMFR) in children of different ages. *International Journal of Audiology*, *43*(6), 339–345. https://doi.org/10.1080/14992020400050043

Phillips, E. (n.d.). Using continuous methods of analysis to examine SBT during naturalistic caregier-infant object play. *Prepint*.

Phillips-Silver, J., & Trainor, L. J. (2005). Psychology: Feeling the beat: Movement influences infant rhythm perception. *Science*, *308*(5727), 1430. https://doi.org/10.1126/science.1110922

Piazza, E. A., Hasenfratz, L., Hasson, U., & Lew-Williams, C. (2020). Infant and Adult Brains Are Coupled to the Dynamics of Natural Communication. *Psychological Science*, *31*(1), 6–17. https://doi.org/10.1177/0956797619878698

Picton, T. W., Skinner, C. R., Champagne, S. C., & Kellett, A. J. C. (1987). Potentials evoked by the sinusoidal modulation of the amplitude or frequency of a tone. *Journal of Acousitic Society of America*, *82*(1), 165–178.

Picton, T., Sasha John, M., Dimitrijevic, A., Purcell, D. (2003). Human auditory steady-state responses. *International Journal of Audiology*, *42*, 177–219.

Poeppel, D., & Assaneo, M. F. (2020). Speech rhythms and their neural foundations. *Nature Reviews Neuroscience*, *21*(6), 322–334. https://doi.org/10.1038/s41583-020-0304-4

Poulin-Dubois, Diane; Brosseau-Liard, P. (2017). The Developmental Origins of Selective Social Learning. *Physiology & Behavior*, *176*(3), 139–148. https://doi.org/10.1177/0963721415613962.The

Povel, D., & Essens, P. (1985). Perception of Temporal Patterns. *Music Perception*, *2*(4), 411–440.

Power, A. J., Mead, N., Barnes, L., & Goswami, U. (2012). Neural Entrainment to Rhythmically Presented Auditory, Visual, and Audio-Visual Speech in Children. *Frontiers in Psychology*, *3*(July), 1–13. https://doi.org/10.3389/fpsyg.2012.00216

Provasi, J. (2016). Comment le rythme vient aux Bébés ? *Spirale*, *76*(4), 50–63. https://doi.org/10.3917/spi.076.0050

Provasi, J., Anderson, D. I., & Barbu-Roth, M. (2014). Rhythm perception, production, and synchronization during the perinatal period. *Frontiers in Psychology*, *5*(SEP), 1–16. https://doi.org/10.3389/fpsyg.2014.01048

Provasi, J., Blanc, L., & Carchon, I. (2021). The importance of rhythmic stimulation for preterm infants in the NICU. *Children*, *8*(8), 1–18. https://doi.org/10.3390/children8080660

Rees, A., Green, G. G. R., & Kay, R. H. (1986). Steady-state evoked responses to sinusoidally amplitude-modulated sounds recorded in man. *Hearing Research*, *23*(2), 123–133. https://doi.org/10.1016/0378-5955(86)90009-2

Refinetti, R. (2010). The circadian rhythm of body temperature. *Frontiers in Bioscience*, *15*(2), 564–594. https://doi.org/10.2741/3634

Regan, D. (1989). *Human brain electrophysiology: Evoked potentials and evoked magnetic fields in science and medicine*.

Reid, V. M. (2007). *The directed attention model of infant social cognition*. *May 2014*. https://doi.org/10.1080/17405620601005648

Reid, V. M., Striano, T., & Iacoboni, M. (2011). Developmental Cognitive Neuroscience Neural correlates of dyadic interaction during infancy. *Accident Analysis and Prevention*, *1*(2), 124–130. https://doi.org/10.1016/j.dcn.2011.01.001

Reindl, V., Gerloff, C., Scharke, W., & Konrad, K. (2018). Brain-tobrain synchrony in parent-child dyads and the relationship with emotion regulation revealed by fNIRS-based hyperscanning. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2018.05.060

Reis, H., & Sprecher, S. (2013). Family Routines and Rituals. *Encyclopedia of Human Relationships*, *20*(4), 284–299. https://doi.org/10.4135/9781412958479.n209

Rekow, D., Baudouin, J. Y., Durand, K., & Leleu, A. (2022). Smell what you hardly see: Odors assist visual categorization in the human brain. *NeuroImage*, *255*(January), 119181. https://doi.org/10.1016/j.neuroimage.2022.119181

Repp, B. H. (2006). Rate Limits of Sensorimotor Synchronization. *Advances in Cognitive Psychology, 2*(2-3), 163–181. https://doi.org/10.2478/v10053-008-0053-9

Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: A review of recent research (2006-2012). *Psychonomic Bulletin and Review*, *20*(3), 403–452. https://doi.org/10.3758/s13423-012-0371-2

Retter, T. L., Rossion, B., & Schiltz, C. (2021). Harmonic amplitude summation for frequency-tagging analysis. *Journal of Cognitive Neuroscience*, *33*(11), 2372–2393. https://doi.org/10.1162/jocn_a_01763

Rickards, F. W., Tan, L. E., Cohen, L. T., Wilson, O, J., Drew, J. H., Clark, G. M. (1994). Auditory steady-state evoked potentials in newborns. *British Journal of Audiology1*, *28*, 327–337.

Rimmele, J. M., Morillon, B., Poeppel, D., & Arnal, L. H. (2018). Proactive Sensing of Periodic and Aperiodic Auditory Patterns. In *Trends in Cognitive Sciences* (Vol. 22, Issue 10, pp. 870–882). Elsevier Ltd. https://doi.org/10.1016/j.tics.2018.08.003

Riquelme, R., Kuwada, S., Filipovic, B., Hartung, K., & Leonard, G. (2006). Optimizing the stimuli to evoke the amplitude modulation following response (AMFR) in neonates. *Ear and Hearing*, *27*(2), 104–119. https://doi.org/10.1097/01.aud.0000201857.99240.24

Robertson, S. S., Watamura, S. E., & Wilbourn, M. P. (2012). Attentional dynamics of infant visual foraging. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(28), 11460–11464. https://doi.org/10.1073/pnas.1203482109

Rocha, S., & Mareschal, D. (2017). Getting into the Groove: The Development of Tempo-Flexibility Between 10 and 18 Months of Age. *Infancy*, *22*(4), 540–551. https://doi.org/10.1111/infa.12169

Rock, A. M., Trainor, L. J., & Addison, T. L. (1999). Distinctive messages in infant-directed lullabies and play songs. *Developmental Psychology*, *35*(2), 527–534. https://doi.org/10.1037/0012-1649.35.2.527

Rojas, D. C., Maharajh, K., Teale, P. D., Kleman, M. R., Benkers, T. L., Carlson, J. P., & Reite, M. L. (2006). Development of the 40 Hz steady state auditory evoked magnetic field from ages 5 to 52. *Clinical Neurophysiology*, *117*(1), 110–117. https://doi.org/10.1016/j.clinph.2005.08.032

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. In *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* (Vol. 336, Issue 1278, pp. 367–373). https://doi.org/10.1098/rstb.1992.0070

Rosenbaum, E., & Morett, C. R. (2009). The effect of parents' joint work schedules on infants' behavior over the first two years of life: Evidence from the ECLSB. *Maternal and Child Health Journal*, *13*(6), 732–744. https://doi.org/10.1007/s10995-009-0488-8

Roß, B., Borgmann, C., Draganova, R., Roberts, L. E., & Pantev, C. (2000). A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones. *The Journal of the Acoustical Society of America*, *108*(2), 679–691. https://doi.org/10.1121/1.429600

Saberi, K., & Hickok, G. (2023). Forward entrainment: Psychophysics, neural correlates, and function. *Psychonomic Bulletin and Review*, *30*(3), 803–821. https://doi.org/10.3758/s13423-022-02220-y

Saby, J. N., & Marshall, P. J. (2012). The Utility of EEG Band Power Analysis in the Study of Infancy and Early Childhood The Utility of EEG Band Power Analysis in the Study of Infancy and Early Childhood. *Developmental Neuropsychology*, 37–41. https://doi.org/10.1080/87565641.2011.614663

Saffran, J. R., & Kirkham, N. Z. (2018). Infant Statistical Learning. *Annual Review of Psychology*, *69*, 181–203. https://doi.org/10.1146/annurev-psych-122216-011805

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*(1), 27–52. https://doi.org/10.1016/S0010-0277(98)00075-4

Saffran, J. R., Werker, J. F., & Werner, L. A. (2007). The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language. In *Handbook of Child Psychology* (pp. 58–108). https://doi.org/10.1002/9780470147658.chpsy0202

Sammon, M. P., & Darnall, R. A. (1994). Entrainment of respiration to rocking in premature infants: Coherence analysis. *Journal of Applied Physiology*, *77*(3), 1548–1554. https://doi.org/10.1152/jappl.1994.77.3.1548

Sandler, A., & Coren, A. (1981). Vestibular Stimulation in Early Childhood: A Review. *Journal of Early Intervention*, *3*(1), 48–55. https://doi.org/10.1177/105381518100300108

Sauseng, P., Klimesch, W., Gruber, W. R., Hanslmayr, S., Freunberger, R., & Doppelmayr, M. (2007). Are event-related potential components generated by phase resetting of brain oscillations? A critical discussion. *Neuroscience*, *146*(4), 1435–1444. https://doi.org/10.1016/j.neuroscience.2007.03.014

Savio, G., Cárdenas, J., Pérez Abalo, M. C., González, A., & Valdés, J. (2001). The low and high frequency auditory steady state responses mature at different rates. *Audiology and Neuro-Otology*, *6*(5), 279–287. https://doi.org/10.1159/000046133

Schaworonkow, N., & Voytek, B. (2021). Longitudinal changes in aperiodic and periodic activity in electrophysiological recordings in the first seven months of life. *Developmental Cognitive Neuroscience*, *47*, 100895. https://doi.org/10.1016/j.dcn.2020.100895

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18. https://doi.org/10.1016/j.tins.2008.09.012

Schroeder, C. E., Lakatos, P., Kajikawa, Y., Partan, S., & Puce, A. (2008). Neuronal oscillations and visual amplification of speech. *Trends in Cognitive Sciences*, *12*(3), 106–113. https://doi.org/10.1016/j.tics.2008.01.002

Shafer, V. L., Yu, Y. H., & Wagner, M. (2015). Maturation of cortical auditory evoked potentials (CAEPs) to speech recorded from frontocentral and temporal sites:

Three months to eight years of age. *International Journal of Psychophysiology*, *95*(2), 77–93. https://doi.org/10.1016/j.ijpsycho.2014.08.1390

Shannon, C. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal*, *196*(4), 519–520. https://doi.org/10.1016/s0016-0032(23)90506-5

Singh, N. C., & Theunissen, F. E. (2003). Modulation spectra of natural sounds and ethological theories of auditory processing. *The Journal of the Acoustical Society of America*, *114*(6), 3394–3411. https://doi.org/10.1121/1.1624067

Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Sciences*, *7*(8), 343–348. https://doi.org/10.1016/S1364-6613(03)00156-6

Smith, L. B., Jayaraman, S., Clerkin, E., & Yu, C. (2018). The Developing Infant Creates a Curriculum for Statistical Learning. *Trends in Cognitive Sciences*, *22*(4), 325–336. https://doi.org/10.1016/j.tics.2018.02.004

Smith, N. A., & Trainor, L. J. (2008). Infant-directed speech is modulated by infant feedback. *Infancy*, *13*(4), 410–420. https://doi.org/10.1080/15250000802188719

Stapells, D. R., Galambos, R., Costello, J. A., & Makeig, S. (1988). Inconsistency of auditory middle latency and steady-state responses in infants 1 David R. Stapells, Robert Galambos, Jamie A. Costello and Scott Makeig. *Electroencephalography and Clinical Neurophysiology*, *71*, 289–295.

Stefanics, G., Háden, G., Huotilainen, M., Balázs, L., Sziller, I., Beke, A., Fellman, V., & Winkler, I. (2007). Auditory temporal grouping in newborn infants.

Psychophysiology, *44*(5), 697–702. https://doi.org/10.1111/j.1469-8986.2007.00540.x

Stefanics, G., Hangya, B., Hernádi, I., Winkler, I., Lakatos, P., & Ulbert, I. (2010). Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. *Journal of Neuroscience*, *30*(41), 13578–13585. https://doi.org/10.1523/JNEUROSCI.0703-10.2010

Stenberg, G. (2013). Do 12-month-old infants trust a competent adult? *Infancy*, *18*(5), 873–904. https://doi.org/10.1111/infa.12011

Stern, D. (1977). *The First Relationship*. Harvard University Press.

Stern, D. (1985). *The Interpersonal World of the Infant*. Karnac Books.

Suanda, S. H., Smith, L. B., & Yu, C. (2017). The Multisensory Nature of Verbal Discourse in Parent – Toddler Interactions The Multisensory Nature of Verbal Discourse in Parent – Toddler. *Developmental Neuropsychology*, *41*(5–8), 324–341. https://doi.org/10.1080/87565641.2016.1256403

Sundara, M., & Scutellaro, A. (2011). Rhythmic distance between languages affects the development of speech perception in bilingual infants. *Journal of Phonetics,* *39*(4), 505–513. https://doi.org/10.1016/j.wocn.2010.08.006

Suppanen, E., Huotilainen, M., & Ylinen, S. (2019). Rhythmic structure facilitates learning from auditory input in newborn infants. *Infant Behavior and Development*, *57*(May), 101346. https://doi.org/10.1016/j.infbeh.2019.101346

Sydor, M., Pop, J., Jasińska, A., & Zabłocki, M. (2022). Anthropo-Mechanical Cradles: A Multidisciplinary Review. *International Journal of Environmental Research and Public Health*, *19*(23). https://doi.org/10.3390/ijerph192315759

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *Journal of Neuroscience*, *16*(13), 4240–4249. https://doi.org/10.1523/jneurosci.16-13-04240.1996

Tamis-LeMonda, C. (2006). *Child Psychology: A Handbook of Contemporary Issues* (C. Balter, L., Tamis-LeMOnda (ed.); 2nd ed.). Psychology Press, Taylor & Francis Group.

Tang, J.-S.-Y., & Maidment, J.-A. (1996). Prosodic Aspects of Child-Directed Speech in Cantonese. *Speech, Hearing and Language*, *9*(1989), 257–276.

Tass, P., Rosenblum, M. G., Weule, J., Kurths, J., Pikovsky, A., Volkmann, J., Schnitzler, A., & Freund, H. (1998). Detection of n:m Phase Locking from Noisy Data: Application to Magnetoencephalography. *Physical Review Letters*, *81*(15), 3291–3294. papers3://publication/uuid/EE6C3B36-05D2-46A0-A8D0-6E09C3C2BF78

Tavano, A., Maess, B., Poeppel, D., & Schröger, E. (2022). Neural entrainment via perceptual inferences. *European Journal of Neuroscience*, *55*(11–12), 3277–3287. https://doi.org/10.1111/ejn.15630

Teichert, T., Gnanateja, G. N., Sadagopan, S., & Chandrasekaran, B. (2022). A linear superposition model of envelope and frequency following responses may help identify generators based on latency. *Neurobiology of Language*, *3*(3), 441–468. https://doi.org/10.1162/nol_a_00072

Ten Oever, S., Schroeder, C. E., Poeppel, D., Van Atteveldt, N., Mehta, A. D., Mégevand, P., Groppe, D. M., & Zion-Golumbic, E. (2017). Low-frequency cortical oscillations entrain to subthreshold rhythmic auditory stimuli. *Journal of Neuroscience*, *37*(19), 4903–4912. https://doi.org/10.1523/JNEUROSCI.3658-16.2017

Thaut, M. H. (2013). Entrainment and the motor system. *Music Therapy Perspectives*, *31*(1), 31–34. https://doi.org/10.1093/mtp/31.1.31

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, *7*(1), 53–71. https://doi.org/10.1207/s15327078in0701_5

Thut, G., Miniussi, C., & Gross, J. (2012). The functional importance of rhythmic activity in the brain. *Current Biology*, *22*(16), R658–R663. https://doi.org/10.1016/j.cub.2012.06.061

Thut, G., Schyns, P. G., & Gross, J. (2011). Entrainment of perceptually relevant brain oscillations by non-invasive rhythmic stimulation of the human brain. In *Frontiers in Psychology* (Vol. 2, Issue JUL, pp. 1–10). https://doi.org/10.3389/fpsyg.2011.00170

Tierney A, Kraus N. Neural entrainment to the rhythmic structure of music. J Cogn Neurosci. 2015 Feb;27(2):400-8. doi: 10.1162/jocn_a_00704. PMID: 25170794.

Tort, A. B. L., Komorowski, R., Eichenbaum, H., & Kopell, N. (2010). Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. *Journal of Neurophysiology*, *104*(2), 1195–1210. https://doi.org/10.1152/jn.00106.2010

Trainor, L. J., & Cirelli, L. (2015). Rhythm and interpersonal synchrony in early social development. *Annals of the New York Academy of Sciences*, *1337*(1), 45–52. https://doi.org/10.1111/nyas.12649

Trainor, L. J., Clark, E. D., Huntley, A., & Adams, B. A. (1997). The acoustic basis of preferences for infant-directed singing. *Infant Behavior and Development*, *20*(3), 383–396. https://doi.org/10.1016/S0163-6383(97)90009-6

Trehub, S. E., & Russo, F. A. (2020). Infant-Directed Singing from a Dynamic Multimodal Perspective: Evolutionary Origins, Cross-cultural variation and Relation to Infant-Directed Speech. *In The Routledge Companion to Interdisciplinary Studies in Singing. Volume 1: Development. Taylor & Francis, New York, USA.*

Trehub, S. E., & Thorpe, L. A. (1989). Infants ' Perception of Rhythm : Categorization of Auditory Sequences by Temporal Structure. *Canadian Journal of Psychology*, *43*(2), 217–229.

Tronick, E. Z., & Cohn, J. F. (1989). Infant-mother face-to-face interaction: age and gender differences in coordination and the occurrence of miscoordination. *Child Development*, *60*(1), 85–92. https://doi.org/10.1111/j.1467-8624.1989.tb02698.x

Truby, H. M., & Lind, J. (1961). *Cry Sounds of the Newborn Infant*.

Tsang, C. D., Falk, S., & Hessel, A. (2017). Infants Prefer Infant-Directed Song Over Speech. *Child Development*, *88*(4), 1207–1215. https://doi.org/10.1111/cdev.12647

Tuck, S. J., Monin, P., Duvivier, C., May, T., & Vert, P. (1982). Effect of a rocking bed on apnoea of prematurity. *Archives of Disease in Childhood*, *57*(6), 475–477. https://doi.org/10.1136/adc.57.6.475

Tunçgenç, B., Cohen, E., & Fawcett, C. (2015). Rock With Me: The Role of Movement Synchrony in Infants' Social and Nonsocial Choices. *Child Development*, *86*(3), 976–984. https://doi.org/10.1111/cdev.12354

Uhlhaas, P. J., Roux, F., Rodriguez, E., Rotarska-Jagiela, A., & Singer, W. (2010). Neural synchrony and the development of cortical networks. In *Trends in Cognitive Sciences* (Vol. 14, Issue 2, pp. 72–80). https://doi.org/10.1016/j.tics.2009.12.002

van Bree, S., Sohoglu, E., Davis, M. H., & Zoefel, B. (2021). Sustained neural rhythms reveal endogenous oscillations supporting speech perception. In *PLoS Biology* (Vol. 19, Issue 2). https://doi.org/10.1371/journal.pbio.3001142

Van Ijzendoorn, M. H., & Hubbard, F. O. A. (2000). Are infant crying and maternal responsiveness during the first year related to infant-mother attachment at 15 months? *Attachment and Human Development*, *2*(3), 371–391. https://doi.org/10.1080/14616730010001596

Van Rullen, R. (2016). Perceptual Cycles. *Trends in Cognitive Sciences*, *20*(10), 723–735. https://doi.org/10.1016/j.tics.2016.07.006

Varlet, M., Williams, R., & Keller, P. E. (2020). Effects of pitch and tempo of auditory rhythms on spontaneous movement entrainment and stabilisation. *Psychological Research*, *84*(3), 568–584. https://doi.org/10.1007/s00426-018-1074-8

Vermillet, A. Q., Tølbøll, K., Litsis Mizan, S., C. Skewes, J., & Parsons, C. E. (2022). Crying in the first 12 months of life: A systematic review and meta-analysis of cross-country parent-reported data and modeling of the "cry curve." *Child Development*, *93*(4), 1201–1222. https://doi.org/10.1111/cdev.13760

Vettori, S., Dzhelyova, M., Van der Donck, S., Jacques, C., Steyaert, J., Rossion, B., & Boets, B. (2020). Frequency-Tagging Electroencephalography of Superimposed Social and Non-Social Visual Stimulation Streams Reveals Reduced Saliency of Faces in Autism Spectrum Disorder. *Frontiers in Psychiatry*, *11*(April), 1–12. https://doi.org/10.3389/fpsyt.2020.00332

Vettori, S., Dzhelyova, M., Van der Donck, S., Jacques, C., Steyaert, J., Rossion, B., & Boets, B. (2019). Reduced neural sensitivity to rapid individual face discrimination in autism spectrum disorder. *NeuroImage: Clinical*, *21*(November 2018), 101613. https://doi.org/10.1016/j.nicl.2018.101613

Virtala, P., & Partanen, E. (2018). Can very early music interventions promote at-risk infants' development? *Annals of the New York Academy of Sciences, 1423(1)* 92-101. https://doi.org/10.1111/nyas.13646

Vogelsang, L., Vogelsang, M., Pipa, G., Diamond, S., & Sinha, P. (2024). Butterfly effects in perceptual development: A review of the 'adaptive initial degradation' hypothesis. *Developmental Review*, *71, 101117*. https://doi.org/10.1016/j.dr.2024.101117

Voloh, B., & Womelsdorf, T. (2016). A role of phase-resetting in coordinating large scale neural networks during attention and goal-directed behavior. *Frontiers in Systems Neuroscience*, *10*(MAR), 1–19. https://doi.org/10.3389/fnsys.2016.00018

Von Stein, A., Chiang, C., & König, P. (2000). Top-down processing mediated by interareal synchronization. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(26), 14748–14753. https://doi.org/10.1073/pnas.97.26.14748

Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science*, *10*(2), 159–164. https://doi.org/10.1111/j.1467-7687.2007.00549.x

Vrugt, D. Ter, & Pederson, D. R. (2017). *The Effects of Vertical Rocking Frequencies on the Arousal Level in Two-Month-Old Infants Author (s): Dick Ter Vrugt and David R . Pederson Published by : Wiley on behalf of the Society for Research in Child Development Stable URL : http://www.jstor.org*. *44*(1), 205–209.

Wallaert, N., Moore, B. C. J., & Lorenzi, C. (2016). Comparing the effects of age on amplitude modulation and frequency modulation detection. *The Journal of the Acoustical Society of America*, *139*(6), 3088–3096. https://doi.org/10.1121/1.4953019

Wang, Y., Ding, N., Ahmar, N., Xiang, J., Poeppel, D., Simon, J. (2011). Sensitivity to temporal modulation rate and spectral bandwidth in the human auditory system: MEG evidence. *Journal of Neurophysiology*, *107*, 2033–2041.

Wang, Y., Lu, L., Zou, G., Zheng, L., Qin, L., Zou, Q., & Gao, J. H. (2022). Disrupted neural tracking of sound localization during non-rapid eye movement sleep. *NeuroImage*, *260*(January), 119490. https://doi.org/10.1016/j.neuroimage.2022.119490

Ward, L. M. (2003). Synchronous neural oscillations and cognitive processes. In *Trends in Cognitive Sciences* (Vol. 7, Issue 12, pp. 553–559). https://doi.org/10.1016/j.tics.2003.10.012

Warreyn, P., Ruysschaert, L., Wiersema, J. R., Handl, A., Pattyn, G., & Roeyers, H. (2013). *Infants' mu suppression during the observation of real and mimicked goal-directed actions*. *2*, 173–185. https://doi.org/10.1111/desc.12014

Wass, S. V. (2021). The origins of effortful control: How early development within arousal/regulatory systems influences attentional and affective control. *Developmental Review*, *61*(July), 100978. https://doi.org/10.1016/j.dr.2021.100978

Wass, S. V., Clackson, K., Georgieva, S. D., Brightman, L., Nutbrown, R., & Leong, V. (2018). Infants' visual sustained attention is higher during joint play than solo play: is this due to increased endogenous attention control or exogenous stimulus capture? *Developmental Science*, *21*(6). https://doi.org/10.1111/desc.12667

Wass, S. V., Perapoch Amadó, M., & Ives, J. (2022). Oscillatory entrainment to our early social or physical environment and the emergence of volitional control. *Developmental Cognitive Neuroscience*, *54*(November 2021). https://doi.org/10.1016/j.dcn.2022.101102

Wass, S. V., Smith, C. G., Clackson, K., Gibb, C., Eitzenberger, J., & Mirza, F. U. (2019). Parents Mimic and Influence Their Infant's Autonomic State through Dynamic Affective State Matching. *Current Biology*, *29*(14), 2415-2422.e4. https://doi.org/10.1016/j.cub.2019.06.016

Wass, S. V., Whitehorn, M., Marriott Haresign, I., Phillips, E., & Leong, V. (2020). Interpersonal Neural Entrainment during Early Social Interaction. *Trends in Cognitive Sciences*, *24*(4), 329–342. https://doi.org/10.1016/j.tics.2020.01.006

Webb, A. R., Heller, H. T., Benson, C. B., & Lahav, A. (2015). Mother's voice and heartbeat sounds elicit auditory plasticity in the human brain before full gestation. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(10), 3152–3157. https://doi.org/10.1073/pnas.1414924112

Weinstein, D., Launay, J., Pearce, E., Dunbar, R. I. M., & Stewart, L. (2016). Singing and social bonding: Changes in connectivity and pain threshold as a function of group size. *Evolution and Human Behavior*, *37*(2), 152–158. https://doi.org/10.1016/j.evolhumbehav.2015.10.002

Wen, X., Mo, J., & Ding, M. (2012). NeuroImage Exploring resting-state functional connectivity with total interdependence. *NeuroImage*, *60*(2), 1587–1595. https://doi.org/10.1016/j.neuroimage.2012.01.079

Widmann, A. (2008). *Pop_eegfiltnew EEGLAB plugin*. The Mathworks, Inc.

Wilder, C. N., & Baken, R. J. (1974). Respiratory patterns in infant cry. *Human Communication*, *Winter*, 18–34.

Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(7), 2468–2471. https://doi.org/10.1073/pnas.0809035106

Winkler, I., Kushnerenko, E., Hovárth, J., Čeponienė, R. Fellman, V., Huotilainen, M., Näätänen, R., S. S. (2003). Newborn infants can organize the auditory world. *PNAS*, *100*(20), 11812–11815.

Wolff, P. (1967). The role of biological rhythms in early psychological development. *Bulletin of the Meninger Clinic*, *31*, 197–218.

Wunderlich, J. L., & Cone-Wesson, B. K. (2006). Maturation of CAEP in infants and children: A review. *Hearing Research*, *212*(1–2), 212–223. https://doi.org/10.1016/j.heares.2005.11.008

Xie, W., & Richards, J. E. (2017). The Relation between Infant Covert Orienting, Sustained Attention and Brain Activity. *Brain Topography*, *30*(2), 198–219. https://doi.org/10.1007/s10548-016-0505-3

Xie, W., Mallin, B. M., & Richards, J. E. (2018). Development of infant sustained attention and its relation to EEG oscillations: an EEG and cortical source analysis study. *Developmental Science*, *21*(3). https://doi.org/10.1111/desc.12562

Yoo, H., Bowman, D. A., & Oller, D. K. (2018). The origin of protoconversation: An examination of caregiver responses to cry and speech-like vocalizations. *Frontiers in Psychology*, *9*(AUG). https://doi.org/10.3389/fpsyg.2018.01510

Yu, C., & Smith, L. B. (2017). *Multiple Sensory-Motor Pathways Lead to Coordinated Visual Attention*. *41*, 5–31. https://doi.org/10.1111/cogs.12366

Yu, L., & Myowa, M. (2021). The early development of tempo adjustment and synchronization during joint drumming: A study of 18- to 42-month-old children. *Infancy*, *26*(4), 635–646. https://doi.org/10.1111/infa.12403

Zangl, R., & Mills, D. L. (2007). Increased brain activity to infant-directed speech in 6- and 13-month-old infants. *Infancy*, *11*(1), 31–62. https://doi.org/10.1207/s15327078in1101_2

Zeifman, D. (2004). Acoustic Features of Infant Crying Related to Intended Caregiving Intervention. *Infant and Child Development*, *13*, 111–122. https://doi.org/10.1002/icd

Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(13), 5768–5773. https://doi.org/10.1073/pnas.1000121107

Zeskind, P. (2013). Infant Crying and the synchrony of arousal. In E. Altenmüller, S. Sabine, & E. Zimmermann (Eds.), *Evolution of Emotional Communication: From Sounds in Nonhuman Mammals to Speech and Music in Man* (1st ed., pp. 155–174). Oxford University Press.

Zeskind, P. S., Klein, L., & Marshall, T. R. (1992). Adults' Perceptions of Experimental Modifications of Durations of Pauses and Expiratory Sounds in Infant Crying. *Developmental Psychology*, *28*(6), 1153–1162. https://doi.org/10.1037/0012-1649.28.6.1153

Zettersten, M. (2019). Patterns in Language and Linguistics: New Perspectives on a Ubiquitous Concept. (2019). Germany: De Gruyter.

Zewdie, E., Ciechanski, P., Kuo, H. C., Giuffre, A., Kahl, C., King, R., Cole, L., Godfrey, H., Seeger, T., Swansburg, R., Damji, O., Rajapakse, T., Hodge, J., Nelson, S., Selby, B., Gan, L., Jadavji, Z., Larson, J. R., MacMaster, F., … Kirton, A. (2020). Safety and tolerability of transcranial magnetic and direct current stimulation in children: Prospective single center evidence from 3.5 million stimulations. *Brain Stimulation*, *13*(3), 565–575. https://doi.org/10.1016/j.brs.2019.12.025

Zhang, Y. S., & Ghazanfar, A. A. (2016). Perinatally influenced autonomic system fluctuations drive infant vocal sequences. *Current Biology*, *26*(10), 1249–1260. https://doi.org/10.1016/j.cub.2016.03.023

Zhang, Y. S., & Ghazanfar, A. A. (2020). A Hierarchy of Autonomous Systems for Vocal Production. *Trends in Neurosciences*, *43*(2), 115–126. https://doi.org/10.1016/j.tins.2019.12.006

Zhou, H., Melloni, L., Poeppel, D., & Ding, N. (2016). Interpretations of frequency domain analyses of neural entrainment: Periodicity, fundamental frequency, and harmonics. *Frontiers in Human Neuroscience*, *10*(June), 1–8. https://doi.org/10.3389/fnhum.2016.00274

Zimmerman, E., & Barlow, S. M. (2012). The effects of vestibular stimulation rate and magnitude of acceleration on central pattern generation for chest wall kinematics in preterm infants. *Journal of Perinatology*, *32*(8), 614–620. https://doi.org/10.1038/jp.2011.177

Zoefel, B., & Heil, P. (2013). Detection of near-threshold sounds is independent of eeg phase in common frequency bands. *Frontiers in Psychology*, *4*(MAY), 1–17. https://doi.org/10.3389/fpsyg.2013.00262

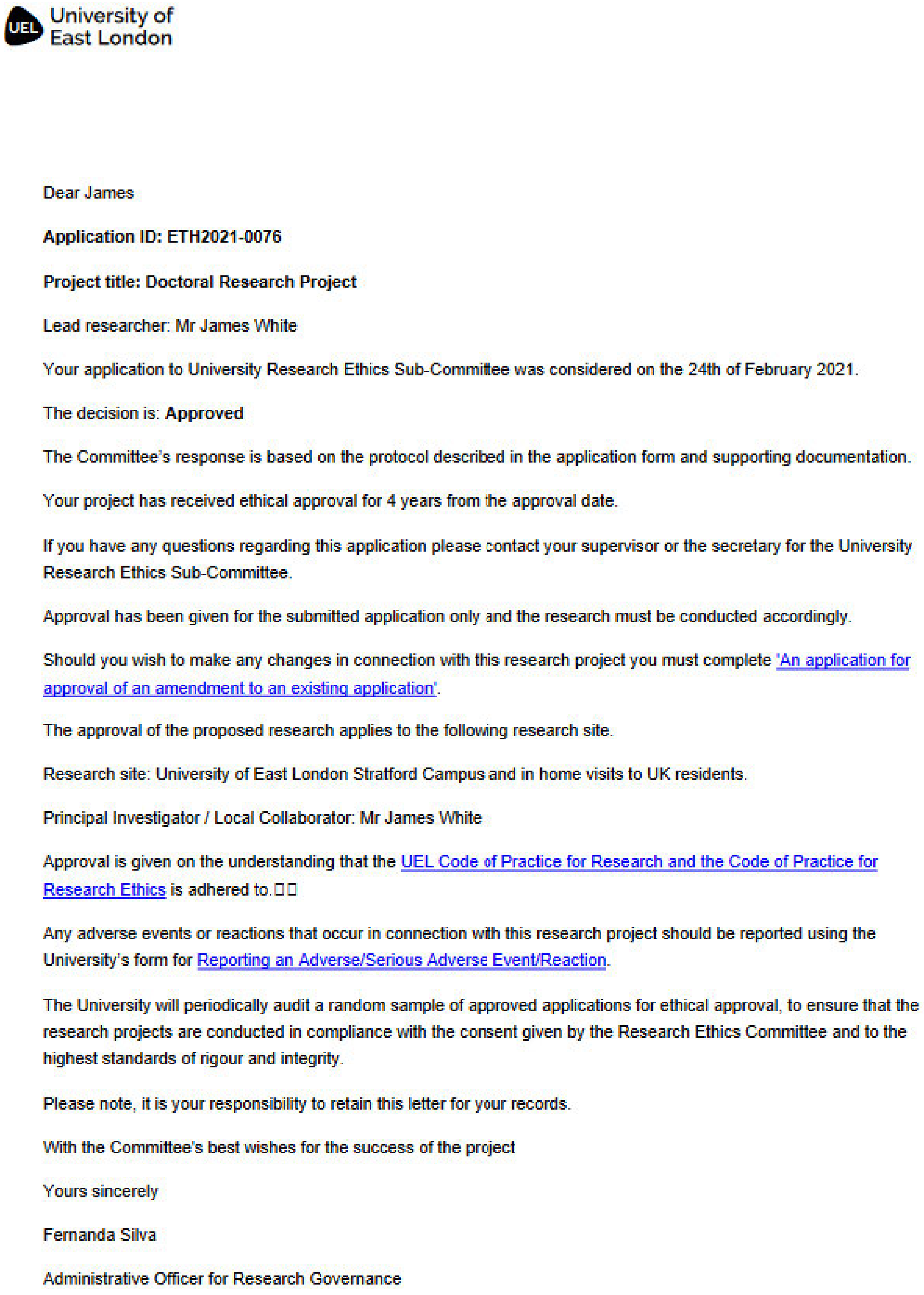
Zoefel, B., & Vanrullen, R. (2015). *The Role of High-Level Processes for Oscillatory Phase Entrainment to Speech Sound*. *9*(December), 1–12. https://doi.org/10.3389/fnhum.2015.00651

Zoefel, B., ten Oever, S., & Sack, A. T. (2018). The involvement of endogenous neural oscillations in the processing of rhythmic input: More than a regular repetition of evoked neural responses. In *Frontiers in Neuroscience* (Vol. 12, Issue MAR, pp. 1–13). https://doi.org/10.3389/fnins.2018.00095

Appendices

Appendix A – Ethics approval and responses

A.1 Ethics approval letter



University of
East London

Dear James

Application ID: ETH2021-0076

Project title: Doctoral Research Project

Lead researcher: Mr James White

Your application to University Research Ethics Sub-Committee was considered on the 24th of February 2021.

The decision is: **Approved**

The Committee's response is based on the protocol described in the application form and supporting documentation.

Your project has received ethical approval for 4 years from the approval date.

If you have any questions regarding this application please contact your supervisor or the secretary for the University Research Ethics Sub-Committee.

Approval has been given for the submitted application only and the research must be conducted accordingly.

Should you wish to make any changes in connection with this research project you must complete 'An application for approval of an amendment to an existing application'.

The approval of the proposed research applies to the following research site.

Research site: University of East London Stratford Campus and in home visits to UK residents.

Principal Investigator / Local Collaborator: Mr James White

Approval is given on the understanding that the UEL Code of Practice for Research and the Code of Practice for Research Ethics is adhered to.□□

Any adverse events or reactions that occur in connection with this research project should be reported using the University's form for Reporting an Adverse/Serious Adverse Event/Reaction.

The University will periodically audit a random sample of approved applications for ethical approval, to ensure that the research projects are conducted in compliance with the consent given by the Research Ethics Committee and to the highest standards of rigour and integrity.

Please note, it is your responsibility to retain this letter for your records.

With the Committee's best wishes for the success of the project

Yours sincerely

Fernanda Silva

Administrative Officer for Research Governance

A.2 Responses to ethics questions

Ethics ETH2021-0076: Mr James White (High risk)

Date 06 Jan 2021

Researcher Mr James White

Student ID 2067263

Project Doctoral Research Project

School Psychology

This is a draft version

Ethics application

Checklist for research projects conducted during a pandemic

Project details

1.1 Is your research project taking place during a pandemic?

Yes

1.2 Where will the research take place?

1. Lab visits - University of East London - Baby Development Lab - Stratford Campus – Arthur Edwards Building

2. Home visits - In participants' homes

1.3 How will the research methodology be amended in response to the pandemic, e.g. change from face-to-face to remote working? How will you address any changing circumstances that may affect the research project and what measures will be put in place?

Ensuring that our study is COVID secure is a top priority and has been a concern that has been well thought through with multiple risk assessments and group discussions.

General:

Our research will be conducted online or remotely as much as possible, including: communication between researchers, coding of collected data, data analysis, write up of experimental results and any other work that could be completed online. Where possible, any questionnaires for participants will be completed online to minimise time spent in the lab and contact with additional resources.

While completing online and remote working, researchers will have no physical contact with each other.

Lab visits:

A specific COVID risk assessment for the lab visits has been attached to this ethics application and details 27 points where changes have been made, reasons for these changes and the risks involved have been scored.

Home visits:

A specific COVID risk assessment for the home visits has been attached to this ethics application, detailing the procedure that we wish to use, the 24 points where changes have been made to ensure that risks have been mitigated and a score for risks involved.

1.4. Is there a plan in place as to how the research project will be conducted in the event of a significant change to the timescale of the research project, changes to the research team, limited access to resources or laboratory space, or restrictions on travel or entering premises?

Yes

1.5. What are the current guidelines/requirements for the location of the research, locally, nationally or/and internationally?

Current Government guidance specifies that labs and research facilities are able to remain open, providing they are COVID secure (Guidance of 6 January 2021). The BabyDev Lab has been cleared by the UEL Gold Committee to continue testing, providing stringent COVID measures are followed; and currently a full research programme is underway. The UK government has published and been updating this guide for working safely during the coronavirus. https://www.gov.uk/guidance/working-safely-during-coronavirus-covid-19/labs-andresearch-facilities This includes guides on lockdown and tier measures, how often to clean and other practices.

International health and safety requirements are not necessary in this scenario as all researchers, participants and research proceedings will be UK based.

1.6. Does your research project comply with all legal and ethical requirements and other applicable guidelines, including those from other organisations and/or countries?

Yes

1.6.1 If yes, please provide details.

As per the government guidance all research facilities can remain open following a COVID-19 risk assessment, and when following COVID safe rules when testing participants. The Baby Development Lab has completed and had a COVID-19 risk assessment approved by a gold level committee. Participants are required to follow the COVID safe rules during testing. No other organisations will be involved in this project.

Research Team

2.1. Who will be responsible for ensuring that the research project adheres to current UK Government guidelines?

Dr Sam Wass (Research Group Leader) and Emily Greenwood (Lab Manager)

2.2. Has consideration been made for supervision and management of the research team members considering remote communication and working from home/off site?

Yes

2.3. Are systems in place to safely share research data with members of the research team that are working remotely?

Yes

Participants

3.1. Are mechanisms in place to ensure that the informed consent process is clear allowing an appropriate forum for questions and right to withdraw, despite resource limitations and in the context of online or remote communication?

Yes

3.1.1 If yes, please provide details.

Participant consent will be collected online via Qualtrics software, all other questionnaire data will be collected online via Qualtrics software. All participants are invited to ask questions via email before, during and after the experiment.

3.2. How will you manage any necessary changes to the participant consenting process should you be unable to recruit participants using your preferred method of data collection?

No changes to participant consent will be needed. Participant consent will be collected online via Qualtrics software, all other questionnaire data will be collected online via Qualtrics software. All participants are invited to ask questions via email before, during and after the experiment.

3.3. Is the research proposal mindful of factors which may impact negatively on equality of participation?

Yes

Risk Assessment

4.1. Have the potential risks to the research or the health, safety and well-being of researchers, research participants and others been fully considered?

Yes

4.1.1 If yes, please provide details.

Please see COVID risk assessments and general risk assessment in the attchment section.

4.2. How will you conduct the research ensuring that you adhere to Government regulations and taking in to account any changes to local and national

requirements, e.g. compliance with any measures that adhere to current infection control guidelines, protective equipment, social distancing?

National guidelines are being monitored:

https://www.gov.uk/guidance/working-safely-during-coronavirus-covid-19/labs-and-research-facilities and as necessary changes are made. Small changes will be made immediately, large changes that require new equipment or furniture will require that testing is paused until the new additions have been procured.

International health and safety requirements are not necessary in this scenario as all researchers, participants and research proceedings will be UK based.

4.3. Have health and safety requirements been evaluated for specific research environments inherent to the project design e.g. laboratory work, access to appropriate equipment, interviews at external organisations or schools?

Yes

4.4. If exposure to the pandemic during the research compromises health and safety, how will you ensure that local, national and where required, international, health and safety requirements are met?

National guideline requirements will be monitored and adhered to. International health and safety requirements are not necessary in this scenario as all researchers, participants and research proceedings will be UK based. https://www.gov.uk/guidance/working-safely-during-coronavirus-covid-19/labs-and-research-facilities

If a researcher or participant is exposed to COVID-19, receives a positive COVID-19 diagnosis and has recently been in the Baby Development Lab these steps will be taken:

1. All researchers and participants who were in the lab on the same day as the exposed participant/researcher, or who have visited the lab since will be informed.

2. An extra thorough deep clean of all rooms, furniture, equipment, toys and other present items will be completed.

3. An exposed researcher will be asked to stay away from the lab for 14 days, and take a COVID test that must be negative before they will be allowed to return. An exposed participant will be asked to stay away from the lab.

4. All other researchers who came into close contact with the person who tested positive must self isolate for at least 5 days before taking a COVID test. Only researchers with negative COVID tests will be allowed to return.

NB. As part of our approved Risk Assessment, researchers and participants are asked to scan the displayed ‘track and trace’ posters at the UEL main reception; and as such will be notified through the app of potential transmissions. Participants who do not have the app will be asked to write down their details in a dated log. The measures go above and beyond Government guidance to ensure the possibility of COVID transmission by an infected researcher or participant is an absolute minimum.

4.5. How will you monitor any changes in national/international Government or equivalent, rules or restrictions and any other health and safety guidance which may affect the research project?

UK National government guidelines for research are being monitored via the government portal:

https://www.gov.uk/guidance/working-safely-during-coronavirus-covid-19/labs-and-research-facilities International health and safety requirements are not necessary in this scenario as all researchers, participants and research proceedings will be UK based.

4.6. Do you have a process in place to maintain compliance with legal and ethical requirements in emergent circumstances, including those from other countries, and staying current as they may change, including consideration of an agreed upon standard if national/regional standards differ?

Yes

4.6.1 If yes, please provide details.

UK National government guidelines for research are being monitored via the government portal:

https://www.gov.uk/guidance/working-safely-during-coronavirus-covid-19/labs-and-research-facilities International health and safety requirements are not necessary in this scenario as all researchers, participants and research proceedings will be UK based.

Data collection and storage

5.1. Have you included sufficient lead time to complete data collection should your research project be suspended or significantly delayed due to unforeseen circumstances?

Yes

Collaborators

6.1. Have agreements been reached or formalised relating to collaborative working, Particularly where the collaborations cross institutional, national, disciplinary or sector boundaries?

Yes

6.2. Is there a plan in place for ensuring that all contributors are given sufficient opportunity to participate in spite of limited resources or difficulties with working during emergent circumstances?

Yes

Contracts

7.1. Has consideration been made for possible modifications to contracts and financial agreements relating to the research, including possible delays in meeting deliverables or changes in research design?

Yes

Intellectual Property and Authorship

8.1. Have agreements related to intellectual property, publication and authorship, or collaboration, including roles and responsibilities of research team members, been agreed and formalised? If so, do agreements need to be renegotiated?

Yes

Conflict of Interest

9.1. Have conflicts of interest relating to the research been identified, declared and addressed in the context of the pandemic, including consideration of pressure to accelerate research projects?

Yes

Monitoring and Audit

10.1. Have monitoring and audit requirements been considered in the context of a pandemic, e.g. secure record keeping, data exchange?

Yes

10.2. Are systems in place to safely share data with those conducting monitoring and auditing?

Yes

Dissemination

11.1. Will the research and its findings be reported accurately, honestly and within a reasonable time frame? Has consideration been given to the best ways to communicate research findings considering any restrictions on certain methods of dissemination due to infection control measures?

Yes

1. Project details

1.1. Title of proposed research project

Oscillatory neural and autonomic correlates of social attunedness during early life: new mechanistic

insights into how we learn to learn from one another

1.2. UEL Researchers

Mr James White

Miss Marta Perapoch Amado

Ms Louise Goupil

Dr Samuel Wass

1.3. Start date of project for which ethical approval is being sought

01 Mar 2021

1.4. Anticipated end date of project for which ethical approval is being sought

31 Aug 2025

1.5. If this project is part of a wider research, please provide the RRDE, SREC, CREB or NHS research ethics approval number.

1.6. If this project is part of a wider research study, please state the start and end dates of the wider study.

01/09/2020 - 31/08/2025

1.7. Where will the research take place?

University of East London Stratford Campus and in home visits to UK residents.

2. Aims and methodology

2.1. Aims and objectives of the project

We are a social species. Most infants, and young children, spend the majority of their early waking lives in the company of others. But, for practical reasons, almost everything that we know about how the brain subserves early attention and learning comes from studies that examined brain function in one individual at a time. This means that we understand lots about how children attend and learn from information presented while they are alone, viewing a computer screen - but little about how attention is shared between people during social interaction. ONACSA will develop new techniques to look, for the first time, at how two brains dynamically interact with one another during early learning exchanges. The project will determine how children's active, participatory bids during learning lead to reactive changes in both members of the dyad – and how these changes, in turn, influence both partners' subsequent attention, and learning. It will also determine how, and why, some infants, and some parents, show greater sensitivity during social exchanges than others. And, using targeted interventions, it will investigate whether social sensitivity can be improved. The question of how two rains dynamically influence one another during learning exchanges has been described as the 'dark matter' of social neuroscience. Yet nobody has looked at these questions before from the perspective of early learning. Our results may help us to move beyond viewing children primarily as passive recipients of information during learning exchanges, to a perspective that better appreciates children's role as active participants in learning. Our findings may also have practical implications for educationalists, and clinicians.

2.2. Methodology, data analysis and recruitment for the project

We aim to collect neuroimaging and physiological data from 150 infants and their parents. The infants will be ages 2-36 months and will have multiple lab and home sessions as they develop at with one of each session at 2, 6, 12, 18 and 36 months. This data will be collected from participants while they engage in screen-based tasks, tabletop free play sessions or naturalistic at home data. I have attached lab and at home COVID risk assessments in the attachments section to demonstrate how all sessions will be kept COVID safe. Questionnaires and interview Once enrolled in the study, participants will be asked to fill in the family background information questionnaire. At each of the testing sessions participants will be asked to complete the following questionnaires in each of the visits: Infant Behaviour Questionnaire, Communicative Development Index, GAD-7, PHQ-9, PSWQ, CHAOS, perceived stress scale. These are all attached in the attachments section

A subset of participants (~30) will be asked to opt-in to the Adult Attachment Interview, a one-hourlong, semi-structured interview. It will be used with a sub-sample of 30 adult participants at the 2-month check. The interview will be conducted once, by one researcher and the adult (parent) participant; either during the lab visit or at home, while the baby is sleeping or otherwise quiet/not needing much adult interactions (so for instance, while feeding).

Lab Sessions

Parent will be seated in a chair across the table from the infant who will be seated in an age appropriate highchair. Both participants will be required to wear specialised head caps which contain 64 electrodes or sensors to record brain activity from the scalp. Additional electrodes will be attached to the participants' face to collect muscle movement data and on the right clavicle and left rib area to record physiological activity, these electrodes will be stuck onto the skin using a small 13 x 8 mm paper washers. Additionally, in accordance with the manufacturing guidelines a small amount of nontoxic water-saline-based gel will be applied to each sensor in order to increase signal quality. Each participant's sensors will then be plugged into individual A/D amplifiers. These amplifiers are electrically isolated units meaning that they only record electrical activity. Both amplifiers will then be connected via a fibre optic cable, and the summed data will be transmitted into an acquisition computer. All cables will be out of reach of infants and where possible held within plastic cable housing. The length of recording time will vary with different ages and will be maintained at an age appropriate level with regular breaks if needed. Participants will also be video recorded, as to allow for additional behavioural analysis post hoc. All video data recorded will be stored on a UEL computer which is not connected to the internet. Additionally, each member of the research team will have an additional copy of the data on an encrypted hard drive. All lab session techniques have been used before in previous studies in labs globally without problems. During initial preparation for the session, infants and parents will be invited into the BabyDevLab, setup in a relaxed format with comfortable chairs and toys. All EEG and physiological apparatus will be applied (as above) to both participants in this room before moving to the testing room. If the participants become distressed, soothing techniques will be encouraged from the parent. If either of the participants becomes too distressed the session will be ended.

During the screen-based task, participants will be asked to watch a screen-based display lasting less than 10 minutes, which is designed to show visual and auditory stimuli that cause neural entrainment. Stimuli include a continuously playing cartoon and audio stimuli with changing visual patterns surrounding the cartoon display and a chirping sound as part of the auditory sound. Participants will be seated in an age appropriate chair, and younger/fussy infants will be asked to sit on their parent's lap in front of the screen. No other objects will be within reach of the participants. Participant response will be passively measured using EEG caps. A researcher will be in the room at all times to monitor the participants for signs of distress. During the screen-based interaction task, participants will be asked to interact with each other via a video link. Participants will be placed in separate rooms so that live audio and visual communication cannot occur without the video link. Infants will be placed in an age appropriate chair positioned in front of a table, on the table may be toys or other age appropriate objects that the infant can play with as well as a screen displaying the parent. The parent will be placed in a chair in front of a table, on the table may be a set of toys or other infant age appropriate objects that the adult can play with as well as a screen displaying the infant. The video screen will be used to manipulate the course of the social interaction. During the tabletop, free play task, participants will be placed on opposite sides of a table in age appropriate chairs. Between the participants may be age appropriate toys or other objects that the pair can play with.

Home Sessions

These sessions take place in the family's home. It involves parent and baby wearing a small wearable device during a day and a night session. This device has been designed to record continuous data for ~10 hours in a variety of parameters such as: heart rate, visual and auditory data, proximity between paired devices and location. All home session techniques have been used before in previous studies without problems. For the set-up, a researcher will travel to the participants' homes at an agreed time and day and fit both participants with our devices. After that, he/she will leave the home. At the end of the session (that is, evening for day-time sessions or the following morning for night-time sessions) the researcher will travel back to their homes to collect the device. The monitor will be worn for periods of the day/night but never during washing. The device is not waterproof, so we ask participants not to go swimming or have wet play whilst wearing it. Other than that, participants can nap/sleep, have her nappy changed, feed and generally do anything they like. During the periods when the monitor is being worn, the parent will be asked to fill in a logbook to record the infant's sleep patterns. The devices will all be encrypted and the raw data will only be accessible to the research team. All adults engaging with infants will have acquired full DBS safety checks prior to any contact with infant.

2.3. Is the data accessed, collected or generated of a sensitive nature?

Yes

2.3.1. If yes, please provide details. Please ensure that all data of a sensitive nature is handled carefully and stored appropriately.

Video and audio recording data - will be personally identifiable, including information of vulnerable (infant) participants. All video data will be stored on UEL's OneDrive for Business. Only one file will be kept in which participant numbers are linked to personally identifiably information (names and addresses). This will be stored on a separate computer and kept in a locked office, on a password-protected computer, in an encrypted file. Data will be backed-up onto OneDrive for Business. Immediately after recording, all video data will be transferred to dedicated, encrypted hard disks that have been purchased specially for this purpose. Recordings will be securely deleted from the device that was used to make the recording. All data collected will only be accessible to members of the research team. It will be stored only on encrypted hard disks, in password-protected format. Between recording sessions, recording devices will be stored in a locked secure location. Questionnaire and interview data - will include questions of a sensitive nature related to demographics, socio-economic status and mental health. All participants will be assigned unique ID numbers. Participants' ID number will be used at all times when managing the research data. All questionnaire data will only be accessible to the research team and will be kept on password protected, encrypted computers and backed up on password protected, encrypted hard drive.

3. About your project

3.1. Is the research project funded?

Yes

3.2. Does the project involve external collaborators?

No

3.3. Does the project involve human participants?

Yes

3.4. Does the project involve non-human animals?

No

3.4.1. If yes, where is the research project taking place?

3.5. Does your project involve access to, or use of, material (including internet use) covered by the Terrorism Act (2006) and / or Counter-Terrorism and Border Security Act (2019) or which could be classified as security sensitive?

No

3.6. Does the project involve secondary research using or analysing an existing data set?

No

3.7. Does the project raise ethical issues that may impact on the natural environment over and above that of normal daily activity?

No

3.8. Does the research involve data collected online or via social media?

No

If yes, please provide details.

3.9. Will the research project take place overseas?

No

3.10. Will the researcher or research team be responsible for the security of all data collected in connection with the proposed research?

Yes

3.11. Does your research project require third-party permission?

No

If yes, please provide details.

3.12. Does your research project involve any circumstances where the professional judgement of you and/or the team is likely to be influenced by personal, institutional, financial or commercial interests?

No

If yes, please provide details.

4. Funding

4.1. Funder(s)

European Research Council (ERC)

4.2. Grant type

Research Council

European Union

If you selected other, please provide further details.

4.3. Value of grant

£ 1355696

4.4. Please upload a letter advising of the award of the grant.

6. Recruitment

6.1. Are the research participants able to give informed consent (in written or verbal form)?

No

6.1.1. If no, is this because they are perceived to lack mental capacity or because they are

vulnerable?

Vulnerable

6.1.2. If the participants are perceived to lack mental capacity, please provide the reason(s).

6.1.3. Further details

6.1.4. If the participants are perceived to be vulnerable, please provide details of the vulnerability.

Research will involve children aged 2-36 months at time of testing. In line with previous studies of this type undertaken by the university, parents will be asked to consent to each test proposed in this study. Written information about each test will be emailed to parents at least a week before their visit to the lab, and they will be encouraged to ask for any clarification they feel they need. All participants are invited to ask questions and voice concerns about our consent and information documents, so that we can respond to or expand on any part of the process that is not clear.

6.1.5. Does the research involve children or young people under the age of 16?

Yes

6.1.6 If yes, are the children or young people able to give informed assent?

No

6.1.7. If no, is this because they are perceived to lack mental capacity or because they are vulnerable?

Vulnerable

6.1.8. If the participants are perceived to lack mental capacity, please provide the reason(s).

6.1.9. Further details

6.1.10. If the participants are perceived to be vulnerable, please provide details of the vulnerability.

The infant participants will be aged between 2 and 36 months at the time of testing. At these ages, the babies will be too young to give informed consent before starting, therefore parents will consent on behalf of their infants/children, as standard in this age cohort. In addition, they will be constantly monitored for signs of distress, and any procedure that is deemed upsetting to either parent, researcher or the baby will be halted. All participants will be informed that they have the right to halt proceedings at any stage. Babies and children will be constantly monitored, and any test will be discontinued should they become distressed or show signs of discomfort.

6.2. How will participants be recruited?

Infants will be recruited in four ways: 1) Infants will be recruited from the existing participant database in the Baby Development Lab at UEL; 2) infants will be recruited from third party infant groups such as local prenatal groups, baby-groups and children's centres. Third party infant groups will be found through social media and search engines and the leadership of these groups will be approached to gain permission to share information about infant research at UEL. No information will be shared with any participants of any infant groups without the permission of the gatekeepers. None of these groups will be NHS run groups (so no NHS ethics application will be required). For this project a live link to recruitment places will be sent to research ethics through Emily Greenwood. In person talks will only be conducted providing that this is legally permitted under COVID-19 regulations and we have gained specific permission for this from the leadership of these groups. 3) during lockdown, social media and word of mouth have proved to be the most fruitful participant recruitment avenues. Infants will be recruited through targeted posts on local Facebook parenting groups, and parents invited to get in contact with the lab. Should the group be private, permission will be sought to enter the group where we will disclose our intentions to highlight the research at the Baby Development Lab at UEL. Parents who contact the lab or who take part in existing studies will be asked to disseminate our information to any other parents/expectant mums they know. 4) infants will be recruited through 'flyering', where leaflets inviting parents to learn more information about this project and other research at the UEL baby development lab will be offered to members of the public in public spaces or by delivering leaflets to houses in the local area. When in public spaces, if relevant, permissions will have been sought from the management of the space. Members of the public will have no obligation to take or act on a leaflet. We will not employ flyering during lockdown due to COVID restrictions. For each of these methods, we will strictly comply with the relevant COVID-19 procedures enforced at the time. Fully informed consent will be obtained before testing commences. All of the parents of babies in this study will have received an information sheet and been given an explanation of the aims of this study before providing their contact details. When parents are initially contacted, they will be given more detailed information about what participation in the study involves via an information sheet, which will be emailed to them. At each interaction, an opportunity will be given to ask any questions or gain clarification. Before any data is collected, the parent/carer will be asked to sign a consent form. All participants are invited to ask questions and voice concerns about our consent and information documents, so that we can respond to or expand on any part of the process that is not clear.

6.4. How many participants are being sought for the project?

150 pairs (one adult, one infant)

6.5. How long will participants be required for the project?

Prior to the testing sessions, parents will be asked to complete a series of questionnaires that should take no more than an hour. A small subset of adult participants (~30) will be asked to complete the Adult Attachment Interview, which should take an hour. Participation in the AAI will be voluntary. Ten testing sessions are planned at infant ages 2, 6, 12, 18 and 36 months. At each age range participants will be invited to the lab for a lab session lasting no longer than a few hours. Both adult and infant participants should be able to complete the entire lab session without needing breaks, however if infants are fussy and need a break then these will be taken flexibly during the session until the infant has calmed down. If the infant remains fussy throughout the break then the testing sessions may be ended early and rescheduled or cancelled. Infants are also welcome to take naps and sessions will be longer to accommodate for this. Participants will also be invited to participate in an at home session. A researcher will travel to the participant's home to provide wearable equipment for the parent and infant, along with instructions on how to put on and remove the wearables. The researcher will also be available to answer any questions that the parents may have, this should take no longer than 45 minutes. The participants will be asked to keep the wearable devices on all day (~8-10 hours) until the researcher returns in the evening to remove and collect the wearable devices. This should take no longer than 20 minutes. Participants will be informed that they are able to stop or suspend recording on either the adult or infant devices at any time, for any reason. However, the devices will be comfortable so that wearing the device for the duration of the day is not an issue. Breaks are not required as part of the at home testing session. The night-time wearable device will also be installed at the same time and any questions may be asked. This will add an extra 20 minutes to the evening visit. The night-time wearable device will be collected in the morning in a separate visit which should take no longer than 20 minutes. As with the daytime at home wearable device, the night-time wearable device has been designed to be comfortable so that it can be worn throughout the night without requiring any breaks. Participants will be informed that they are able to stop or suspend recording at any time for any reason.

6.6. Will the participants be remunerated for their contribution?

Yes

6.6.1. If yes, please specify monetary value of cash or giftcard / vouchers.

Participants will be offered the cost of travel costs to and from the lab for lab sessions.

Participants will also be offered a £10 Love To Shop voucher for each session that is completed.

7. DBS

7.1. Do you require Disclosure Barring Service clearance (DBS) to conduct the research project?

Yes

7.2. Is your DBS clearance valid for the duration of the research project?

Yes

7.2.1. If you have current DBS clearance, please provide your DBS certificate number.

001717370584

8. Medical

8.1. Is your project a clinical trial and / or involves the administration of drugs, substances or agents, placebos or medical devices?

No

8.1.1. If yes, please provide clarification as to why your project does not fall under the Medicines for Human Use (Clinical Trials) Regulations (2004) or Medical Devices Regulations (2002) or any subsequent amendments to the regulations.

8.2. Does your project involve the collecting, testing or storing of human tissue / DNA including organs, plasma, serum, saliva, urine, hair, nails or any other associated material?

No

8.2.1. If yes, please provide clarification as to why your project does not fall under the Human Tissue Act (2004).

9. Risk

9.1. Does the project have the potential to cause physical or psychological harm or offence to participants and / or researchers?

No

9.1.1. If yes, please provide details of the risk or harm explaining how this will be minimised.

9.1.2. Please complete and upload a risk assessment form.

9.2. Does the project involve potential hazards and / or emotional discomfort / distress?

Yes

9.2.1. If yes, provide an outline of support, feedback or debriefing protocol.

Infants sometime express mild distress on the application of the EEG equipment. However, we have received extensive training in how to minimise this. We do not obtain any usable EEG recordings if infants are at all distressed, so we take every possible step to minimise this when it occurs. The process of recording EEG data from infants is in place in numerous other research labs across the world, and the process we follow will be identical to those used in other labs. Some questions within the questionnaires (all attached in the attachments section) or the adult attachment interview (interview questions also attached in the attachments section) may be considered to involve emotional discomfort or distress to some participants. All participants will be informed that they do not need to complete any question or questionnaire that they don't want to. The adult attachment interview will be opt in for a small (~30) subset of participants and all participants will be informed that they can skip any question or stop the interview at any time. After the participant has finished the session they will be given a debriefing form (see attached), in the unlikely event that the infant or adult participant has come to any distress they will be able to contact the principal investigator (Dr. Sam Wass) or the ethics subcommittee, both of which are listed in the debriefing sheet.

9.3. Provide an outline of any measures you have in place in the event of an adverse event or reaction or unexpected outcome, the potential impact on the project and, if applicable, the participants.

We do not expect adverse events arising from the study procedures. Under circumstances of an unexpected adverse event, the participants' health and safety will be the highest priority. If health and safety is in any way compromised, the participant will be withdrawn from the study, with clear reasoning given.

10. Anonymisation

10.1. Will the participants be anonymised at source?

No

10.1.1. If yes, please provide details of how the data will be anonymised.

10.2. Are participants' responses anonymised or are an anonymised sample?

No

10.2.1. If yes, please provide details of how the data will be anonymised.

10.3. Are the samples and data de-identified?

Yes

10.3.1. If yes, please provide details of how the data will be anonymised.

Direct and indirect identifiers will be removed from data and participants will be assigned a participant code. This will be entered along with the date and time of testing on the files containing raw data, and will be used to record all other data collected during pre-post assessments. Participant names and ID numbers will be stored in a separate password protected database. This is so that participants can withdraw their data up to the point at which it is included in the final analysis. Consent forms will be kept securely and separately from the raw data. Only members of the research team will have access to both the raw data and consent forms. Data will be retained in a secure place at the end of the project as, should funding allow, we might wish to follow up the sample over a longer time period. Video recording of the behavioural paradigms will only be viewed by members of the research team.

10.4. Please provide details of data transcription.

All data will be analysed using MATLAB software. Behavioural paradigms will be coded according to their manuals. UEL students may be recruited as research assistants in order to code videos; assistants will be fully briefed on ethics related to the study and will not be allowed access to consent forms, code keys, or any other data that would allow them to identify the participant. Research assistants will only be permitted to code data on university computers (that are disconnected from the internet) and will not be allowed to copy or move video files from the university. Research assistants will be asked to complete a confidentiality agreement before undertaking any work on the project. An ethics amendment form will be completed and submitted to add research assistants to this ethics application before data transcription and coding starts.

10.5. If applicable, will all members of the research team know how the code links the data to the individual participant?

Yes

10.5.1. If no, in the event of a researcher's absence please specify the process should access to the research data be required.

10.6. Will participants be anonymised in publications that arise from the research?

Yes

10.6.1. If no, please provide details.

10.7. Will participants have the option of being identified in the study and dissemination of research findings and / or publication?

No

10.7.1. If yes, please provide details.

16. Data security

16.1. Will the researcher or research team be responsible for the security of all data collected in connection with the proposed research?

Yes

16.1.1. If no, please provide details.

16.2. Will the research data be stored safely on a password protected computer?

Yes

16.2.1. If no, please provide details.

16.3. Will the research data be stored on a UEL data managed device?

Yes

16.3.1. If no, please specify where the electronic data will be stored and how the data will be kept secure.

16.4. Will you keep research data, codes and identifying information in a separate location?

No

16.4.1. If yes, please explain how you will store the research data.

16.5. Will the raw data be shared with individuals outside of the research team?

No

16.5.1. If yes, please specify the names, positions and their relationship to the research.

Name

Position

Relationship to research

16.6. Will participants be audio and/or video recorded?

Yes

16.6.1. If yes, please explain how you will transfer, store and, where relevant, dispose of audio and/or video recordings.

Digital audio-video recordings will be transferred onto the UEL secure computer network that only the research team will have access. Video files may need to be stored externally; in this case, recordings will be stored and transferred on a password-encrypted hard drive with access limited to members of the research team.

16.7. If audio and/or video recordings will be retained, please provide details and state how long the recordings will be kept.

Recording will be kept for up to a year after the completion of the study and data analysis unless further funding can be obtained and the data is used again in further analyses, for which ethics approval of analyses of secondary data will be sought.

16.8. Will you retain hard copies of the data?

No

16.8.1. If yes, please provide details of how the data will be transported safely and, where relevant, undergo secure disposal.

16.9. Will the research data be encrypted and transferred inside of European Economic Area (EEA)?

No

16.9.1. If no, provide details of where the research data will be stored and measures in place

to keep the data secure.

Research data will be kept in the UK only. Research data will be stored on password protected, encrypted researcher laptops, dedicated password protected hard disks and will be backed up to UEL's OneDrive for Business. Only research team members will have access to this data.

16.10. How long will the research data that details personal identifiers be stored?

Any personal identifiers will only be retained until the research has been published and it is no longer possible for participants to request to withdraw their data.

18. Dissemination

18.1. Will the results be disseminated?

Yes

18.1.1. If yes, how will the results of the research be reported and disseminated?

Dissertation / Thesis

Peer reviewed journal

Conference presentation

Written feedback to research participants

Books or chapters

Blogs

18.1.2. If you selected other, please provide further details.

18.1.3. If the results of the research will not be reported and disseminated, please provide a reason.

20. Attachments

Generate your Participant Information Sheet and Consent form using answers provided in your ethics approval application. The Word files generated can be edited. Then upload your final files before submitting your application.

20.1. Upload any additional files to support your application which have not already been uploaded within your application.

Appendix B - Supplementary Materials for: At Which Low Amplitude Modulated Frequency Do Infants Best Entrain? A Frequency Tagging Study

B.1 Participant details

Table B.1. Showing the participant details, blanks indicate where the adult participant did not give an answer.

Participant number	Adult age (years)	Infant age (days)	Infant gender	Infant gestation (weeks)	Other children
1009	38.39	155	M	44	N
1021		156	F		
1028	32.86	153	F	42	N
1029	36.82	185	F	44	N
1032	33.22	177	F	44	N
1039	33.73	150	M	40	N
1040	35.61	144	M		Y
1042	35.39	153	M	37.5	Y
1046	40.18	184	F	42	Y
1051	32.84	161	M	44	N
1052	32.37	172	F	47	N
1053	34.88	155	M	39	Y
1054	31.21	159	M	38.5	Y
1055	34.55	161	M	39	Y
1059		178	M		
1060		174	F		
1061	42.45	178	F		Y
1062		172	F		
1063	36.21	151	M	46	N
1064	31.71	182	M	41	Y
1065	40.11	190	M	43	Y
1066	34.58	167	M	39	N
1068	41.07	151	M	39	Y
1069	39.28	193	M	40.29	N
1070		166	F		
1071	31.54	159	F	42	N
1073	38.19	163	F	47	Y
1076	35.98	190	F		Y
1077	35.72	166	M	39	N

1078	34.35	181	M	38	N
1079	33.65	158	F	39.71	N
1082	29.81	194	M		N
1083	34.80	160	M	41	N
1084	39.54	173	M	42.14	N
1085	36.88	151	M		Y
1086	43.10	144	M	43	Y
1087	32.76	136	M	37.29	N
1088	39.24	153	M	37.86	Y
1089	28.64	196	F	40	
1091		187	M		
1092		183	M		
1093	33.47	149	F	42	N
1096	40.03	145	M		N
1097	34.49	160	M	38	N
1099	38.08	166	F	41.29	
1102		166	M		
1103	36.90	169	M	41	N
1104	34.74	153	F	38	Y
1108	38.26	186	F	39.57	Y
1109	37.23	137	F		N
1114	37.14	154	M	43	Y
1117		172	F	37	Y
1128	36.41	144	M	47	N
1129	32.47	210	M	41	
1135	35.99	218	F	42	N
1142	37.27	141	M		Y

B.2 Conditions completed (post preprocessing)

Table B.2. Showing which participants completed which conditions.

Participants	Rest		2Hz		4Hz		6Hz		8Hz		10Hz		12Hz	
	Inf	Mum	Inf	Mum	Inf	Mum	Inf	Mum	Inf	Mum	Inf	Mum	Inf	Mum
1009					X	X			X	X				
1021									X	X				
1028									X	X				
1029										X				
1032									X	X				
1039									X	X				
1040									X	X				
1042			X	X	X	X	X	X	X	X	X	X	X	X
1046									X	X				
1051			X	X	X	X	X	X	X	X	X		X	X
1052			X	X	X	X	X	X	X	X	X		X	X
1053			X	X	X	X	X	X	X	X	X	X	X	X

1053			X	X	X	X	X	X	X	X	X	X	X	X
1054			X	X	X	X		X	X	X		X		
1055				X		X	X	X	X	X			X	X
1059				X		X		X		X		X		X
1060							X	X	X	X			X	X
1061			X	X	X	X	X	X	X	X	X	X	X	X
1062			X	X	X	X	X	X	X	X	X	X	X	X
1063				X		X		X	X	X		X		X
1064	X	X								X				
1065		X		X		X		X		X		X		X
1066	X	X							X	X				
1068	X	X	X	X	X	X		X	X	X	X	X		X
1069	X	X			X	X	X	X	X	X	X	X	X	X
1070	X	X			X	X	X	X	X	X	X	X	X	X
1073	X	X	X		X		X		X	X	X	X	X	X
1076	X	X					X	X	X	X				
1077	X	X					X	X	X	X				
1078	X	X		X	X	X	X	X	X	X	X	X	X	X
1079	X		X		X		X				X	X	X	
1082	X		X		X		X						X	
1083	X	X	X	X	X	X		X		X		X	X	X
1084	X	X	X	X	X	X	X	X					X	X
1085	X	X	X	X	X	X	X	X	X	X	X	X	X	X
1086	X	X	X	X	X	X	X	X	X	X	X	X	X	X
1087	X	X	X	X	X	X	X	X	X	X		X	X	X
1088					X	X	X	X	X	X	X	X	X	X
1089	X	X	X	X	X	X	X	X	X	X		X	X	X
1091	X	X	X	X			X	X					X	X
1092	X	X	X	X	X	X	X	X			X	X	X	X
1093	X	X	X	X	X	X	X	X		X	X	X	X	X
1096	X	X	X	X	X	X		X		X		X		X
1097	X	X	X	X		X		X	X	X		X		X
1099	X	X	X	X	X	X	X	X	X	X	X	X	X	X
1102	X	X	X	X	X	X	X	X	X	X	X	X	X	X
1103	X	X	X	X	X	X	X	X	X	X	X	X	X	X
1104	X	X	X	X	X	X	X	X	X	X	X	X	X	X
1108	X		X		X		X		X		X		X	
1109	X	X	X	X		X		X	X	X		X		X
1114	X	X	X	X	X	X	X	X	X	X	X	X	X	X
1117		X		X	X	X		X						
1128	X	X	X	X	X	X	X	X	X	X	X	X		
1129	X	X	X	X	X	X	X	X	X	X	X	X	X	X
1135		X								X				
1142	X	X	X	X	X	X	X	X	X	X	X	X	X	X

B.3 Data quality metrics per participant and condition

Table B.3. showing the data quality metrics for infant and adult participants by condition. Blue shaded columns represent the percentage of channels identified as noisy or bridged. Any condition with 1-25% channels identified as noisy or bridged had these channels interpolated (see section 2.5), any condition with 25-75% of channels identified as noisy or bridged had these channels rejected and removed from the dataset. White shaded columns show the percentage of one second segments zeroed out (see section 2.5). All values shown are percentages, any blank conditions indicate that either the participant didn't complete this condition or the data were rejected. At the end of each participant section there are rows showing the N for each condition, the average percentage of the number of channels rejected and the percentage of 1 second segments that were zeroed out, the average number of channels rejected above and below the 25% threshold for interpolation, the total number of channels in each condition across all participants and the total number of seconds of EEG in each condition.

Infant participants														
	Rest		**2Hz**		**4Hz**		**6Hz**		**8Hz**		**10Hz**		**12Hz**	
1009					48.00	13.96			23.00	6.70				
1021									16.00	0.00				
1028									25.00	4.66				
1032									13.00	0.00				
1039									2.00	1.65				
1040									0.00	0.00				
1042			3.00	0.00	9.00	0.00	6.00	0.00	5.00	0.90	6.00	0.00	5.00	0.42
1046									0.00	0.00				
1051			8.00	18.69	8.00	3.00	6.00	13.13	8.00	29.85	5.00	2.10	8.00	15.53
1052			11.00	0.00	11.00	1.75	11.00	0.00	9.00	0.00	11.00	1.06	9.00	0.00
1053			25.00	0.00	23.00	0.70	17.00	0.00	36.00	1.96	20.00	0.56	22.00	0.00
1054			47.00	0.00	48.00	0.56			52.00	1.56				
1055							45.00	0.00	45.00	0.00			45.00	0.00
1060							5.00	0.35	3.00	0.00			19.00	5.83
1061			28.00	0.00	17.00	0.00	31.00	0.00	38.00	0.00	17.00	0.00	30.00	0.00
1062			25.00	0.00	11.00	0.00	34.00	10.21	31.00	0.00	13.00	0.00	13.00	0.70
1063									30.00	0.00				
1064	52.00	34.54												
1066									50.00	30.63				
1068	47.00	8.25	41.00	0.91	52.00	0.00			47.00	2.33	39.00	0.56		
1069	47.00	0.00			67.00	0.00	67.00	0.21	42.00	0.00	67.00	0.00	59.00	0.00
1070	17.00	0.35			41.00	2.95	20.00	0.00	20.00	0.00	22.00	0.00	22.00	23.19
1073	23.00	0.00	22.00	0.00	25.00	0.00	23.00	0.00	41.00	0.00	23.00	0.70	27.00	4.93
1076	20.00	0.00					19.00	0.00	19.00	0.00				
1077	9.00	25.62					27.00	35.85	17.00	33.33				
1078	14.00	0.63			16.00	0.00	20.00	0.70	17.00	0.00	16.00	0.00	17.00	0.00
1079	6.00	0.00	6.00	0.00	6.00	12.08	5.00	13.99			9.00	11.44	3.00	0.00
1082	9.00	17.76	11.00	3.83	8.00	0.00	6.00	0.00					9.00	0.00
1083	27.00	0.00	23.00	0.00	42.00	38.91							55.00	44.47
1084	13.00	0.00	11.00	0.00	11.00	0.21	11.00	3.24					11.00	10.03

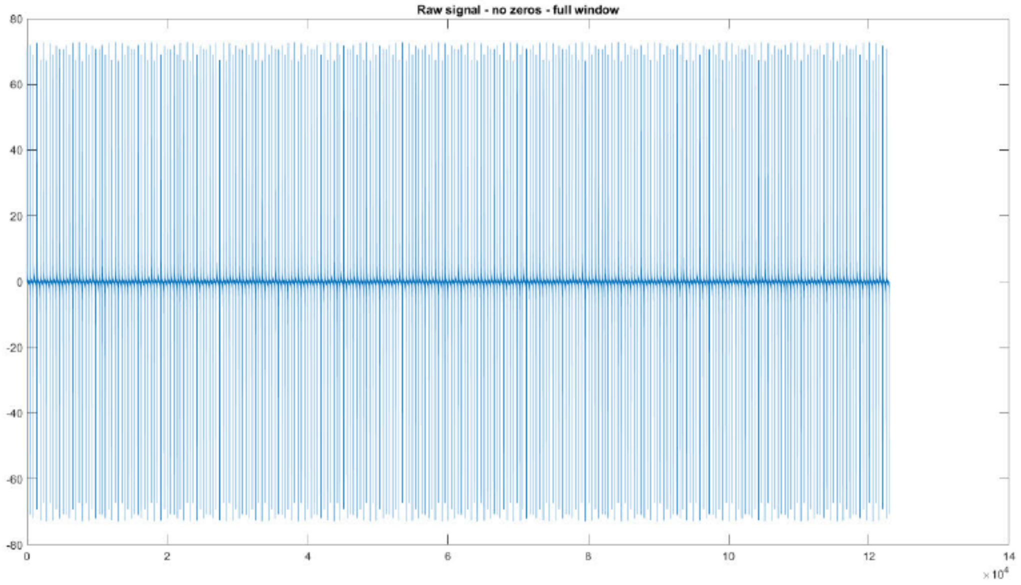
1085	19.00	0.00	17.00	0.00	17.00	0.00	19.00	0.00	28.00	0.00	17.00	1.41	16.00	0.00
1086	0.00	2.73	0.00	1.87	0.00	3.99	2.00	4.77	3.00	3.01	2.00	0.98	5.00	0.00
1087	33.00	0.00	58.00	39.82	45.00	0.00	34.00	0.00	22.00	29.15			33.00	20.62
1088					20.00	0.00	22.00	0.00	22.00	1.97	20.00	0.00	20.00	0.00
1089	48.00	0.00	45.00	0.00	50.00	0.00	45.00	0.28	42.00	0.77	19.00	0.63	52.00	0.28
1091	0.00	0.00	11.00	3.44			11.00	0.49					13.00	16.44
1092	20.00	2.11	16.00	0.00	14.00	0.28	14.00	0.70					25.00	0.00
1093	14.00	0.00	17.00	0.00	17.00	0.00	14.00	0.00			23.00	0.21	20.00	1.55
1096	34.00	2.04	19.00	1.18	28.00	0.56								
1097	39.00	0.00	45.00	9.29					23.00	7.25				
1099	13.00	2.17	8.00	0.00	11.00	1.82	11.00	0.00	14.00	4.42	11.00	2.10	8.00	0.00
1102	16.00	0.00	19.00	0.00	14.00	0.00	16.00	0.00	66.00	50.57	39.00	55.11	17.00	0.00
1103	6.00	0.00	13.00	0.00	2.00	0.00	2.00	0.00	6.00	0.00	5.00	0.00	2.00	0.00
1104	52.00	0.91	42.00	0.74	41.00	3.01	36.00	7.09	52.00	15.55	47.00	19.78	38.00	3.23
1108	33.00	7.84	39.00	4.31	25.00	4.00	13.00	0.35	45.00	0.35	39.00	0.28	39.00	0.56
1109	39.00	0.00	45.00	9.29					23.00	7.25				
1117					70.00	5.39								
1128	23.00	1.41	36.00	1.33	36.00	0.00	39.00	0.00	45.00	9.10	34.00	0.28		
1129	6.00	4.66	2.00	0.98	0.00	0.00	0.00	0.00	2.00	0.00	0.00	2.81	16.00	5.89
N	29		30		33		32		39		24		30	
Avg %	23.41	3.83	23.10	3.19	25.24	2.82	19.72	2.86	25.18	6.23	21.00	4.17	21.93	5.12
Avg % <25	12.67		13.35		12.62		11.87		12.70		13.28		13.33	
Avg % > 25	41.00		39.67		44.14		39.78		42.06		44.17		40.30	
Total elecs	1421		1477		1578		1644		1868		1213		1501	
Total segs		6693		6972		7696		7458		8776		5520		6831
Adult participants														
	Rest		**2Hz**		**4Hz**		**6Hz**		**8Hz**		**10Hz**		**12Hz**	
1009					19.00	7.03			19.00	14.54				
1021									16.00	5.80				
1028									3.00	37.91				
1029									20.00	37.65				
1032									5.00	5.47				
1039									19.00	24.69				
1040									6.00	5.21				
1042			0.00	2.39	0.00	5.26	0.00	0.00	0.00	2.35	0.00	6.52	0.00	1.20
1046									11.00	22.18				
1051			9.00	7.24	47.00	45.68	2.00	0.00	5.00	0.70			2.00	0.00
1052			56.00	12.09	47.00	8.39	42.00	2.50	47.00	5.89			44.00	1.61
1053			9.00	11.41	8.00	19.46	8.00	16.64	9.00	11.66	8.00	29.89	8.00	17.15
1054			19.00	61.04	17.00	52.69	14.00	61.30	13.00	52.53	38.00	52.98		
1055			5.00	1.55	3.00	0.00	6.00	0.56	6.00	3.83			9.00	0.00
1059			6.00	17.77	6.00	20.72	11.00	14.89	13.00	14.67	8.00	6.91	8.00	8.31
1060							9.00	12.04	8.00	19.64			9.00	13.36

1061			17.00	6.67	19.00	4.51	17.00	6.75	19.00	21.04	17.00	1.89	17.00	1.46
1062			31.00	4.76	31.00	9.60	31.00	16.12	33.00	12.35	31.00	9.46	31.00	7.01
1063			23.00	5.81	27.00	10.64	30.00	15.49	23.00	14.22	23.00	8.34	20.00	3.00
1064	11.00	13.78							13.00	21.90				
1065	17.00	1.05	47.00	0.00	44.00	0.56	45.00	3.98	17.00	3.82	45.00	0.00	45.00	1.39
1066	6.00	2.95							6.00	11.97				
1068	5.00	26.59	2.00	24.79	3.00	14.29	2.00	16.34	3.00	18.38	2.00	22.52	2.00	14.42
1069	11.00	3.73			11.00	3.16	11.00	1.75	11.00	0.70	11.00	4.69	11.00	2.11
1070	17.00	12.96			19.00	13.38	19.00	23.20	16.00	9.55	17.00	17.34	19.00	14.89
1073	8.00	0.00							8.00	0.00	8.00	5.64	11.00	51.34
1076	9.00	36.73					8.00	48.61	6.00	51.04				
1077	14.00	23.95					13.00	11.07	13.00	13.62				
1078	25.00	29.11	30.00	1.62	30.00	1.62	31.00	0.00	28.00	14.79	30.00	5.77	30.00	0.00
1079											42.00	24.20		
1083	38.00	37.31	23.00	23.59	16.00	28.28	33.00	30.19	16.00	18.36	25.00	36.84	31.00	38.40
1084	11.00	8.39	27.00	15.45	8.00	1.46	8.00	4.85					8.00	2.45
1085	13.00	16.32	8.00	20.80	5.00	22.61	6.00	17.97	11.00	8.60	13.00	4.97	8.00	19.24
1086	38.00	39.32	38.00	36.47	41.00	20.38	39.00	24.23	39.00	28.64	39.00	37.72	39.00	41.89
1087	2.00	3.91	2.00	9.76	3.00	8.46	6.00	4.82	2.00	11.96	2.00	12.71	2.00	6.30
1088					5.00	9.36	3.00	15.71	6.00	15.58	5.00	0.98	2.00	2.31
1089	22.00	8.90	22.00	0.94	22.00	0.35	22.00	8.46	22.00	0.63	22.00	6.03	22.00	0.42
1091	2.00	0.00	0.00	0.00			2.00	0.00					2.00	10.93
1092	13.00	3.91	14.00	3.24	13.00	2.93	14.00	5.86			13.00	7.68	13.00	9.71
1093	6.00	11.07	6.00	8.98	6.00	15.32	9.00	8.19	11.00	10.43	5.00	11.19	8.00	10.64
1096	42.00	26.62	42.00	6.78	42.00	22.79	42.00	23.54	44.00	22.14	45.00	23.41	44.00	20.76
1097	19.00	2.31	19.00	5.64	19.00	4.25	22.00	4.07	22.00	2.24	22.00	2.94	22.00	0.70
1099	27.00	7.57	27.00	19.12	20.00	1.47	20.00	0.00	22.00	5.05	25.00	1.68	22.00	0.21
1102	44.00	45.28	58.00	35.62	44.00	41.68	39.00	38.72	53.00	38.25	44.00	46.95	39.00	32.77
1103	27.00	13.11	27.00	1.85	25.00	1.69	23.00	0.70	22.00	0.00	25.00	0.49	23.00	0.91
1104	17.00	49.23	14.00	45.54	16.00	43.75	11.00	52.22	17.00	42.27	19.00	38.73	13.00	48.53
1109	19.00	2.31	19.00	5.64	19.00	4.25	22.00	4.07	22.00	2.24	22.00	2.94	22.00	0.70
1114	8.00	15.35	6.00	10.72	8.00	11.75	8.00	15.27	8.00	20.47	11.00	14.24	8.00	13.29
1117	9.00	36.02	9.00	37.28	9.00	45.94	9.00	41.55						
1128	56.00	2.12	61.00	5.28	59.00	0.98	63.00	2.80	63.00	0.77	61.00	1.75		
1129	42.00	1.54	41.00	0.56	42.00	2.44	41.00	0.84	42.00	1.26	42.00	1.40	41.00	3.65
1135	45.00	10.30							45.00	9.99				
N	31		33		36		39		47		32		35	
Avg %	20.10	15.86	21.73	13.65	20.92	14.09	19.00	14.24	18.36	14.83	22.50	14.03	18.14	11.46
Avg % <25	12.00		11.05		11.96		10.89		12.34		13.77		11.19	
Avg % > 25	38.40		40.42		39.92		39.64		43.78		37.85		38.22	
Total elecs	1587		1654		1824		2022		2458		1589		1836	
Total segs		6260		6840		7422		8025		9608		6602		7435

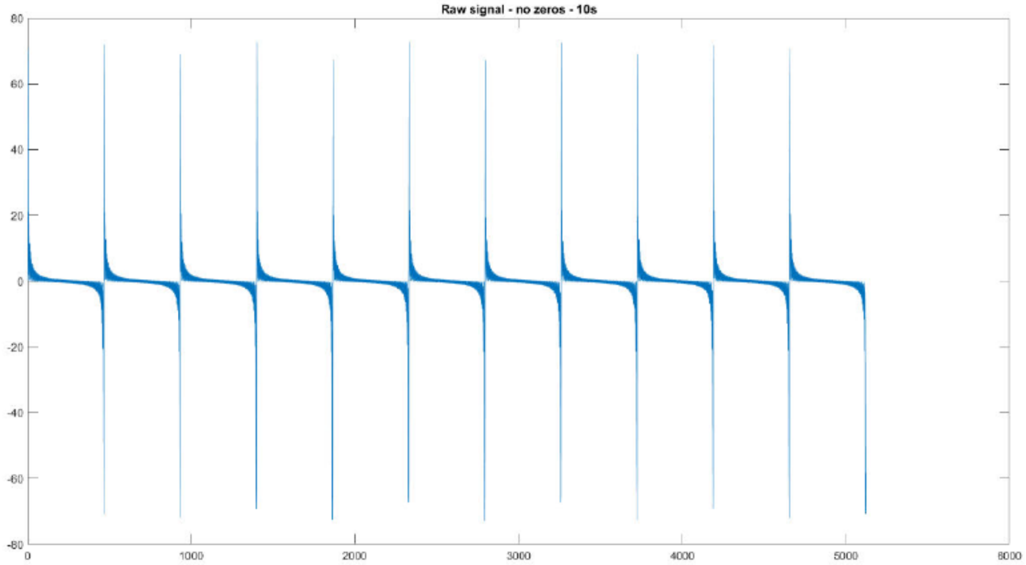
B.4 Check for impact of zero padding the results

In MATLAB, a composite sinewave was produced by adding a series of sinewaves with amplitude modulated frequencies of 1-100 + 10%, i.e. 1.1, 2.2, 3.3 through to 110, the time series can be seen in the below figure (a), along with a close up (b) and the spectral series (c). Then for each one second segment a random number of segments was zeroes out (up to 25% which is the maximum threshold in this study), which can be seen in the spectral (d) and temporal (e) series. While there was a reduction in amplitude of the recorded frequencies between c and e there was no distortion of the spectral make up of the signal. There was some added low level noise broadband across the spectra.

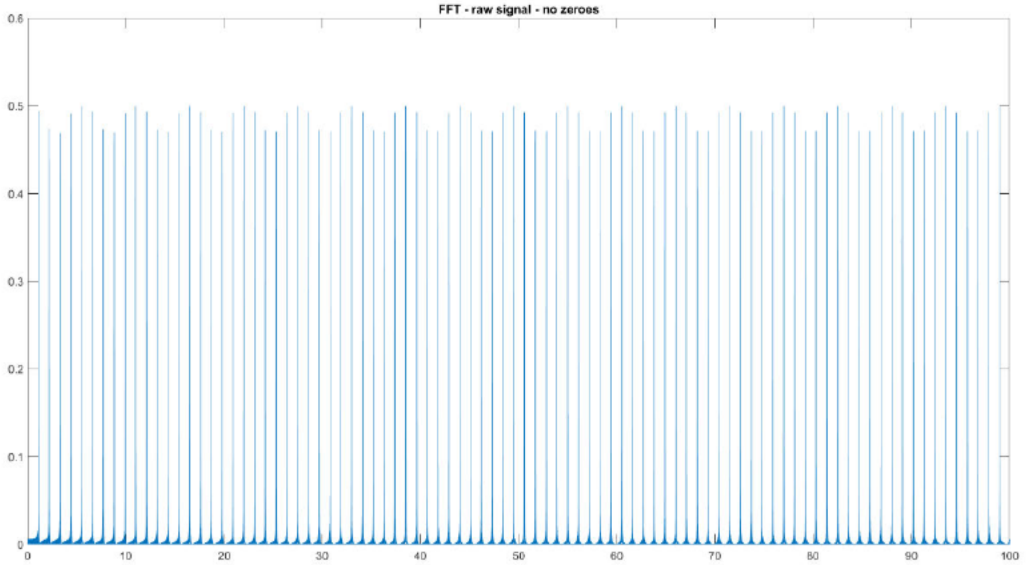
a



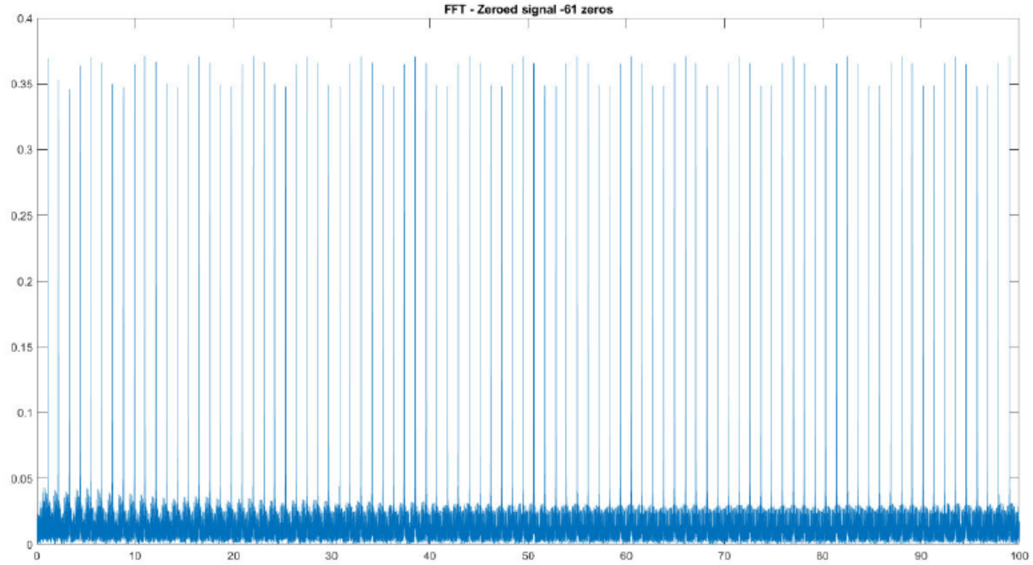
b


Raw signal - no zeros - 10s
80
60
40
20
0
-20
-40
-60
-80
0
1000
2000
3000
4000
5000
6000

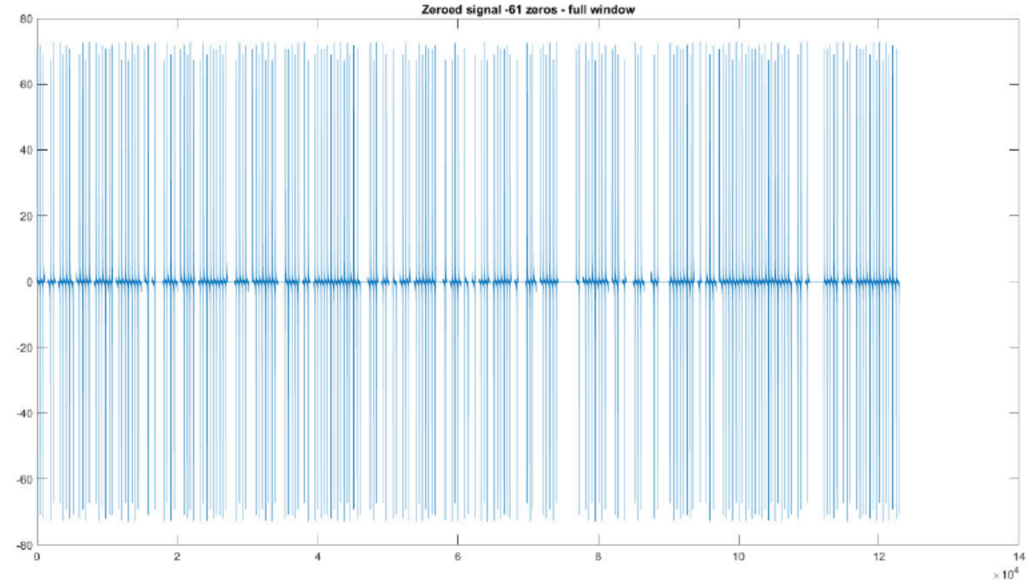
c


FFT - raw signal - no zeroes
0.6
0.5
0.4
0.3
0.2
0.1
0
0
10
20
30
40
50
60
70
80
90
100

d

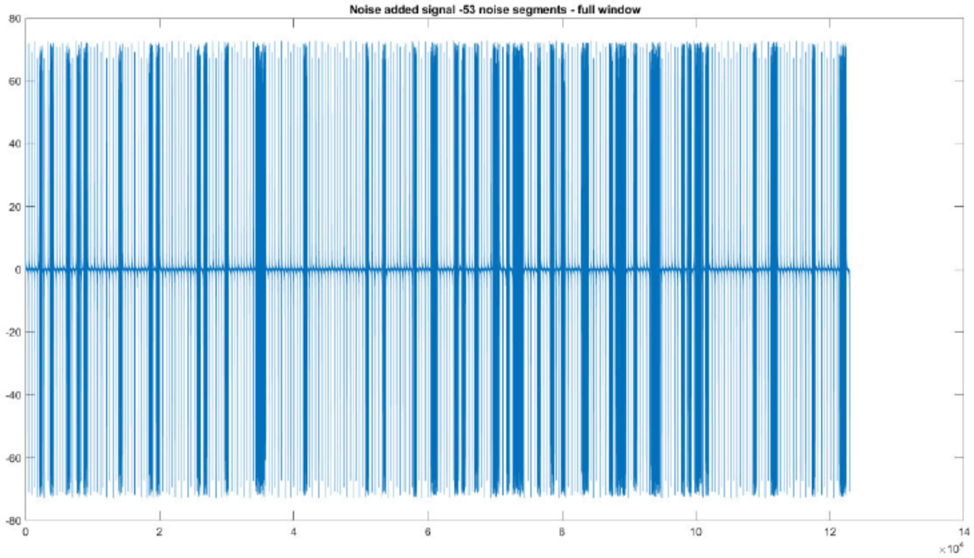


e

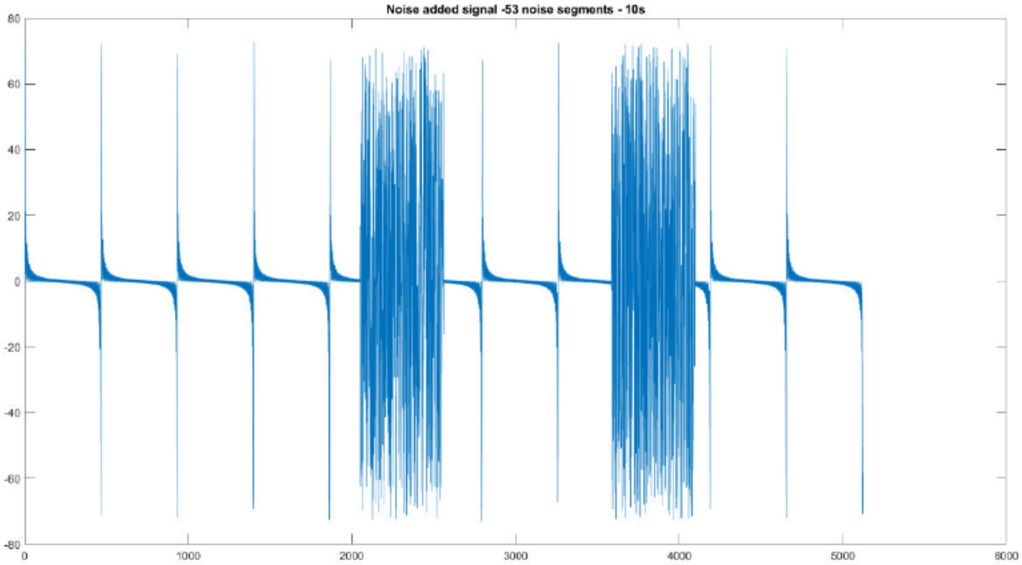


To ensure that this was better than the alternative, white noise was generated which was at least 70% of the amplitude of the sinusoidal waves and randomly added into the original (a) timeseries with up to 25% of the one second segments being replaced. Timeseries with added noise is shown in (f), with a close up (g) and a spectral decomposition (h). There is a lot of distortion to the original signal (a), along with a lot of noise added broadband across the spectra, which suggests that zeroing out noisy segments is worth doing.

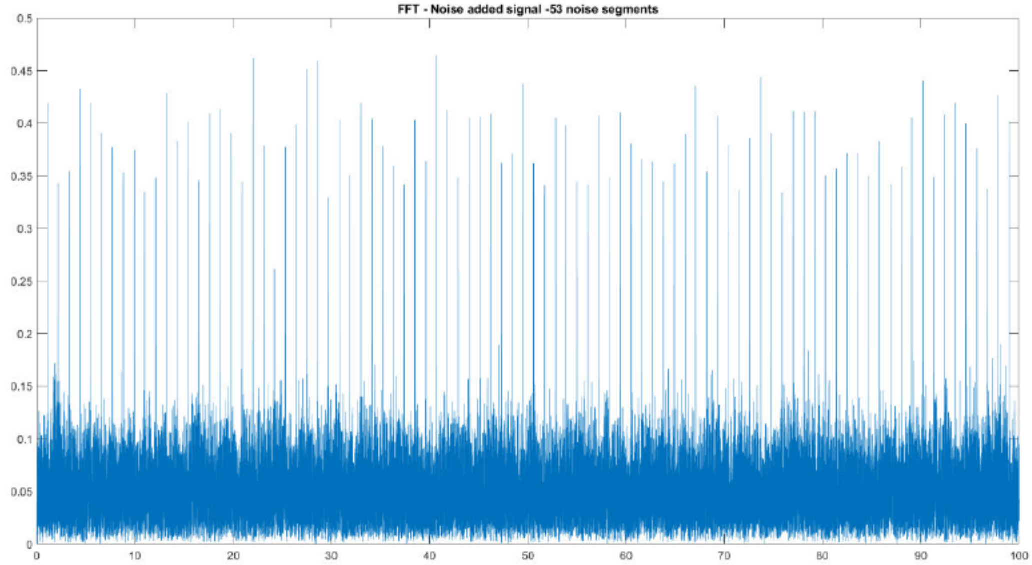
f


Noise added signal -53 noise segments - full window
80
60
40
20
0
-20
-40
-60
-80
0
2
4
6
8
10
12
14
$\times 10^4$

g


Noise added signal -53 noise segments - 10s
80
60
40
20
0
-20
-40
-60
-80
0
1000
2000
3000
4000
5000
6000

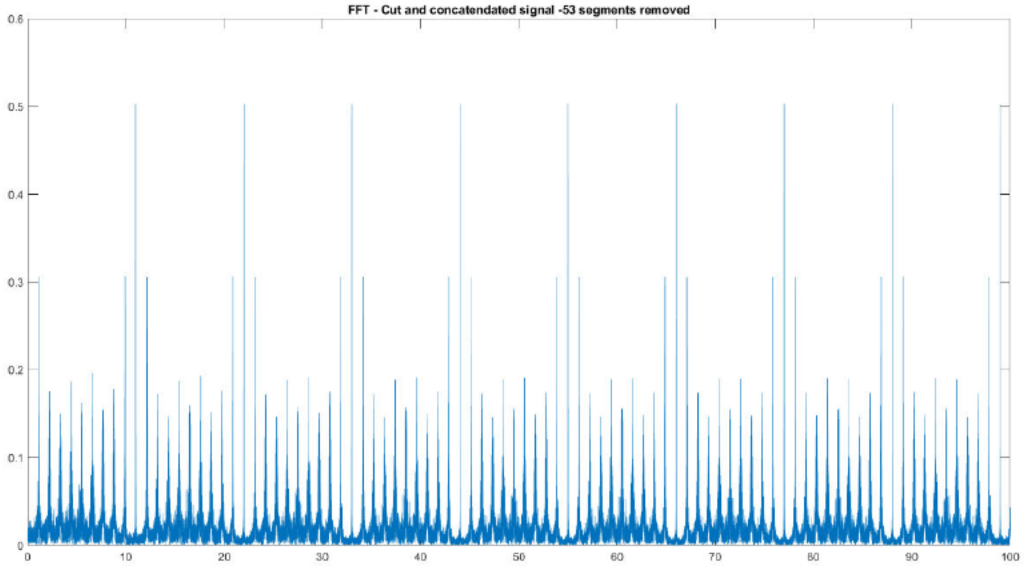
h



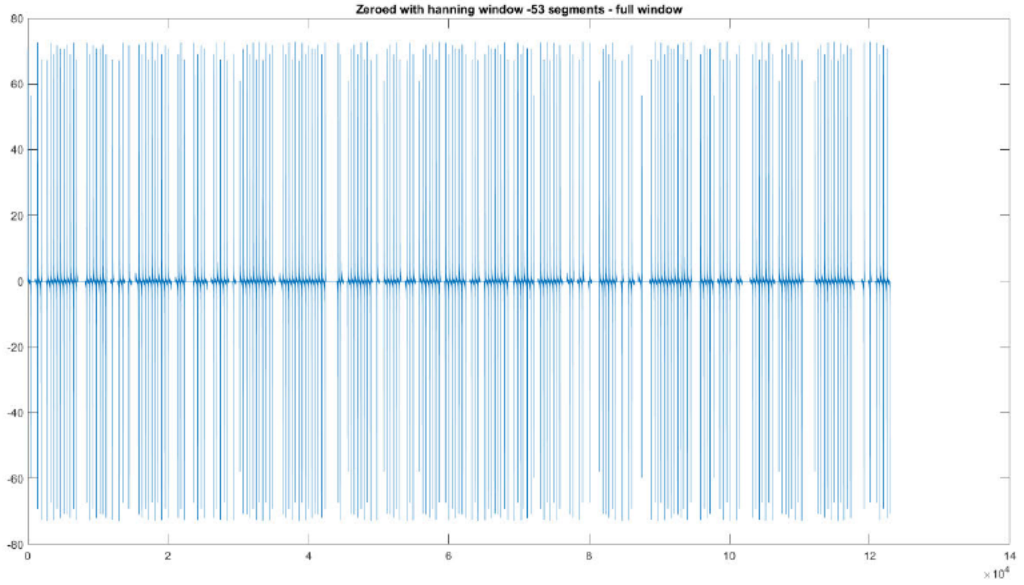
Two alternatives were also tested, first cutting out bad segments and concatenating zeroed out segments to the end of the temporal timeseries before spectral decomposition (i) showed that this greatly distorts the spectral timeseries, likely due to edge effects that are exacerbated.

To combat edge effects a version using a hanning window was trialled timeseries (j), close up (k) and spectral decomposition (l) are shown below. This ended up having the same effect as zeroing out one second segments.

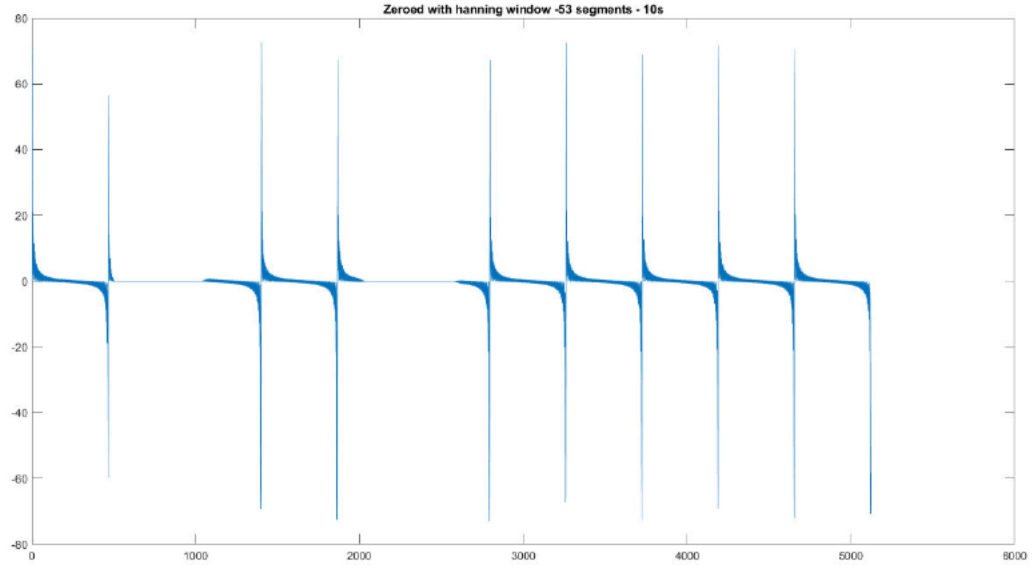
i


FFT - Cut and concatendated signal -53 segments removed
0
0.1
0.2
0.3
0.4
0.5
0.6
0
10
20
30
40
50
60
70
80
90
100

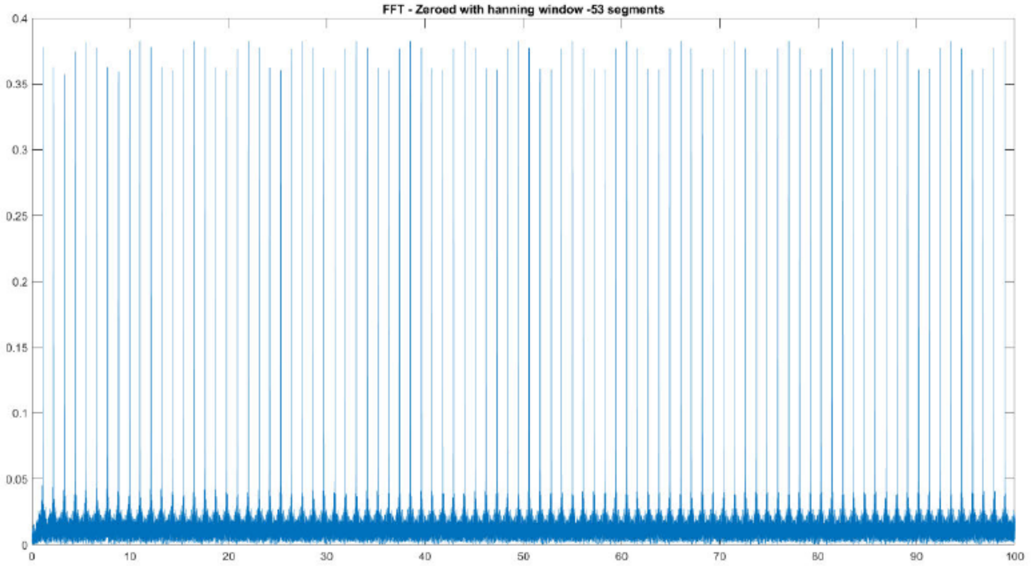
j


Zeroed with hanning window -53 segments - full window
-80
-60
-40
-20
0
20
40
60
80
0
2
4
6
8
10
12
14
×10^4

k


Zeroed with hanning window -53 segments - 10s
80
60
40
20
0
-20
-40
-60
-80
0
1000
2000
3000
4000
5000
6000

l


FFT - Zeroed with hanning window -53 segments
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
10
20
30
40
50
60
70
80
90
100

B.5 SNR scores for rest vs 2Hz condition

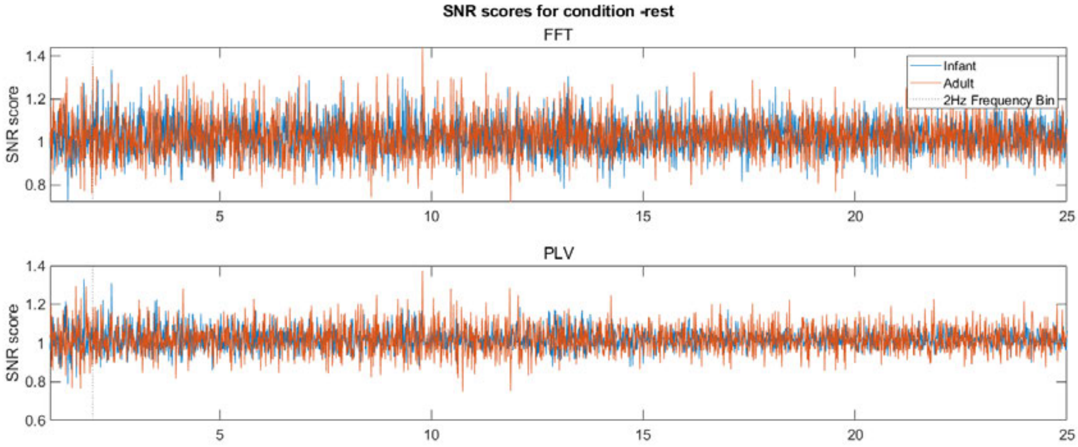


Figure B.5.1 Showing the full range of SNR scores between 1-25Hz for both the FFT and PLV analyses in the rest condition. Note that there is no particular spike above chance. 2Hz frequency bin has been marked on the left.

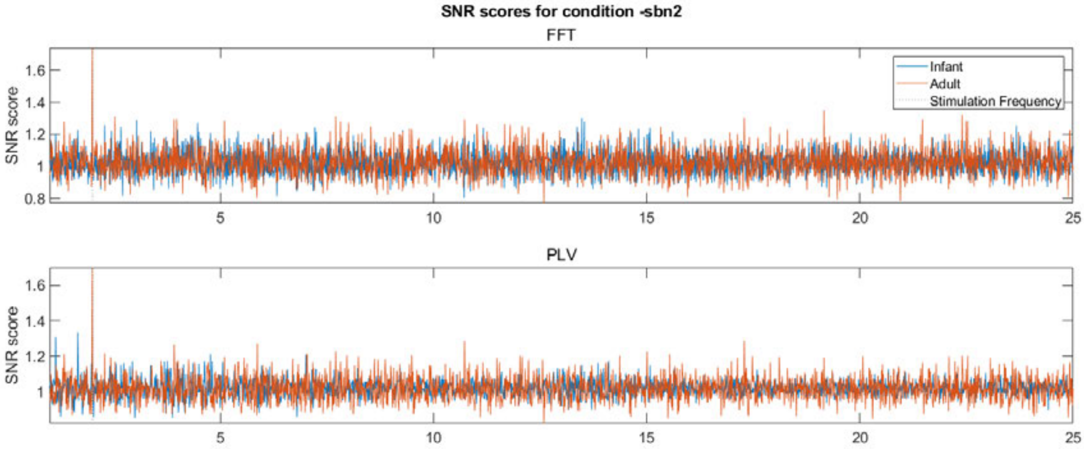


Figure B.5.2 Showing the full range of SNR scores between 1-25Hz for both the FFT and PLV analyses in the 2Hz condition. Note the spike on the left is at exactly 2Hz.

B.6 Bivariate correlation between infant and careful SNR scores

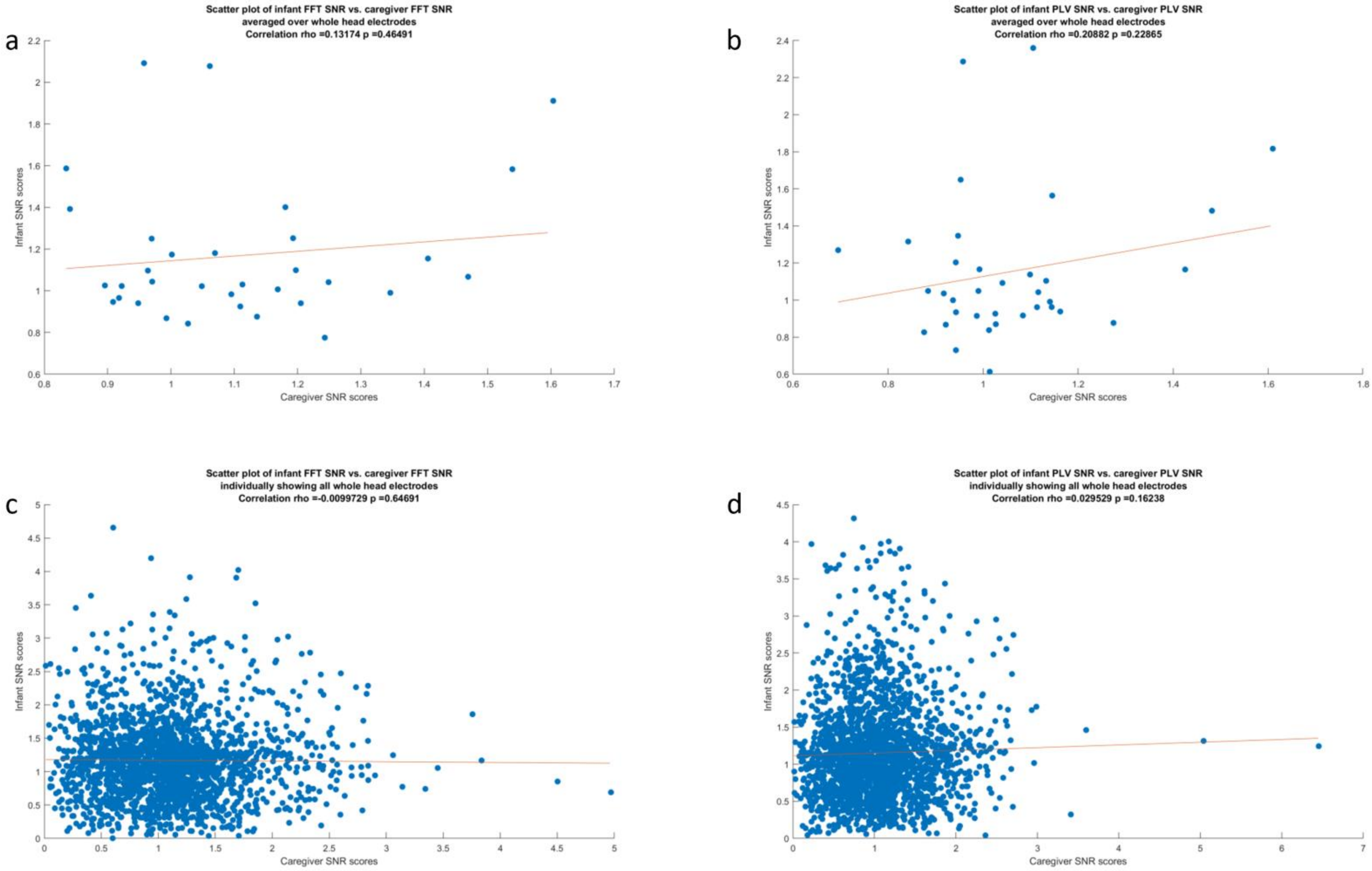


Figure B.6 Showing bivariate correlations of infant and caregiver data between SNR scores of FFT (a, c) and PLV (b, d) that have either been averaged per participant over all electrodes (a, b) or showing individual electrodes (c, d). Correlation scores and p values are as follows: a) SNR of FFT over averaged data r(31) = 0.13, p = 0.46, b) SNR of PLV over averaged data r(31) = 0.21, p = 0.23, c) SNR of FFT for each electrode r(2110) = -0.01, p = 0.65, d) SNR of PLV for each electrode r(2110) = 0.03, p = 0.16.

Appendix C - Supplementary Materials for: Do Adults and Infants Entrain to the Stimulus Beat or the Isochronous Rhythm? Evidence From a Novel Jittered Audio Approach

C.1 Conditions completed (post preprocessing)

Table C1. Showing which participants completed each of the conditions.

Participants	Rest		8Hz		8Hz 5%		8Hz 10%		8Hz 20%	
	Inf	Mum	Inf	Mum	Inf	Mum	Inf	Mum	Inf	Mum
1009			X	X						
1021			X	X						
1028			X	X	X	X	X	X	X	X
1029				X		X		X		X
1032			X	X	X	X	X	X	X	X
1039			X	X	X	X	X	X	X	X
1040			X	X	X	X	X	X		
1041					X	X	X	X		
1042			X	X	X	X	X	X	X	X
1046			X	X	X	X				
1051			X	X						
1052			X	X						
1053			X	X						
1054			X	X						
1055			X	X						
1059				X						
1060			X	X						
1061			X	X						
1062			X	X						
1063			X	X						
1064	X	X		X	X	X	X	X		X
1065		X		X		X		X		X
1066	X	X	X	X		X		X		X
1068	X	X	X	X	X	X	X	X	X	X
1069	X	X	X	X	X	X			X	X
1070	X	X	X	X	X	X	X	X	X	X
1073	X	X	X	X	X	X	X	X	X	X
1076	X	X	X	X					X	X
1077	X	X	X	X					X	X
1078	X	X	X	X	X	X		X	X	X
1079	X									
1082	X									
1083	X	X		X		X				

1084	X	X								
1085	X	X	X	X	X	X	X	X	X	X
1086	X	X	X	X	X	X	X	X	X	X
1087	X	X	X	X	X	X	X	X	X	X
1088			X	X	X	X	X	X	X	X
1089	X	X	X	X	X	X	X	X	X	X
1091	X	X								
1092	X	X								
1093	X	X		X		X		X		X
1096	X	X		X		X		X		X
1097	X	X	X	X	X	X	X	X		X
1099	X	X	X	X	X	X	X	X	X	X
1102	X	X	X	X		X				
1103	X	X	X	X	X	X	X	X	X	X
1104	X	X	X	X	X	X	X	X	X	X
1108	X		X		X		X		X	
1109	X	X	X	X	X	X	X	X		X
1114	X	X	X	X	X	X				
1117		X								
1128	X	X	X	X	X	X	X	X	X	X
1129	X	X	X	X	X	X	X	X	X	X
1135		X		X		X		X		

Table C2. Participant ages, infant age, gestation and whether the infant has siblings for visit 1.

Participant number	Adult age (years)	Infant age (days)	Infant gender	Infant gestation (weeks)	Other children
1009	38.39	155	M	44	N
1021		156	F		
1028	32.86	153	F	42	N
1029	36.82	185	F	44	N
1032	33.22	177	F	44	N
1039	33.73	150	M	40	N
1040	35.61	144	M		Y
1041	37.54	154	M	42	Y
1042	35.39	153	M	37.5	Y
1046	40.18	184	F	42	Y
1051	32.84	161	M	44	N
1052	32.37	172	F	47	N
1053	34.88	155	M	39	Y
1054	31.21	159	M	38.5	Y
1055	34.55	161	M	39	Y
1059		178	M		
1060		174	F		
1061	42.45	178	F		Y
1062		172	F		
1063	36.21	151	M	46	N
1064	31.71	182	M	41	Y
1065	40.11	190	M	43	Y
1066	34.58	167	M	39	N
1068	41.07	151	M	39	Y
1069	39.28	193	M	40.29	N
1070		166	F		
1071	31.54	159	F	42	N
1073	38.19	163	F	47	Y
1076	35.98	190	F		Y
1077	35.72	166	M	39	N
1078	34.35	181	M	38	N
1079	33.65	158	F	39.71	N
1082	29.81	194	M		N
1083	34.80	160	M	41	N
1084	39.54	173	M	42.14	N
1085	36.88	151	M		Y
1086	43.10	144	M	43	Y
1087	32.76	136	M	37.29	N
1088	39.24	153	M	37.86	Y
1089	28.64	196	F	40	
1091		187	M		
1092		183.0	M		

1093	33.47	149	F	42	N
1096	40.03	145	M		N
1097	34.49	160	M	38	N
1099	38.08		F	41.29	
1102		166	M		
1103	36.90	169	M	41	N
1104	34.74	153	F	38	Y
1108	38.26	186	F	39.57	Y
1109	37.23	137	F		N
1114	37.14	154	M	43	Y
1117		172	F	37	Y
1128	36.41	144	M	47	N
1129	32.47	210	M	41	
1135	35.99	218	F	42	N

C.2 Interstimulus intervals after random assignment of flat distribution of lags

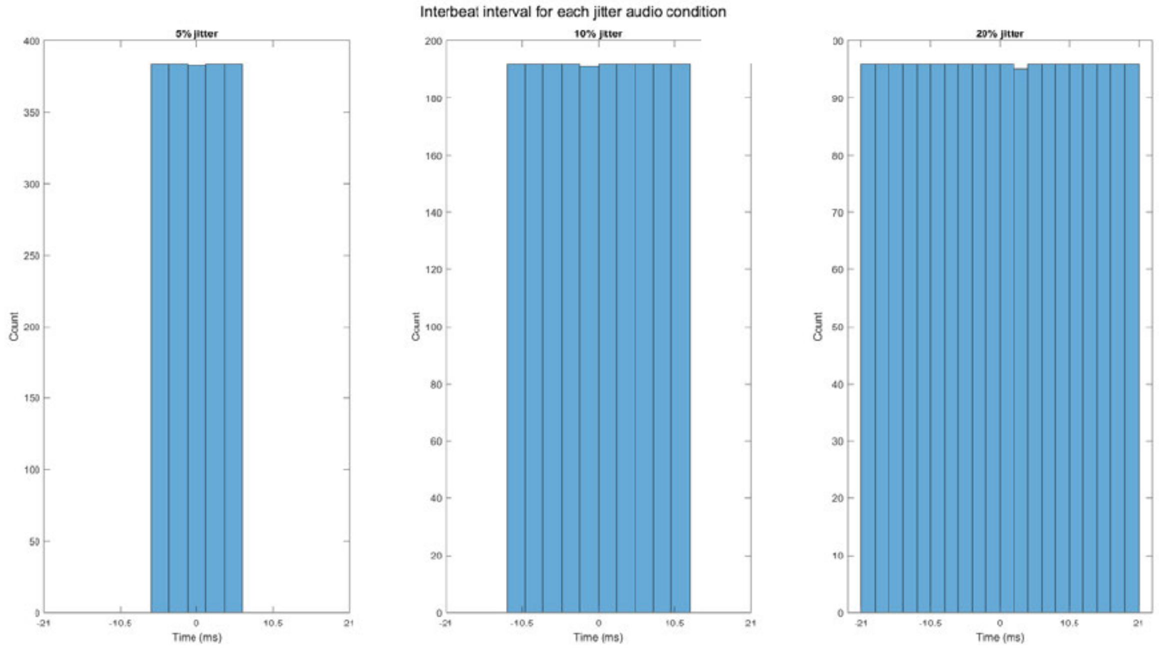


Figure C.2 Histograms showing the distribution interbeat intervals of beats around the rhythm. The x axis shows time 0 representing where the isochronous rhythm would be.

C.3 Data quality metrics per participant and condition

Table C.3. showing the data quality metrics for infant and adult participants by condition. Blue shaded columns represent the percentage of channels identified as noisy or bridged. Any condition with 1-25% channels identified as noisy or bridged had these channels interpolated (see section 2.5), any condition with 25-75% of channels identified as noisy or bridged had these channels rejected and removed from the dataset. White shaded columns show the percentage of one second segments zeroed out (see section 2.5). All values shown are percentages, any blank conditions indicate that either the participant didn't complete this condition or the data were rejected. At the end of each participant section there are rows showing the N for each condition, the average percentage of the number of channels rejected and the percentage of 1 second segments that were zeroed out, the average number of channels rejected above and below the 25% threshold for interpolation, the total number of channels in each condition across all participants and the total number of seconds of EEG in each condition.

Infant participants										
	Rest		**0% jitter**		**5% jitter**		**10% jitter**		**20% jitter**	
1009			23.00	6.70						
1021			16.00	0.00						
1028			25.00	4.66	14.00	7.39	14.00	1.19	14.00	0.42
1032			13.00	0.00	9.00	0.00	11.00	0.00	13.00	0.21
1039			2.00	1.65	0.00	0.21	3.00	0.00	2.00	0.42
1040			0.00	0.00	0.00	0.00	0.00	1.68		
1041					27.00	1.82	19.00	0.00		
1042			5.00	0.90	56.00	6.71	44.00	0.49	45.00	1.82
1046			0.00	0.00	2.00	0.84				
1051			8.00	29.85						
1052			9.00	0.00						
1053			36.00	1.96						
1054			52.00	1.56						
1055			45.00	0.00						
1060			3.00	0.00						
1061			38.00	0.00						
1062			31.00	0.00						
1063			30.00	0.00						
1064	52.00	34.54			0.00	0.00	14.00	4.23		
1066			50.00	30.63						
1068	47.00	8.25	47.00	2.33	41.00	2.75	45.00	0.00	75.00	38.26
1069	47.00	0.00	42.00	0.00	64.00	0.63			69.00	0.00
1070	17.00	0.35	20.00	0.00	61.00	3.15	72.00	15.26	20.00	0.00
1073	23.00	0.00	41.00	0.00	36.00	0.00	39.00	0.00	33.00	10.87
1076	20.00	0.00	19.00	0.00					22.00	1.25
1077	9.00	25.62	17.00	33.33					9.00	36.03
1078	14.00	0.63	17.00	0.00	16.00	0.00			16.00	0.00

1079	6.00	0.00								
1082	9.00	17.76								
1083	27.00	0.00								
1084	13.00	0.00								
1085	19.00	0.00	28.00	0.00	27.00	2.11	19.00	1.62	31.00	3.73
1086	0.00	2.73	3.00	3.01	2.00	5.20	2.00	6.25	5.00	7.21
1087	33.00	0.00	22.00	29.15	22.00	0.56	22.00	0.00	22.00	20.64
1088			22.00	1.97	22.00	0.28	22.00	0.84	22.00	0.00
1089	48.00	0.00	42.00	0.77	44.00	1.12	44.00	0.00	45.00	0.63
1091	0.00	0.00								
1092	20.00	2.11								
1093	14.00	0.00								
1096	34.00	2.04								
1097	39.00	0.00	23.00	7.25	30.00	7.81	23.00	18.95		
1099	13.00	2.17	14.00	4.42	25.00	1.19	22.00	3.38	20.00	4.37
1102	16.00	0.00	66.00	50.57						
1103	6.00	0.00	6.00	0.00	5.00	0.00	2.00	3.51	14.00	2.18
1104	52.00	0.91	52.00	15.55	61.00	16.21	42.00	27.77	59.00	14.58
1108	33.00	7.84	45.00	0.35	33.00	9.71	33.00	27.87	28.00	27.43
1109	39.00	0.00	23.00	7.25	30.00	7.81	23.00	18.95		
1128	23.00	1.41	45.00	9.10	59.00	2.38	25.00	1.50	52.00	1.91
1129	6.00	4.66	2.00	0.00	2.00	0.00	6.00	0.00	2.00	0.00
N	29		39		26		23		22	
Avg %	23.41	3.83	25.18	6.23	26.46	3.00	23.74	5.80	28.09	7.82
Avg % <25	12.67		12.70		9.15		14.19		13.92	
Avg % > 25	41.00		42.06		42.43		43.00		48.56	
Total elecs	1421		1868		1226		1123		1013	
Total segs		6693		8776		6052		5200		4868
Adult participants										
	Rest		**0% jitter**		**5% jitter**		**10% jitter**		**20% jitter**	
1009			19.00	14.54						
1021			16.00	5.80						
1028			3.00	37.91	3.00	36.69	3.00	38.70	3.00	33.89
1029			20.00	37.65	17.00	23.68	9.00	27.78	8.00	25.79
1032			5.00	5.47	5.00	4.61	5.00	5.38	5.00	4.61
1039			19.00	24.69	16.00	10.56	11.00	12.92	16.00	14.11
1040			6.00	5.21	6.00	12.59	6.00	1.19		
1041					23.00	16.70	31.00	39.98		

1042			0.00	2.35	2.00	3.77	2.00	0.84	2.00	0.49
1046			11.00	22.18	8.00	12.18				
1051			5.00	0.70						
1052			47.00	5.89						
1053			9.00	11.66						
1054			13.00	52.53						
1055			6.00	3.83						
1059			13.00	14.67						
1060			8.00	19.64						
1061			19.00	21.04						
1062			33.00	12.35						
1063			23.00	14.22						
1064	11.00	13.78	13.00	21.90	17.00	16.88	13.00	18.27	13.00	24.05
1065	17.00	1.05	17.00	3.82	17.00	0.00	14.00	1.33	16.00	0.91
1066	6.00	2.95	6.00	11.97	6.00	14.71	6.00	17.02	8.00	11.90
1068	5.00	26.59	3.00	18.38	2.00	19.94	2.00	24.00	2.00	21.38
1069	11.00	3.73	11.00	0.70	11.00	3.29			11.00	2.17
1070	17.00	12.96	16.00	9.55	16.00	14.14	20.00	31.56	16.00	21.32
1073	8.00	0.00	8.00	0.00	8.00	0.49	8.00	6.64	8.00	14.55
1076	9.00	36.73	6.00	51.04					5.00	44.77
1077	14.00	23.95	13.00	13.62					16.00	4.12
1078	25.00	29.11	28.00	14.79	30.00	1.69	30.00	1.54	31.00	0.63
1083	38.00	37.31	16.00	18.36	31.00	21.61				
1084	11.00	8.39								
1085	13.00	16.32	11.00	8.60	6.00	4.97	9.00	9.73	2.00	11.90
1086	38.00	39.32	39.00	28.64	39.00	29.86	45.00	47.02	45.00	35.69
1087	2.00	3.91	2.00	11.96	2.00	10.35	2.00	9.37	2.00	8.52
1088			6.00	15.58	3.00	28.68	2.00	17.71	6.00	25.45
1089	22.00	8.90	22.00	0.63	22.00	0.56	22.00	0.00	19.00	5.20
1091	2.00	0.00								
1092	13.00	3.91								
1093	6.00	11.07	11.00	10.43	6.00	19.57	6.00	27.53	8.00	15.76
1096	42.00	26.62	44.00	22.14	44.00	5.46	45.00	2.24	45.00	29.99
1097	19.00	2.31	22.00	2.24	22.00	1.82	22.00	6.87	28.00	5.33
1099	27.00	7.57	22.00	5.05	22.00	0.28	22.00	1.96	20.00	6.80
1102	44.00	45.28	53.00	38.25	77.00	33.85				
1103	27.00	13.11	22.00	0.00	25.00	0.91	25.00	0.00	41.00	8.97
1104	17.00	49.23	17.00	42.27	13.00	53.52	16.00	27.81	16.00	45.21
1109	19.00	2.31	22.00	2.24	22.00	1.82	22.00	6.87	28.00	5.33
1114	8.00	15.35	8.00	20.47	6.00	15.71				
1117	9.00	36.02								
1128	56.00	2.12	63.00	0.77	56.00	2.25	56.00	0.00	59.00	0.77

1129	42.00	1.54	42.00	1.26	41.00	4.62	42.00	1.27	42.00	1.61
1135	45.00	10.30	45.00	9.99	45.00	6.29	45.00	18.17		
N	31		47		34		29		29	
Avg %	20.10	15.86	18.36	14.83	19.68	12.77	18.66	13.92	17.97	14.87
Avg % <25	12.00		12.34		11.77		11.23		9.62	
Avg % > 25	38.40		43.78		43.11		39.88		39.88	
Total elecs	1587		2290		1687		1448		1463	
Total segs		6260		9608		7120		5990		5924

Appendix D - Supplementary Materials for: Do Infants Show More Neural Tracking to More Rhythmic Speech and Song?

D.1 Conditions completed (post preprocessing)

Table D1. Participant completion of each condition at each age visit.

	ID speech			**ID song**		
Participant number	**Visit 1**	**Visit 2**	**Visit 3**	**Visit 1**	**Visit 2**	**Visit 3**
1001						X
1003					X	X
1006		X	X		X	X
1010		X			X	X
1013		X			X	
1015		X	X		X	
1016		X			X	
1017						X
1018	X			X		
1022					X	
1023	X	X	X	X		X
1024	X	X	X		X	X
1026	X	X		X	X	X
1028	X			X		
1029		X			X	X
1030		X				
1032				X		
1033	X					
1034		X				
1036	X			X		
1037		X			X	
1039	X			X		
1040	X	X		X		
1042					X	
1043	X			X		
1044	X		X			X
1046	X					
1048				X		
1049	X		X			X
1050					X	X
1052	X	X			X	
1054	X	X			X	
1055	X			X		
1057	X					

1058	X		X			
1060	X			X		
1061	X			X		
1062	X	X	X		X	X
1063		X	X		X	X
1064	X			X		
1065				X		
1066	X	X		X		
1067	X	X		X	X	
1068	X					
1069	X			X		
1070	X			X		
1071				X		
1072		X				
1073	X					
1076	X			X		
1077	X			X		
1078			X		X	
1079				X		
1080		X			X	
1081	X			X		
1083	X			X		
1084	X			X		
1086	X			X		
1087	X	X		X	X	
1088				X		
1091	X			X		
1092					X	
1093	X			X		
1095	X			X		
1096	X					
1097	X					
1098				X		
1099	X			X		
1101	X			X		
1105					X	
1111				X		
1112	X					
1115	X			X		
1116	X			X		
1118				X		
1119	X			X		
1126				X		
1128	X					
1133	X			X		

1134	X					

D.2 Participant details

Table D.2 shows adult and infant age, gestation, gender and infant siblings. Timepoints v1-3 represent infants at 5, 10 and 15 months respectively. Blank cells represent either that no usable data was collected, the infant did not age into that timepoint by the data cut off data or the adult caregiver did not want to answer the questions.

Participant number	Adult age (years)			Infant age (months)			Infant gender	Infant gestation (weeks)	Other children
	V1	V2	V3	V1	V2	V3			
1001			39.23			15.11	M		
1003					10.82	17.44	M		
1006		26.48	26.92		10.26	15.61	F	42	N
1010		38.28	38.72		11.61	16.89	F		
1013		36.32			10.16		F	41	N
1015		28.78	29.32		9.61	16.00	F	43	N
1016		38.42			10.13		M	39	N
1017			36.28			16.56	F	41	Y
1018	40.56			4.72			M	43	Y
1022		40.16			11.54		M	38.5	
1023				4.95	9.77	15.08	M		
1024				5.97	9.41	14.98	M		
1026	31.01	31.39	31.79	5.15	9.70	14.43	F		N
1028	32.86			5.02			F	42	N
1029		37.14	37.53		9.90	14.52	F	44	N
1030		35.66			10.16		F	41	N
1032	33.22			5.80			F	44	N
1033				5.05			M		
1034					10.33		M		
1036	31.58			4.66			F		
1037		34.01			10.62		M		N
1039	33.73			4.92			M	40	N
1040	35.61	36.07		4.72	10.26		M		Y
1042		35.76			9.38		M	37.5	Y
1043	39.98			4.20			F	45	Y
1044	37.34		38.11	5.61		14.85	M	44	Y
1048	38.48			7.28			M	42	Y
1049	34.54		35.12	9.18		16.13	F	40	Y
1050		44.18	44.60		10.10	15.18	F	44	N
1052	32.37	32.91		5.64	12.16		F	47	N
1054	31.21	31.58		5.21	9.70		M	38.5	Y

1055	34.55			5.28			M	39	Y
1057	32.33			5.41			M	43	Y
1058	38.29		39.19	4.98		15.80	M	41	Y
1060				5.70			F		
1061	42.45			5.84			F		Y
1062					10.66	14.10	F		
1063		36.59	37.10		9.57	15.57	M	46	N
1064	31.71			5.97			M	41	Y
1065	40.11			6.23			M	43	Y
1066	34.58	35.00		5.48	10.49		M	39	N
1067	39.07	39.46		5.51	10.26		M	40.5	N
1068	41.07			4.95			M	39	Y
1069	39.28			6.33			M	40.29	N
1070				5.44			F		
1071	31.54			5.21			F	42	N
1072					13.15		M		
1073	38.19			5.34			F	47	Y
1076	35.98			6.23			F		Y
1077	35.72			5.44			M	39	N
1078		34.69	35.20		10.00	16.03	M	38	N
1079	33.65			5.18			F	39.71	N
1080					5.64	15.28	F		
1081	37.56			5.28			M	41	Y
1083	34.80			5.25			M	41	N
1084	39.54			5.67			M	42.14	N
1086	43.10			4.72			M	43	Y
1087	32.76	32.87		4.46	5.80		M	37.29	N
1088	39.24			5.02			M	37.86	Y
1091				6.13			M		
1092					11.84		M		
1093	33.47			4.89			F	42	N
1095	40.21			6.62			F	41	N
1096	40.03			4.75			M		N
1097	34.49			5.25			M	38	N
1098	33.93			5.05			F	45	Y
1099	38.08			7.02			F	41.29	
1101	35.70			5.08			M	43	N
1105		34.73			9.97		F	42	N
1111	34.46			8.36			M	37	Y
1112	35.04			5.11			F	44	
1115	30.63			2.92			M		N
1116	34.11			3.05			M		N
1118	37.20			5.11			F	40	
1119	34.14			13.34			M	46	

1126	30.47			4.89			M		N
1128	36.41			4.72			M	47	N
1133	33.30			6.10			F		N
1134	33.79			6.66			F	43	

D.3 Data quality metrics per participant and condition

Table D.3. showing the data quality metrics for infant participants by condition and age point. Blue shaded columns represent the percentage of channels identified as noisy or bridged. Any condition with 1-25% channels identified as noisy or bridged had these channels interpolated (see section 2.5), any condition with 25-75% of channels identified as noisy or bridged had these channels rejected and removed from the dataset. White shaded columns show the percentage of one second segments zeroed out (see section 2.5). All values shown are percentages, any blank conditions indicate that either the participant didn't complete this condition or the data were rejected. At the end of each participant section there are rows showing the N for each condition, the average percentage of the number of channels rejected and the percentage of 1 second segments that were zeroed out, the average number of channels rejected above and below the 25% threshold for interpolation, the total number of channels in each condition across all participants and the total number of seconds of EEG in each condition.

	ID speech						**ID song**					
Ppt number	**Visit 1**		**Visit 2**		**Visit 3**		**Visit 1**		**Visit 2**		**Visit 3**	
1001											69	**0.00**
1003									30	15.05	36	**3.50**
1006			66	51.19	27	49.47			48	27.24	36	**24.90**
1010			66	33.99					66	40.65	78	**42.16**
1013			23	46.05					64	37.22		
1015			36	1.52	67	0.00			38	9.44		
1016			55	57.13					42	47.69		
1017											52	**24.02**
1018	58	8.93					47	16.34				
1022									72	32.53		
1023	17	0.00	75	29.61	31	14.73	17	1.03			5	**21.67**
1024	22	1.00	41	0.87	36	1.47			48	30.71	19	**1.67**
1026	16	42.10	70	49.33			75	51.90	72	46.44	44	**11.87**
1028	55	14.72					70	20.15				
1029			23	50.55					27	0.90	61	**17.57**
1030			39	9.87								
1032							66	23.87				
1033	73	55.15										
1034			30	59.14								
1036	67	31.52					30	2.54				
1037			36	4.28					33	1.95		

1039	0	9.33					2	4.00				
1040	6	5.09	44	18.64			0	1.83				
1042									36	9.17		
1043	25	5.77					9	0.45				
1044	58	11.80			2	0.00					3	**0.22**
1046	3	0.71										
1048							66	12.89				
1049	73	23.17			63	0.00					66	**0.00**
1050									38	15.52	53	**11.64**
1052	6	5.84	77	13.36					11	1.58		
1054	75	49.81	23	1.75					72	20.15		
1055	55	2.47					33	16.47				
1057	69	39.54										
1058	22	1.38			67	28.13						
1060	39	0.00					22	0.00				
1061	36	4.87					14	3.05				
1062	33	10.62	63	72.53	17	6.06			33	39.51	6	**1.50**
1063			11	0.00	44	11.57			8	0.49	20	**6.75**
1064	67	17.56					69	8.91				
1065							23	56.91				
1066	69	2.20	8	2.51			64	0.46				
1067	75	50.56	56	1.11			55	26.21	64	2.06		
1068	27	40.85										
1069	31	10.21					20	5.02				
1070	28	10.33					38	2.81				
1071							31	11.12				
1072			39	63.87								
1073	20	2.66										
1076	61	66.56					28	24.97				
1077	8	2.54					33	12.54				
1078					55	41.42			67	7.12		
1079							61	53.28				
1080			34	2.39					30	40.26		
1081	33	52.29					39	45.19				
1083	30	63.95					30	45.95				
1084	42	3.59					73	35.74				
1086	5	13.45					5	22.73				
1087	59	14.89	19	52.80			75	19.73	13	38.28		
1088							6	30.92				
1091	45	37.59					30	46.89				
1092									67	28.36		
1093	30	0.56					28	0.00				

1095	56	11.78					23	29.73				
1096	70	1.52										
1097	61	41.94										
1098							56	36.43				
1099	44	18.76					39	31.25				
1101	42	6.06					38	5.38				
1105									20	0.41		
1111							48	23.09				
1112	38	48.99										
1115	52	23.29					52	23.63				
1116	9	2.85					3	0.35				
1118							53	67.32				
1119	9	60.40					53	9.71				
1126							58	8.07				
1128	44	0.69										
1133	13	1.03					16	0.59				
1134	77	5.62										**0.00**
N	50		23		11		43		24		15	
Avg %	39.86	19.11	42.45	28.29	40.90	15.28	38.05	19.99	43.43	21.42	39.14	19.11
Avg % <25	12.07		17.83		9.50		12.31		13.00		10.60	
Avg % > 25	51.34		51.69		48.75		49.59		49.84		55.00	
Total elecs	1887		811		379		1666		834		546	
Total segs		9513		3785		2032		8067		4338		9513

Appendix E - Supplementary Materials for: Rocking as soothing behaviour Impact of Dynamic Caregiver Rocking Characteristics on Infant Soothing, Evidence from Day Long Home Recordings

E.1 Participant details

Table E.1 showing participant details, adult and infant age, infant gestation, whether the family had other children and the duration of the home recording.

Participant number	Adult age (years)	Infant age (days)	Infant gender	Infant gestation (weeks)	Other children	Recording Duration (Hours)
1001	38.36	141	M	42	Y	9.24
1003		149	M		N	7.12
1004	36.93	178	M	44	Y	5.10
1005	45.65	144	M	38	Y	8.18
1006	26.13	188	F	42	N	9.60
1011	32.46	167	F	41	Y	11.02
1013	35.95	176	F	41	N	7.78
1015	28.49	186	F	43	N	6.36
1017	35.31	151	F	41	Y	12.43
1022	39.68	176	M	38.5		7.24
1029	36.84	191	F	44	N	7.01
1030	35.31	181	F	51	N	7.60
1031	35.29	169	F	45	N	8.08
1032	33.22	178	F	44	N	7.58
1034		180	M			7.95
1037	33.62	184	M		N	8.70
1038	31.65	182	M	39	N	6.93
1039	33.81	182	M	40	N	9.11
1041	37.58	166	M	42	Y	6.14
1042	35.47	181	M	37.5	Y	8.18
1043	40.09	166	F	45	Y	12.43
1044	37.36	178	M	44	Y	8.73
1048	38.50	172	M	42	Y	8.91
1049	34.26	179	F	40	Y	10.88
1050	43.86	191	F	44	N	7.25
1052	32.45	201	F	47	N	7.86
1053	34.90	161	M	39	Y	8.11
1055	34.58	175	M	39	Y	8.33
1057	32.35	262	M	43	Y	8.00
1058	38.38	187	M	41	Y	9.12

1059		192	M			6.99
1060		188	F			8.09
1062		174	F			12.43
1063	36.28	179	M	46	N	8.08
1065	40.18	216	M	43	Y	6.81
1066	34.65	194	M	39	N	6.46
1068	41.12	172	M	39	Y	12.43
1070		181	F			7.70
1071	31.60	181	F	42	N	2.76
1073	38.28	195	F	47	Y	9.51
1076	36.00	200	F		Y	6.72
1077	35.73	167	M	39	N	8.43
1078	34.39	137	M	38	N	9.61
1079	33.67	165	F	39.71429	N	7.54
1080		173	F			12.43
1081	37.56	161	M	40.85714	Y	7.23
1082	29.84	204	M		N	6.35
1083	34.92	204	M	41	N	9.42
1086	43.12	152	M	43	Y	12.43
1087	32.85	171	M	37.28571	N	7.60
1088	39.28	168	M	37.85714	Y	7.28
1089	28.67	209	F	40		6.19
1093	33.52	170	F	42	N	7.17
1095	40.27	221	F	41	N	7.27
1096	40.06	154	M		N	12.43
1098	33.97	169	F	45	Y	7.53
1099	37.92	207	F	41.28571		12.43
1100	38.11	211	F		N	7.86
1101	35.88	220	M	43	N	8.02
1102		170	M			5.91
1103	37.09	239	M	41	N	7.28
1105	34.39	178	F	42	N	6.19
1108	38.39	232	F	39.57143	Y	6.17
1110	30.62	214	F	40	Y	12.43
1118	37.24	170	F	40		12.43

E.2 Example of audio and actigraphy preprocessing

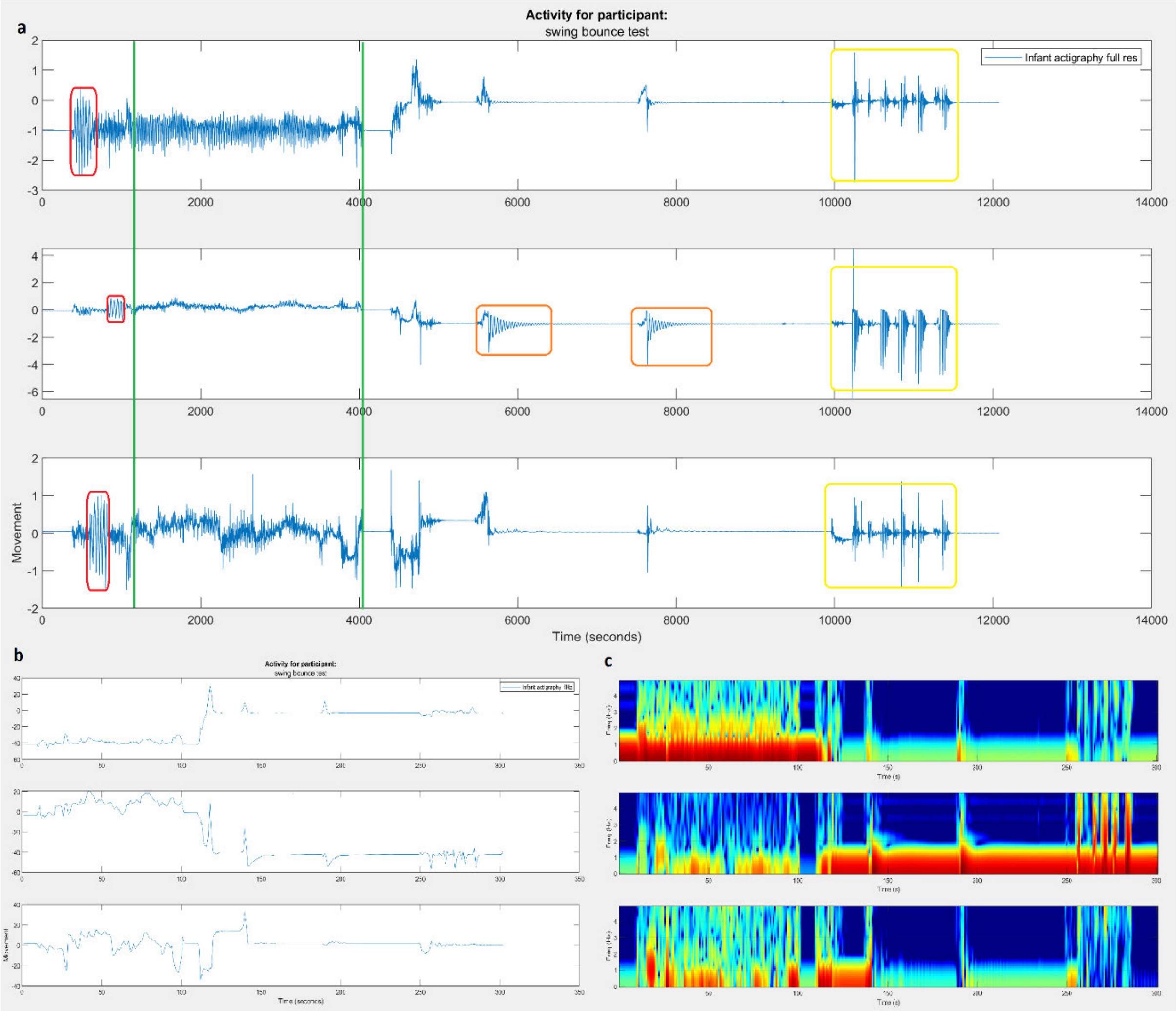


Figure E.2. Example raw actigraphy data produced using controlled settings. a) red shows times when controlled movements in each axial direction were made, between the green lines shows natural walking (thought to be a potential confound to finding oscillations), in the orange box lateral swinging was produced and in the yellow box vertical bouncing was produced; b) difference scores showing the level of activity for each of the axes does not distinguish rocking behaviours; c) time-frequency plots of the actigraphy data shows spurious oscillations up to 1Hz where there are now oscillations.

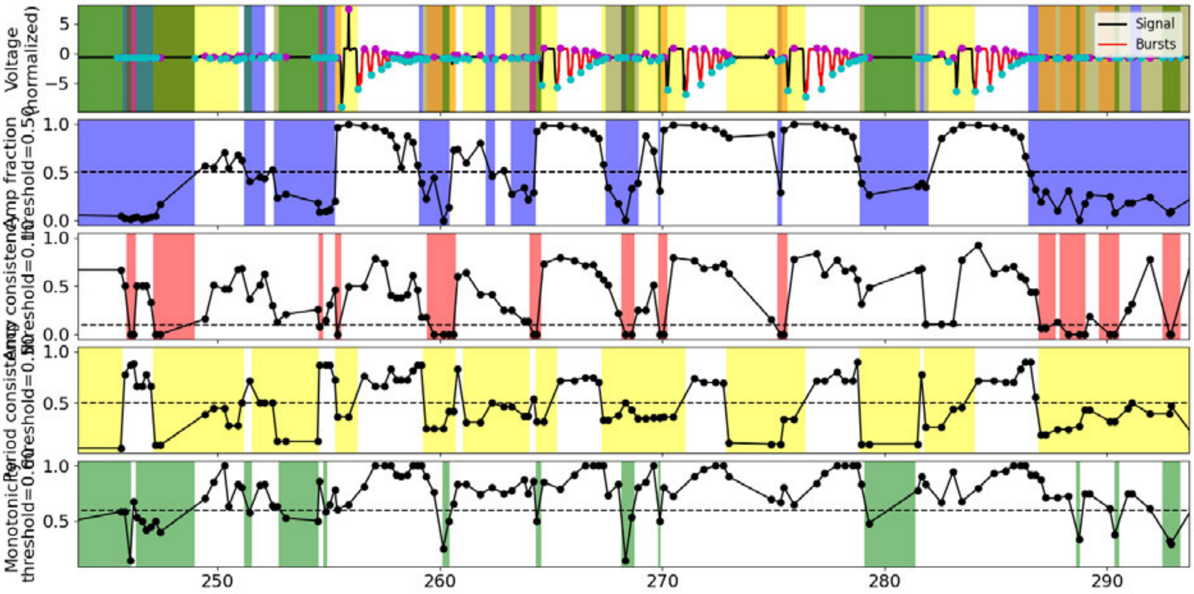


Figure E.3 Showing the parameters of the cycle-by-cycle analysis for the actigraphy data. Using the data from the orange box in figure 2a. a) shows the raw data and an overlay of any violations of the parameters below, any data that is determined to be an oscillation is highlighted in red; b-e show data highlighted for violating threshold of minimum amplitude (b), amplitude consistency (c), period consistency (d), and monotonicity (e).

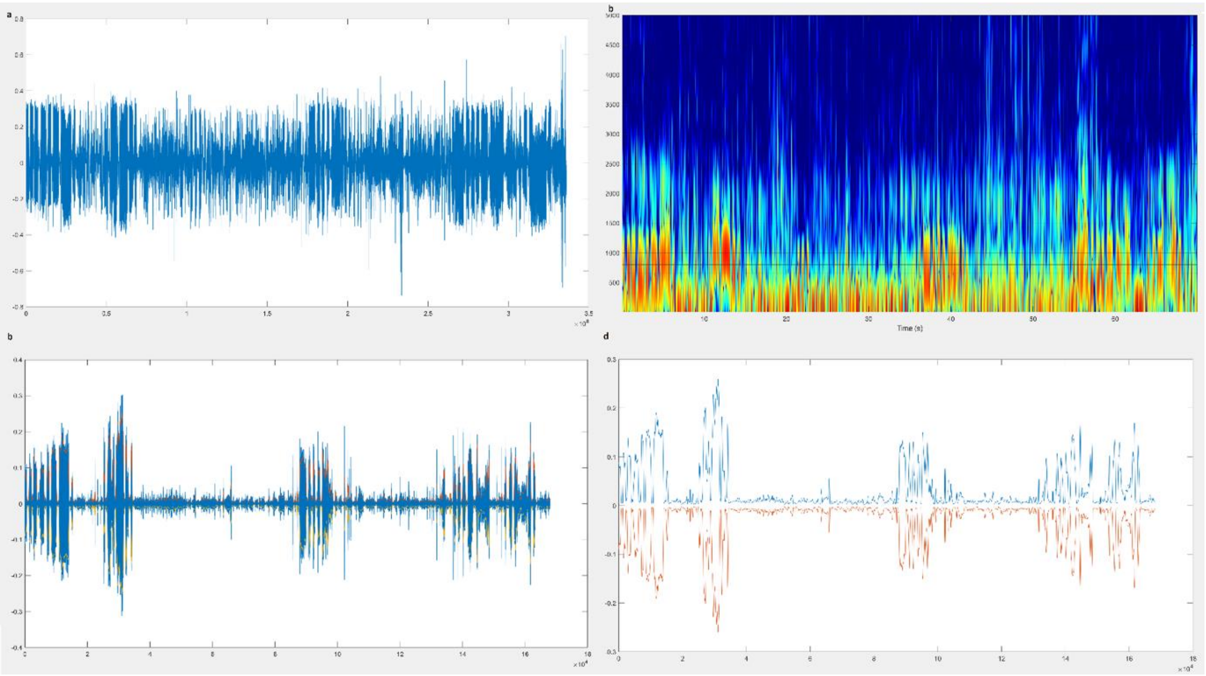


Figure E.4 showing a) a section of raw audio data, b) the spectral decomposition, and an indication of where infant crying has been detected; c) filtered signal between 800-100-Hz, d) amplitude envelope of the filtered audio signal

E.3 Proximity calibration and reliability testing

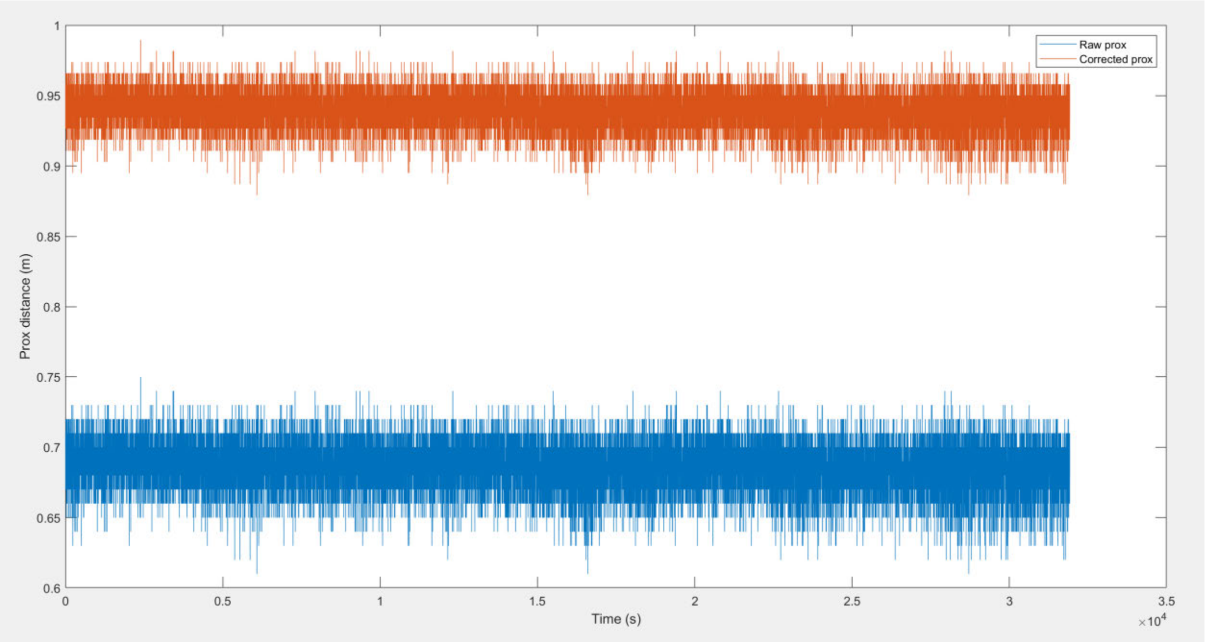


Figure E.5 Stationary proximity test, wearable devices were set up and left 1 metre apart for 9 hours to test whether there would be a drift in proximity readings. Raw proximity readings are shown along with data corrected with a polynomial distribution.

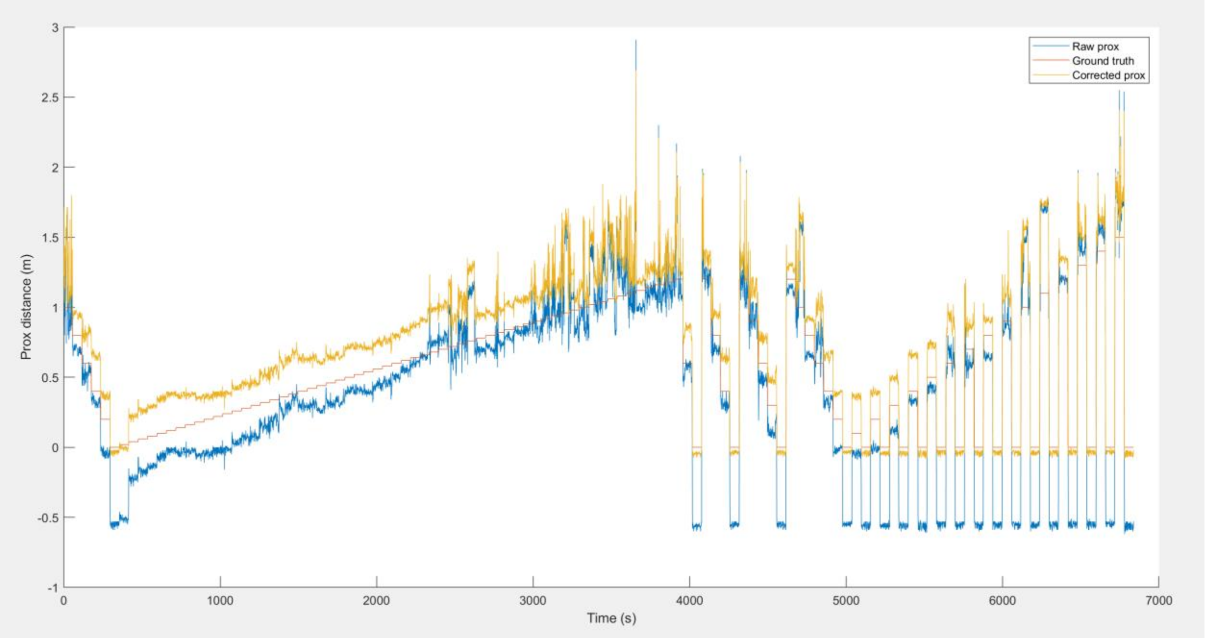


Figure E.6 Showing a calibration proximity test conducted by placing one wearable device a fixed distance from another for a minute at a time. Ground truth shown in orange shows the distance that the devices were kept from each other and the timings, the raw proximity values are shown in blue and the corrected values are shown in yellow.

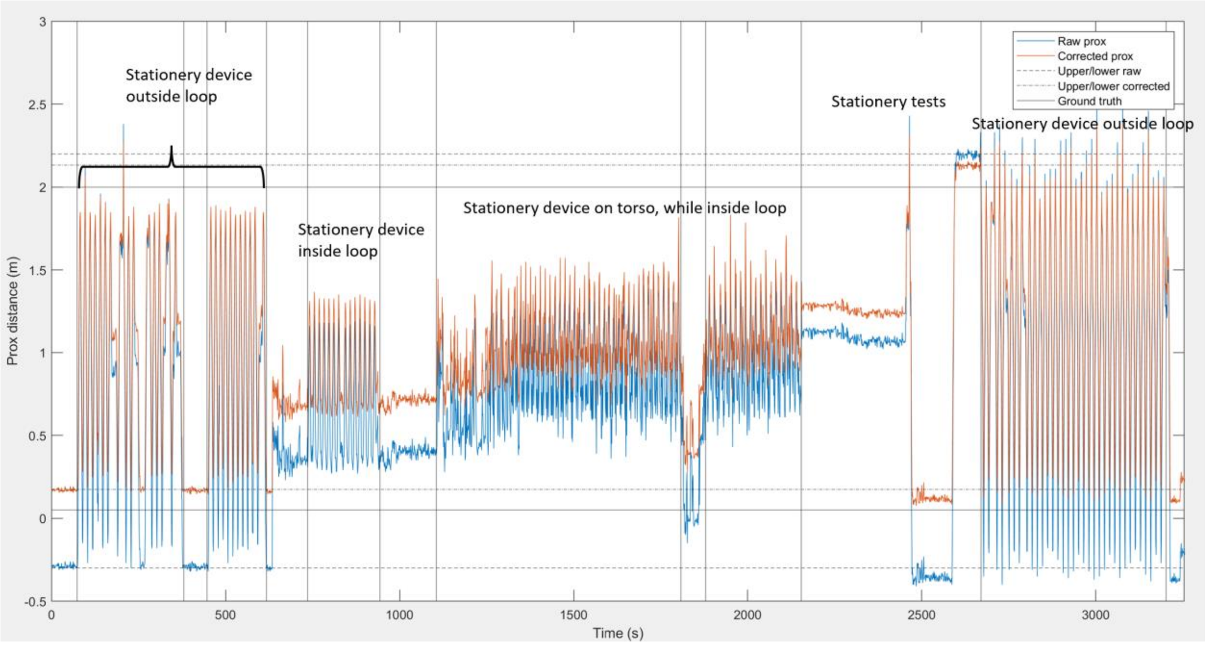


Figure E.7 *Moving proximity test conducted by placing one wearable device in a fixed location and the other on a moving toy train with a fixed looping track. Raw proximity data are shown in blue, while corrected proximity data are shown in orange. The test involved placing the stationary device outside the track then inside the track, then inside the track by strapped to a torso, then stationary, then back outside the track. This was to test whether movement was registered in the same way from different distances, whether the angle of the devices had an impact on the readings, whether obstacles such as a torso would impact readings, whether there would be discrepancies from stationarity and to retest the original test.*

E.8 Descriptives showing the amount of crying and rocking within bouts

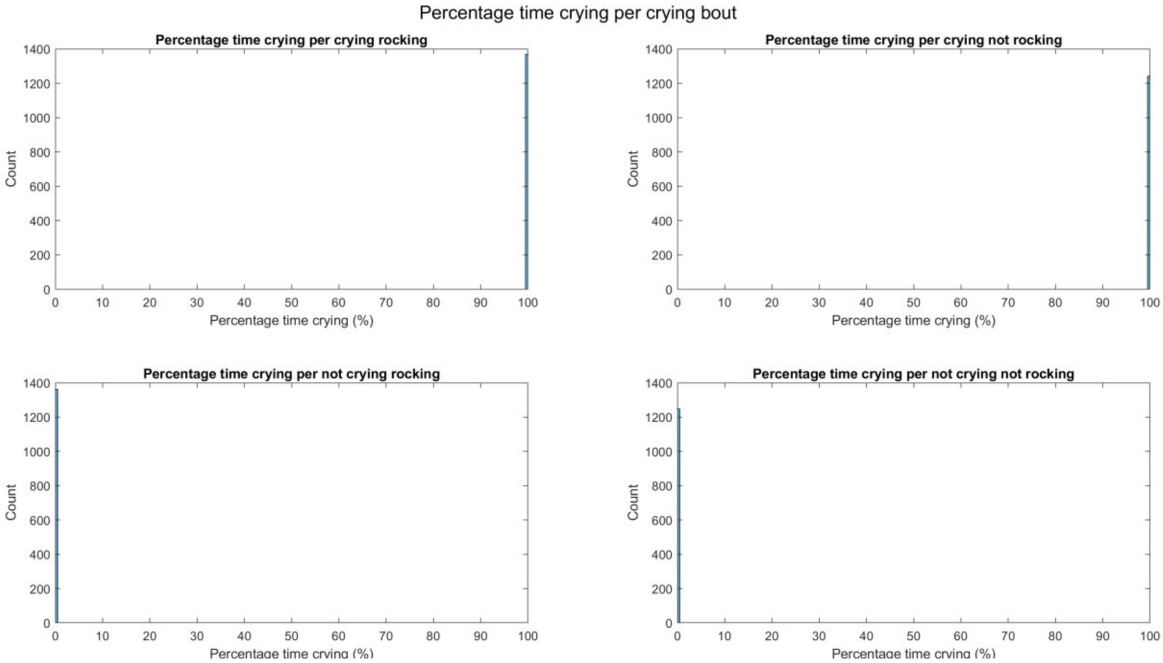


Figure E.8 Histograms showing the percentage of each crying bout spent with detected cries. Top left crying-rocking, top right, crying-not rocking, bottom left not crying-rocking, bottom right not crying-not rocking.

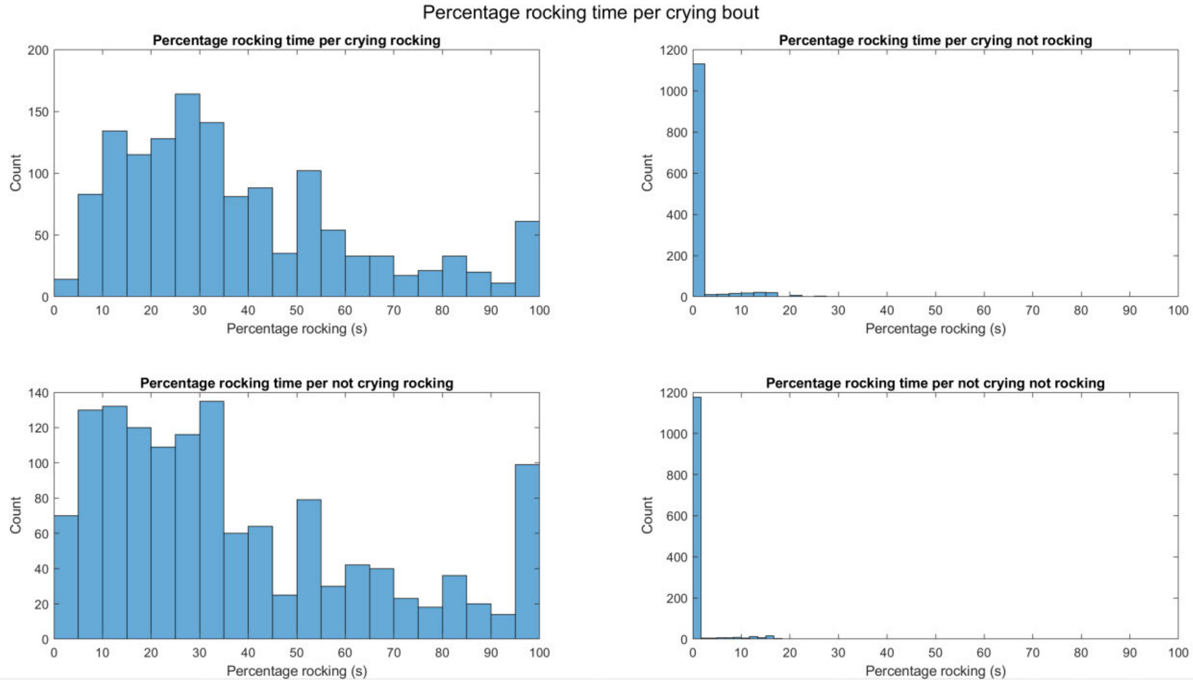


Figure E.9 Histograms showing the percentage time spent rocking during each crying bout. Top left crying-rocking, top right, crying-not rocking, bottom left not crying-rocking, bottom right not crying-not rocking.

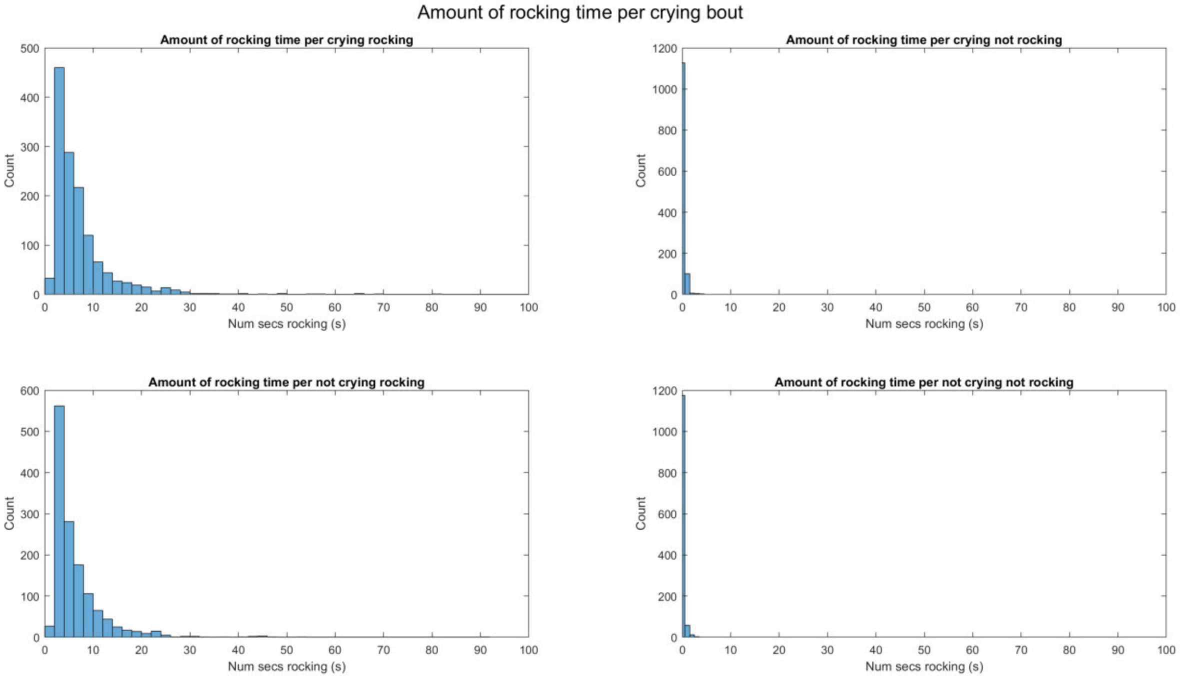


Figure E.10 Histogram showing the total number of seconds of rocking for all crying bouts for each of the conditions. Top left crying-rocking, top right, crying-not rocking, bottom left not crying-rocking, bottom right not crying-not rocking.