Novel Biomarkers in Infants: Developing Optical Imaging Solutions for the Measurement of Early Brain Development

A thesis submitted to The University of Manchester for the degree of Doctor of Philosophy (PhD)
in the Faculty of Biology, Medicine and Health

2018

Chen Zhao

School of Health Sciences

List of Contents

Word Count: 55,585

List of Tables

List of Figures

List of Abbreviations

ANOVA	Analysis of Variance
dB	Decibel
EEG	Electroencephalogram
ERP	Event Related Potential
fMRI	functional Magnetic Resonance Imaging
fNIRS	functional Near-Infrared Spectroscopy
FDR	False Discovery Rate
FFT	Fast Fourier Transformation
ICA	Independent Component Analysis
MACI	Manchester Assessment of Caregiver-Infant Interaction
NHS	National Health Service
SNR	Signal-to-Noise Ratio
SPL	Sound Pressure Level
STC	Superior Temporal Cortex
WHO	World Health Organization
ΔHbO2	Oxy-Haemoglobin concentration changes
ΔHbR	Deoxy-Haemoglobin concentration changes

Abstract

A PhD thesis (Medicine) submitted to the University of Manchester for the Faculty of Biology, Medicine and Health by Chen Zhao, September 2018.

Language and social-emotional atypicalities may act as important early markers of potential developmental problems in children. Advances in psychophysiological tools have led to an understanding of the neural correlates that underpin infant voice and vocal emotion processing as the precursors to language and social-emotional brain development. However, there is little evidence for the longitudinal developmental trajectory of infant voice and vocal emotion processing. In addition, the influence of early experience, especially maternal caregiving behaviour, on the infant voice and vocal emotion processing is less clear. To address research gaps relating to the early infant language and social-emotional neural development, this PhD study involved developing an acceptable and reliable method to observe changing voice and vocal emotion processing in the infant brain in the first year of life. The study aimed (a) to develop and pilot a functional Near-Infrared Spectroscopy (fNIRS) protocol that could be acceptably and feasibly applied in a clinical setting in order potentially to monitor and evaluate the neurological underpinnings of early language and social-emotional development; (b) to track longitudinal changes in neural correlates of human voice and emotional vocalisation processing in typically developing infants at the age of 6, 9 and 12 months; (c) to explore the role of early maternal caregiving behaviour as a possible mediator in the healthy development of voice and vocal emotion processing in infants.
Chapter 2 describes methodological considerations in the longitudinal assessment of infant voice and vocal emotion neural processing. Chapter 3 compared fNIRS data analysis methods between standardised analysis and Independent Component Analysis (ICA) implanted emotion-sensitive components identification procedures (paper 1). Chapter 4 (paper 2) reports the acceptability and feasibility of piloting the fNIRS paradigm in a non-research, community-based clinical setting in typically developing infants at their ages of 6, 9 and 12 months. Chapter 5 to 7 (paper 3 to 5) report the observation of infant neural development in processing emotional vocalisations, human voice and non-vocal sounds over time, and the influence of maternal caregiving behaviour on the infant's neural response to vocal emotion. Mother-infant pairs successfully completed the longitudinal study with high retention and satisfaction rate. We found consistent temporal cortical activations to auditory stimuli, particularly to happy vocalisations and non-vocal sounds, but the activation locations within temporal cortices were inconsistent over time. Right temporal cortical responses to angry vocalisations were significantly increased with age. Six-month-old infant's neural responses to vocal anger were significantly associated with maternal directiveness. We also report a lack of significantly enhanced neural responses to voice than to non-vocal stimuli. Findings suggest that the fNIRS paradigm has a promising outlook for the clinical use in infants and children. The infant neural responses to human voice stimuli are continuously being "fine-tuned" in temporal cortices. We also report a possible interaction with the quality of maternal caregiving. Infants may have distinct neural developmental trajectories of processing positive and negative emotional vocalisations. Future research will benefit from exploring links between neural responses to human vocalisations and language and social-emotional development in typically and atypically developing infants and children. Future study is also needed to optimise infant fNIRS data analysis, especially with the implementation of ICA.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification at this or any other University or institute of learning.

Copyright Statement

i The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property University IP Policy and/or Reproductions described in it (see http://documents.manchester.ac.uk/display.aspx?DocID=24420), in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in The University's policy on Presentation of Theses.

Acknowledgements

First all of all I would like to thank all the mothers and infants participated in this longitudinal study, your support and dedication of time have built the basis of this PhD project and added a block to this important area.

I would like to thank my supervisors: Prof Kathryn M Abel, Dr Ingo Schiessl, Dr Ming Wai Wan, and Dr Georgia Chronaki, for giving me the opportunity to become a PhD student. Had it not been for your continued support, patient guidance and expertise, I would not have been able to complete this PhD.

I would also like to give my thanks to the China Scholarship Council for their generous support and funding for my PhD study in the UK.

My special thanks go to my office buddies. I am blessed for being surrounded by such amazing people. You have been so kind and generously offering help to me. Every time I think about Amy Degnan, Rebekah Shallcross, Henna Lemetyinen, Heather Mitchell, Sabina Vatter, Gemma Stringer, Hein Heuvelman, I feel so warm from the depth of my heart, and I shall always remember the joyful days with you guys.

Finally, I would like to thank my mum for her endless love, support and encouragement for all these years.

Chapter 1 Introduction

1.1 The importance of early language and social-emotional development

1.1.1 The importance of early vocal and language development

Effective communication skills represent a pivotal component of social and cognitive development and are an essential building block for healthy relationships. Not only speech cues, but also non-speech cues (e.g. vocal cues), are used in social communication to allow an individual to understand information (e.g. social status, physical condition, or affective state) of the other (Belin, Fecteau, & Bédard, 2004). It has been widely acknowledged that social cognition skills, including early and concurrent language ability, are robustly associated with future successful peer interactions and interpersonal relationships (Hebert-Myers, Guttentag, Swank, Smith, & Landry, 2006).

The literature has supported links between atypical language development and developmental disorders and mental health problems in children and adults. For example, delayed language development has been found to be associated with lower IQ and reading difficulties at the age of seven (Silva, Mcgee, & Williams, 1983). In addition, psychiatric and behaviour problems, such as internalising and externalising behaviours, often co-occur with childhood language problems (Beitchman et al., 1996; Beitchman, Nair, Clegg, Ferguson, & Patel, 1986; Maggio et al., 2014; Toppelberg & Shapiro, 2000). This high prevalence of mental health problems among language delayed children may persist in adolescence (Beitchman et al., 2001; Conti-Ramsden & Botting, 2008), and adulthood (Beitchman et al., 2001; Botting, Durkin, Toseeb, Pickles, & Conti-Ramsden, 2016; Carton, Kessler, & Pape, 1999). Beitchman et al.'s 14-year longitudinal study reported sustained comorbidity

of anxiety disorder and anti-social personality disorder in children and young adults who had speech and language impairments (Beitchman et al., 1996; Beitchman et al., 2001). Because of the wide variation in normal language development (Bates, Thal, & Janowsky, 1992; Cole, Tamang, & Shrestha, 2006; Fenson et al., 1994; Kitayama, Markus, & Matsumoto, 1995), many, if not most children with language or social-emotional developmental problems are not diagnosed until between the ages of 3 - 6 years old (Palfrey, Singer, Walker, & Butler, 1987; Pinto-Martin, Dunkle, Earls, Fliedner, & Landes, 2005).

Language development begins in an infant from birth; research has increasingly moved towards the infant's pre-linguistic development and its association with future language abilities. An infant is born with the ability to distinguish syllables irrespective of language, and according to Kuhl's native language neural commitment concept, at around the sixth month, neural processing of native speech patterns starts to stabilise while the infant's ability to discriminate non-native language diminishes (Kuhl, 2004). This early neural basis for native speech perception ensures the smooth language learning process. By evaluating the development of early infancy speech discrimination, it is possible to predict future language acquisition. Empirical evidence from behavioural and cognitive neuroscience studies support Kuhl's theory (Kuhl, Conboy, Padden, Nelson, & Pruitt, 2005; Kuhl, Ramirez, Bosseler, Lin, & Imada, 2014; Kuhl & Rivera-Gaxiola, 2008; Molfese, 2000; Molfese & Molfese, 1985, 1997; Tsao, Liu, & Kuhl, 2004). Speech perception to both native and non-native language in 7-month-old infants predicted the vocabulary acquisition and expressive language (sentence complexity and the length of utterances) at 2 years (Kuhl et al., 2008; Kuhl et al., 2005; Tsao et al., 2004). Although there is a large inter-individual variation in infant pre-linguistic

development, this variation remains relatively stable throughout the early toddlerhood (Määttä, Laakso, Tolvanen, Westerholm, & Aro, 2016). Up to 53.3% of childhood language outcome variance by 8 years can be explained by pre-linguistic communication growth in the first 6 to 18 months of life (Määttä et al., 2016). Earlier studies have supported these findings in categorising 3, 5 and 8 years old children into different language development groups depending on their neonatal Event Related Potential (ERP) responses to speech and non-speech sounds (Molfese, 2000; Molfese & Molfese, 1985, 1997). Results showed better discrimination ability in the older age groups, with an accuracy of 89% in grouping 5-year-old children into two verbal performance groups; and 81% in determining language status at 8 years as defined by poor reader, dyslexia, or control language development group.

Furthermore, neuroimaging studies from as early as infancy support the idea of applying contemporary neuroimaging tools to observe typical and atypical language brain development (Blasi et al., 2015; Grossmann, Oberecker, Koch, & Friederici, 2010; Lloyd-Fox, Blasi, et al., 2013; Lloyd-Fox, Blasi, Mercure, Elwell, & Johnson, 2012). Both functional Magnetic Resonance Imaging (fMRI) and functional Near-Infrared Spectroscopy (fNIRS) studies showed identifiable dissimilarity in human voice sensitivity between high-risk autism infants and typically developing controls (Blasi et al., 2015; Lloyd-Fox, Blasi, et al., 2013). Four to six months old infants at high risk of autism had significantly reduced sensitivity when they listened to human vocalisations compared with non-voice sounds (Blasi et al., 2015; Lloyd-Fox, Blasi, et al., 2013). Cantiani studied 6-month-old Italian infants and found their auditory processing discrimination ability predicted language development at the age of 20 months. Furthermore, their ERP components could separate high and low-risk language learning impairment groups as expected (Cantiani et al., 2016).

This evidence emphasised the importance and possibility of early assessment of infant vocal and language development.

1.1.2 The importance of early social-emotional development

A large amount of our communication is conveyed non-verbally through gestures, facial expressions and vocal emotional information - emotional prosody (Burgoon, Guerrero, & Floyd, 2016). Despite the importance of language development, the ability to grasp the emotional information conveyed in another's speech is equally vital for successful communication (Hargrove, 1997).

Emotional prosody, similar to facial emotional expression, is a core element in vocalisation; the changes in intonations convey emotional information across cultural (Hargrove, 1997; Scherer, Banse, & Wallbott, 2001). Studies using emotional prosody have found that difficulty in distinguishing emotional prosody has been associated with language deficits and mental well-being in both adults and children (Baum & Nowicki, 1998; Carton et al., 1999; Nowicki & Duke, 1992). In Carton's study, even after controlling for depression, a correlation was found between the ability to decode nonverbal emotional sounds and the well-being of relationship (Carton et al., 1999). Children with low accuracy in interpreting tones of voices were less popular, had less control for external behaviour and performed worse academically among their peers (Nowicki & Duke, 1992). Chronaki also reported that children's hyperactivity and conduct problems were negatively associated with the accuracy with which they distinguished vocal emotional expressions in a speech discrimination task (Chronaki, Garner, et al., 2015). This study has highlighted the need for the application of vocal emotional expression tasks in the assessment of children with behaviour problems. In a subsequent ERP

study, attention deficit hyperactivity disorder (ADHD)-diagnosed children also showed altered neural responses, in terms of N1 (or N100) component amplitude compared with healthy control children while listening to non-linguistic vocal anger prosody (Chronaki, Benikos, Fairchild, & Sonuga-Barke, 2015).

Atypical social-cognition in children and adults may have its roots in infancy, and social-emotional atypicalities may act as important early markers of potential developmental problems in children (Blasi et al., 2015; Chronaki, Benikos, et al., 2015; Chronaki, Garner, et al., 2015; Chronaki, Wigelsworth, Pell, & Kotz, 2018). For example, Blasi et al.'s study found that in addition to decreased sensitivity to human vocalisation, infants at risk of autism also exhibited reduced sensitivity to emotional vocalisations displayed in listening to the contrast of sad and neutral vocalisations (Blasi et al., 2015).

Social-emotional development is a key element in communication and mental well-being; however, compared with vocal and language development, vocal emotion is less studied, especially in early infancy.

1.2 Early language and social-emotional developmental trajectory

The cognitive skills needed for language and social-emotional cognition develop dramatically in infants' first year of life and equip infants with the skills they will need to navigate healthy social lives.

1.2.1 Early vocal and language development

A fetus can hear sounds from around 22 weeks of gestational age (Hepper & Shahidullah, 1994a, 1994b; Moore, 2002) and can respond behaviourally to sounds

at around 23-35 weeks of pregnancy (Hepper & Shahidullah, 1994b). After the maturation of the auditory system at around 38 weeks gestation, a fetus shows more interest in the own mother's voice than other, strange female's voices (Kisilevsky et al., 2003).

The infant receptive language initiates with the interest in the mother's voice and generally to human voices, which becomes evident in infancy (Blasi et al., 2015; Blasi et al., 2011; Dehaene-Lambertz et al., 2010; Grossmann, Oberecker, et al., 2010; Lloyd-Fox et al., 2012). Comprehension of words develops from around 8 to 10 months of age (Bates et al., 1992; Fenson et al., 1994). At this point, infants can understand, and may respond to, contextual words (such as the inhibitory word "no"), name-calling and play or communication routines (e.g. "peekaboo", waving bye-bye). Word comprehension development does not follow a linear trend by the end of infancy. According to parent report, a 12-month-old infant can comprehend an average of 86 words, which nearly doubles to 156 two months later; this 'vocabulary burst' happens after 12 months of age (Fenson et al., 1994).

The infant's expressive language develops when the infant starts babbling from 6 to 8 months of age after several months of listening practice (receptive language development); imitation and word-like sounds begin around 10 months of age (Bates et al., 1992; Fenson et al., 1994). Although there is much individual variation in the development of word comprehension and production, word production develops slowly, with fewer than 10 words produced at the age of 12 months (Bates, Bretherton, & Snyder, 1991). Infants and toddlers generally comprehend 50 words before producing 10 words (Benedict, 1979). A child may be able to differentiate between two words or even recognise a sentence before the accurate production of a word or sentence. Similar to the burst of word comprehension, infants reach a

detectable word production ‘burst’ after the age of 13 months (Benedict, 1979; Fenson et al., 1994). By the age of 30 months, vocabulary growth can be tenfold compared to the previous period. There is a great individual variation in this word boost and the individual differences may carry over through the early years (Määttä et al., 2016). This pattern of early human language acquisition is a universal phenomenon across a variety of cultures and languages.

1.2.2 Early social-emotional development

As with early language development, emotional development moves from an understanding of emotions to an ability to respond behaviourally to others’ emotional expression; this process is part of early social-emotional development (Rosenblum, Dayton, & Muzik, 2009). An infant is not only born with the ability to detect language, but also to detect and differentiate emotions (Mastropieri & Turkewitz, 1999). Impressive changes take place in the first year of life, as infants sharpen their skills in order that they can ‘fit in’ to the communicating world.

Infants cry from birth as a sign of emotional expression, with the ability to smile developing around 6-8 weeks post birth and laugh from around 3 months (Rosenblum et al., 2009). Complex emotions, such as jealousy, shame and guilt appear in late infancy or toddlerhood (Barrett, Zahn-Waxler, & Cole, 1993; Hart, Carrington, Tronick, & Carroll, 2004). Social-emotional development is also greatly affected by the infant’s temperament and other social and cultural factors (such as religions, values and perspectives, (Cole et al., 2006; Kitayama et al., 1995). Infants have less capacity to regulate their own emotions; they gradually become more capable of regulating their emotions through a combination of internal and external ‘learning’ processes. Perception and neurophysiological development (especially the

development of frontal lobes), caregivers' assistance, and regulation behaviours all contribute to the development of emotion regulation (Thompson, 1991).

An older infant not only understands the caregiver's emotional cues, her (or his) behaviours are also, in turn, regulated by these signals. From around 10 months after birth, the infant's understanding of emotions in the communication motivates behavioural reactions (Barrett, Campos, & Emde, 1996; Hornik, Risenhoover, & Gunnar, 1987; Mumme, Fernald, & Herrera, 1996; Sorce, Emde, Campos, & Klinnert, 1985). This social-emotional development also follows a selective timeline, beginning with parents and then spreading to strangers (Barrett et al., 1996; Walker-Andrews, Krogh-Jespersen, Mayhew, & Coffield, 2011).

In summary, infants' skills in language and social-emotion are developing at a fast pace in the first year of life. Despite a great number of developmental changes taking place, there is also a wide variation in this development across typically developing infants (Bates et al., 1992; Cole et al., 2006; Fenson et al., 1994; Kitayama et al., 1995; Määttä et al., 2016). This individual variation makes the language and social-emotional development hard to evaluate in infants and children, which may lead to a delay in diagnosis (Palfrey et al., 1987; Pinto-Martin et al., 2005).

A recent study reported that language and social-emotional developmental delays/deficits are identifiable at a group level in emergent developmental disorders at around 9 to 13 months of age from behavioural measures (Wan, Green, & Scott, 2018). With contemporary psychophysiological methods (e.g. Electroencephalogram-EEG, fMRI, and fNIRS), identifiable differences in neural responses to social stimuli were found between typically developing infants and infants at risk of developmental disorder from as young as 4 months (Blasi et al.,

2015; Elsabbagh et al., 2012; Lloyd-Fox, Blasi, et al., 2013). The use of human vocalisations (voice and vocal emotional sounds) as stimuli supported the identification of typical and atypical language and social-emotional brain development in infants (Blasi et al., 2015; Grossmann, Oberecker, et al., 2010; Kuhl et al., 2014; Lloyd-Fox, Blasi, et al., 2013) and children (Chronaki, Benikos, et al., 2015; Tsao et al., 2004). Infant neural responses to human voice and vocal emotional stimuli are recognised as indicators of language and social-emotional development. Here we reviewed infant development of voice and vocal emotion processing in the first year of life.

1.3 The development of voice and vocal emotional processing in infancy

The exploration of brain responses to voice and emotion processing originated from work in brain-injury patients. Studies on participants with and without brain lesions support findings of left hemisphere dominance in semantic processing by showing the left frontal lobe involvement in both word production and word recognition tasks (Benton, 1968; Posner, Petersen, Fox, & Raichle, 1988). Left temporal lobe was also found to play a role in lexico-semantic processing (Demonet et al., 1992). Patients with damage in the right hemisphere show impaired ability at recognising and understanding emotional vocalisations compared to patients with damage in the left hemisphere and healthy controls (Blonder, Bowers, & Heilman, 1991; Heilman, Scholes, & Watson, 1975). Right hemisphere damage was correlated with voice recognition, while voice discrimination was related to bi-hemispheric damage (Vanlancker, Kreiman, & Cummings, 1989).

Modern neuroimaging studies in healthy adults have offered some insights about the localisation of social stimuli processes (Adolphs, 2009). Areas associated with sensitivity to the human voice have been identified in bilateral temporal lobes when contrasted with responses to non-human sounds with somewhat stronger activation in the right hemisphere (Belin, Zatorre, Lafaille, Ahad, & Pike, 2000; Grandjean et al., 2005). Meanwhile, Fecteau and colleagues reported enhanced activation in the left prefrontal cortex to human voices, as well as to emotional compared to neutral vocalisations (Fecteau, Armony, Joanette, & Belin, 2005). Adult neural responses to happy vocalisations were found stronger in bilateral middle temporal gyrus and right inferior frontal gyrus (Johnstone, van Reekum, Oakes, & Davidson, 2006). Similar findings in bilateral temporal gyrus were also reported by Ethofer et al using both happy and angry intonations (Ethofer, Anders, Wiethoff, et al., 2006). Even though both left and right temporal lobes were activated by emotionally prosodic stimuli compared to neutral sounds, the right hemisphere showed more intense activation than the left hemisphere, especially in the middle and superior temporal gyrus (Mitchell, Elliott, Barry, Cruttenden, & Woodruff, 2003). Deeper brain regions like bilateral amygdalae have also been found to be sensitive to emotional vocalisations (Fecteau, Belin, Joanette, & Armony, 2007).

Schimmer and Kotz proposed a vocal emotional comprehension model that consists of three sub-processing elements. First, the brain undertakes sensory processing; then it integrates the auditory ‘what’ pathway, and finally the processing model rests on the evaluation of semantic processing (Schirmer & Kotz, 2006). Sensory processing is a bottom-up mechanism which evokes bilateral primary auditory sensory cortical (bilateral superior temporal gyrus) responses, while the integration of emotional auditory information works in the right anterior superior temporal sulcus as a top-

down mechanism. Higher order cognitive processes evaluate the valence and sematic properties of stimuli, and are mediated by the right inferior gyrus, the orbitofrontal cortex, and the left inferior frontal gyrus, respectively (Schirmer & Kotz, 2006).

In summary, adult neural responses to voice in contrast to non-vocal sounds are located in the bilateral temporal and frontal regions, which are also part of 'social brain' (Adolphs, 2009). Vocal emotional processing, although informed by adult brain lesion studies to be lateralised to the right hemisphere, also evidenced in bilateral superior temporal and inferior frontal regions.

1.3.1 Voice processing in infancy

The processing of voice starts alongside the development of hearing during fetal life. Infants' ability to recognise human voice develops via exposure to pre- and post-natal environmental stimuli. Voice recognition and the ability to discriminate different vocalisations develops as a precursor to speech processing (Friederici, 2005). Fetuses from 36 weeks gestation and neonates showed an increase in heart rate when they heard familiar voices, such as that of their parents', and a decrease in heart rate when exposed to unfamiliar voices, such as unknown female voices (Kisilevsky et al., 2003; Ockleford, Vince, Layton, & Reader, 1988). Fetuses' heart rate slowed when hearing low intensity speech during sleep (Groome et al., 1999). In addition, neonates and 3-month-old infants responded to speech with higher-amplitude sucking behaviour and longer looking time compared to other sounds. (Decasper & Fifer, 1980; Shultz & Vouloumanos, 2010; Vouloumanos, Hauser, Werker, & Martin, 2010; Vouloumanos & Werker, 2007).

In addition to the behavioural responses, the neonatal brain, especially the frontal and temporal regions, have the ability to differentiate syllable structures in human

language (Gervain, Macagno, Cogoi, Pena, & Mehler, 2008; Gómez et al., 2014; Pena et al., 2003). Neonates 'prefer' ABB tri-syllables, evidenced by increased brain haemodynamic responses compared to ABC. No differences were seen in infants hearing ABA tri-syllables compared with ABC, indicating the recognition of such speech grammar (Gervain et al., 2008). Preference for syllable structure was also found in Gomez et al.'s new-born study (Gómez et al., 2014). Moreover, neonates significantly responded to 'forward speech' as the recognition of normal speech and showed reduced responses to atypical speech (i.e. backward speech) (Pena et al., 2003). These studies reported left hemisphere superiority to speech stimuli.

From birth to around six months, infants can discriminate between phonemes (Kuhl, 2004). Indexed by head turn, 7-month-old infants have high accuracy in discriminating phonemes in both native and foreign languages (Kuhl, 2004). Non-native language discrimination skills gradually disappeared in the age group of 8-12 months, with the accuracy of non-native language discrimination remarkably declined (Werker & Tees, 2002). This is because infancy phonetic perception and discrimination capacity between native and foreign language is naturally shaped by the particular language environment in which the infant is predominantly living (Kuhl et al., 2014; Werker & Tees, 2002). The non-native language perception fades from 7 months onward till the end of the first year of life (Kuhl et al., 2014; Werker & Tees, 2002). Both head turn behaviours and brain activations have shown that the infant's brain is gradually forming adult-like language perception responses (Kuhl et al., 2008; Kuhl et al., 2014; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Rivera-Gaxiola, Silva-Pereyra, & Kuhl, 2005; Werker & Tees, 2002). An ERP study with 7- and 11-months-old infants showed that both native and non-native language evoked significant negative potentials at 250-550 ms and positive going potentials at

150-250 ms (Rivera-Gaxiola et al., 2005). There was a clear growth in native language perception in 11-month-old infants compared to 7-month-old infants (Rivera-Gaxiola et al., 2005). A MEG study reported that both superior temporal and inferior frontal regions were activated when infants aged 7, 11 and 12 months, as well as adults, listened to speech sounds. (Kuhl et al., 2014). Despite activations in the same brain regions to native and non-native language across age groups, the pattern of activation varied at different ages. The magnitude of activation to both native and non-native language was equal in frontal and temporal regions in 7-month-old infants, while greater activation was found in temporal regions for native language and greater responses in frontal regions to non-native language at 11 to 12 months of age and in adults (Kuhl et al., 2014). No hemispheric dominance was reported in infants' responses to native and non-native language phonemes. While, Dehaene-Lambertz and colleagues' fMRI study found that the adult-like auditory structural asymmetry was already evident in 2-month-old infants, with the left temporal region more strongly activated by speech sounds than by music sounds (Dehaene-Lambertz et al., 2010). Minagawa-Kawai et al. showed 4-month-old infants' neural responses to speech (native and non-native) were lateralised to the left temporal cortex, and non-speech emotional vocalisations were localised to the right temporal cortex (Minagawa-Kawai, van der Lely, et al., 2011).

Lloyd-Fox and colleagues provided evidence of neural responses to auditory social stimuli from as early as 2 months of age and up to 24 months (Lloyd-Fox et al., 2017; Lloyd-Fox et al., 2012; Lloyd-Fox et al., 2014). They compared neural responses evoked by human non-speech vocalisations (i.e. laugh, cry, yawn and cough) with those elicited by non-vocal environmental sounds (i.e. running water, rattles and bells). Their findings suggest that infants aged 4 months and younger had

not developed distinct responses to human vocalisations vs non-vocal environmental sounds. Infants aged between 0 to 2 months display greater oxygenated haemoglobin concentration changes to non-vocal environmental sounds than to human vocalisations in posterior temporal regions in fNIRS studies (Lloyd-Fox et al., 2017; Lloyd-Fox, Blasi, et al., 2013). By 4-8 months, they reported distinct left and right temporal responses both to human vocalisations and to non-human sounds and suggested this may represent a transition period (Lloyd-Fox et al., 2017; Lloyd-Fox et al., 2012). Lloyd-Fox and colleagues' findings further suggest that infants from the age of 9 months have robust and stable neural responses to human vocalisations compared to non-human sounds located in the right temporal cortex (Lloyd-Fox et al., 2017). fMRI studies support Lloyds-Fox et al.'s developmental time course for the processing of auditory social stimuli. During passive listening, 3- to 7- month-old infants' superior and middle temporal gyri and superior and middle frontal gyri responded to human social vocalisations, and the left superior temporal gyrus responded to non-human sounds (Blasi et al., 2015; Blasi et al., 2011). Grossmann et al.'s fNIRS study, which compared infant brain activation to human vocalisations (neutral speech and non-speech) versus non-vocal sounds reported overlapping findings: in 7-, but not 4-month-old infants, bilateral superior temporal cortical activation was significantly greater to human vocalisations than to non-vocal sounds, suggesting that sensitivity to the human voice emerges sometime between 4 and 7 months of age (Grossmann, Oberecker, et al., 2010). The greater fronto-temporal activation to human vocalisation than non-vocal sounds is also consistent with child and adult voice recognition (Belin & Zatorre, 2000; Belin et al., 2000; Rogier, Roux, Belin, Bonnet-Brilhault, & Bruneau, 2010).

In summary, from birth, infants behaviourally distinguish the mother's voice and unfamiliar voices, and show an interest to speech over other sounds. In addition, infants' ability to discriminate phonemes in native and non-native languages is equally accurate before the age of 6 months. Even though neural responses representing the preference to human vocalisations over non-vocal sounds are relatively vague before the age of 9 months, studies involving stimuli with elements of speech and emotion obtained overlap but different results. It is unclear when infants develop specialised neural responses to human vocalisations, and whether speech and emotion influence the neural responses.

1.3.2 Vocal emotional processing in infancy

Human vocal communication typically includes linguistic meaning (semantics) as well as emotional information. Emotional vocalisations carry crucial information for social communication and these elements are believed to stimulate increased awareness and attention in infants (Blasi et al., 2015; Blasi et al., 2011; Grossmann, Oberecker, et al., 2010; Grossmann, Striano, & Friederici, 2005; Missana, Altvater-Mackensen, & Grossmann, 2017), children (Chronaki, Benikos, et al., 2015; Chronaki, Garner, et al., 2015; Rogier et al., 2010) and adults (Ethofer, Anders, Wiethoff, et al., 2006; Fecteau et al., 2007; Grandjean et al., 2005; Zhang, Zhou, & Yuan, 2018). Before infants understand semantic meaning of words or sentences, they are sensitive to different intonations in speech, and they process, even respond to emotional vocalisations in the first year of life (Fernald, 1993; Hornik et al., 1987; Mumme et al., 1996; Sorce et al., 1985).

Infants are born with the ability to distinguish different emotional vocalisations. New-borns have higher eye opening scores when they hear happy vocalisations,

especially in the native language, compared with angry, sad and neutral ones (Mastropieri & Turkewitz, 1999). One-month-old infants could identify their own mother's voice with intonation but not their mother's speech without prosodic aspects (Mehler, Bertoncini, Barriere, & Jassikgerschenfeld, 1978), suggesting infants' preference for prosodic features in speech. From 3 months, infants can discriminate between emotions vocally and habituate when shown the same emotional vocalisation repeatedly (Walker-Andrews & Grolnick, 1983). However, discrimination effects appear to be limited to a change from sad to happy vocalisations; 3 to 5 months, this ability starts to stabilise and become more robust (Walker-Andrews & Grolnick, 1983).

From the age of 5 months, infants respond differently to positive and negative infant-directed speech, irrespective of language. Their positive affect scores were higher when listening to positive vocalisations and negative affect scores were higher when hearing negative vocalisations (Fernald, 1993). Six-month-old infants seem to prefer positive affect (happy) speech, they show longer looking-time to positive compared to neutral or negative (sad) speech (Singh, Morgan, & Best, 2002).

Emotional expressions have an important influence on an infant's subsequent capacity to regulate their affect and behaviour around 12 months of age. For instance, positive emotions, such as joy or interest, encourage 12-month-old infants to cross the visual cliff; while negative emotions, such as anger or fear reduce the cross behaviour (Sorce et al., 1985). Similarly, evidence from Mumme et al suggest that mothers' fearful vocalisations significantly affected 12-month-old infants' toy exploration behaviour and affect (Mumme et al., 1996). It has been suggested that negative expression (such as fear, disgust) has the most influence on affect and behavioural regulation in 12-month-old infants, compared to positive (such as happy,

surprise) and neutral expressions (Hornik et al., 1987). In addition, even without facial expressions, fearful voices were sufficient to stimulate infants' corresponding behaviours, which include longer looking time to the caregiver, less exploration of a toy and more negative affect expressions (Hornik et al., 1987; Mumme et al., 1996).

These studies demonstrate how 12-month-old infants are not only capable of distinguishing different emotional expressions, but can also regulate their behaviours by reference to the emotional information from others. Moreover, some researchers argue that infants show a 'bias' to negative emotions (Vaish, Grossmann, & Woodward, 2008). This may suggest that infants' learning process starts from a point of needing to know how to avoid dangers, which is commonly associated with negative emotions (Peeters & Czapinski, 1990).

Results from infant neuroimaging studies are in line with behavioural findings evidencing infants' distinct patterns of brain activation to emotional compared to neutral vocalisations (Blasi et al., 2015; Blasi et al., 2011; Grossmann, Oberecker, et al., 2010; Grossmann et al., 2005; Missana et al., 2017; Zhang, Zhou, Hou, Cui, & Zhou, 2017).

In neonates, increased blood oxygenated haemoglobin concentration changes have been observed in temporal and parietal regions in relation to passive listening of emotional vocalisations (Zhang et al., 2017). Significantly larger hemodynamic responses were reported for emotional vocalisations (fearful, angry and happy) than for neutral vocalisations in the right temporal region; and larger responses for fearful, than happy and neutral, were located in the right parietal region (Zhang et al., 2017).

Based on ERP methodology, 7-month-old infants showed stronger positive slow potentials at around 500 ms over bilateral temporal areas following angry and happy vocalisations compared to neutral vocalisations (Grossmann et al., 2005). By contrast, words with an angry prosody elicited enhanced negative ERP responses at around 450 ms over frontal-central regions compared to words with happy or neutral prosody (Grossmann et al., 2005). In the same age group, oxygenated blood flow increased in superior temporal and inferior frontal regions when infants listened to angry and happy voices compared to neutral sounds (Grossmann, Oberecker, et al., 2010). Specifically, angry sounds evoked significantly larger responses than happy and neutral vocalisations in the right superior temporal region; happy generated significantly greater activation in the right frontal region (Grossmann, Oberecker, et al., 2010). The right superior temporal region, in reaction to emotional prosodies, also belonged to the voice sensitive activation pattern in 7-month-old infants (Grossmann, Oberecker, et al., 2010). This was the first study which reported an overlap in patterns of brain activation to voice and vocal emotional processing in infants. Following this, a recent study reported 8-month-old infants' ERPs to peers' laugh (positive) and cry (negative) sounds compared to adult hummed speech (neutral) (Missana et al., 2017). They found increased negativity in temporal regions to crying sounds than to laughing and neutral sounds, and enhanced positivity in central regions to laughing sounds than to crying and neutral sounds (Missana et al., 2017).

Blasi et al. (2011) presented 3- to 7-month-old infants with three types of adult non-speech vocalisations (neutral, positive and negative) while asleep in the MRI scanner (Blasi et al., 2011). They reported stronger activation to negative (sad) than to neutral vocalisations in the insula with no differences observed between positive

(happy) and neutral vocalisations, suggesting early functional specialisation for processing negative emotions (Blasi et al., 2011). More recent fMRI research has also reported stronger sensitivity to sad vocalisations in the right fusiform gyrus and left hippocampus in infants aged 4 to 7 months (Blasi et al., 2015). However, in this study, activation patterns for vocal emotional processing did not overlap with those elicited for voice processing (Blasi et al., 2015; Blasi et al., 2011).

The brain activation patterns to vocal emotional processing are somewhat inconsistent in infant studies. It is less clear whether vocal emotional processing shares the same neural basis as voice processing, because only few studies have compared these processing mechanisms. There are no consistent findings of brain localisations of infant neural responses to voice and emotional vocalisations, but the activation regions in the brain are within the temporal-fronto network. As suggested by the development of 'social brain', although less developed and specialised, infant 'social brain' shows similar neural activations to social stimuli as in the adult brain (Grossmann & Johnson, 2007). Considering the speed of development in infancy, it is crucial for us to be able to pinpoint the activation patterns at each age point associated with the different elements of voice and vocal emotional processing. Longitudinal studies that allow observations in the developmental process at different ages may help to understand infant voice and vocal emotion developmental trajectory and indicate the best observation time period for obtaining robust, more consistent findings across studies.

1.4 Maternal sensitivity and language and social-emotional development in infancy

1.4.1 Maternal sensitivity

The primary caregiver, usually the mother, provides the main socio-emotional environment in which the infant's brain develops functionality. Although mother-infant interaction has bidirectional qualities (the infant's behaviour affecting the mother's, and vice versa), a positive relationship relies heavily on the mother's sensitive responsiveness behaviour and tuning towards her infant (Ainsworth, 1978). In the past few decades, research has indexed the quality of maternal caregiving behaviour through the concept of 'maternal sensitivity'.

Sensitive maternal behaviour refers to the mother's perceptiveness of her infant's cues, signals and behaviours and to provide, in response to these, prompt and appropriate responses. Sensitive caregiving involves a mother being ready to accommodate the infant's changing developmental needs. Sensitivity also involves giving the infant the autonomy to explore the world, to create new things (Landry, Smith, MillerLoncar, & Swank, 1997) and provide positive and encouraging narratives in the interactions with infants (Ainsworth, 1978). In turn, the infant provides active responses in the mutual interaction which may reinforce or modify the mother's caregiving behaviour (Eisenberg et al., 2010; Kivijarvi et al., 2001). Maternal sensitivity is embedded in the daily interaction with the infant (such as in a play session), when the mother's behaviour may be quantified and qualified with standardised scales (such as Manchester Assessment of Caregiver-Infant Interaction, (Wan et al., 2012, 2013)).

By contrast, maternal directiveness conveys explicit or implicit expectations about the infant's behaviour, including whether s/he attends to or does something, or prohibits an action (Guzell & Vernon-Feagans, 2004). While this caregiver strategy may help ensure safety and provide clear guides for behaviour, directiveness will typically involve vocal and/or behavioural demands, intrusions and critical utterances, which may take explicitly emotionally negative forms, or may be positively valenced, such as repeated 'overbright' name calls as attentional bids and sharp 'playful' vocalisations ("ah ah ah") signalling disapproval. Thus, being the recipient of high directiveness likely entails receiving high emotional vocal input that is self-relevant (whether positive or negative, as with infants who experience high maternal sensitivity). However, negative emotion may be more often encountered and appraised as a guide to acceptable behaviour, giving rise to a neural bias towards the processing of negative prosody. One study to date has attempted to link maternal behaviour (intrusiveness) with infant neural vocal response: no significant linear relationship was found in infants with high and low risk of autism (Blasi et al., 2015).

1.4.2 Maternal care behaviour and early infant language and social-emotional development

As mentioned above, before birth, an infant is familiar with the mother's voice during fetal life. Abundant evidence from studies measuring infants' heart rates, sucking and looking behaviours display this familiarity and preference (Decasper & Fifer, 1980; Kisilevsky et al., 2003; Ockleford et al., 1988). A mother, as the main caregiver, scaffolds the infant's social and language development and has an important, if not central, influence on this process.

At 3 to 5 months, an infant begins to learn that vocalisations are meaningful through both verbal and non-verbal interactions; and a child learns the relevance and impact of the own vocalisations and those of the mother. A certain type of vocalisation represents a specific goal, such as soothing sounds when the infant is distressed, or encouraging sounds during play (Goldstein, Schwade, & Bornstein, 2009; Tronick, Als, Adamson, Wise, & Brazelton, 1979). A responsive mother, who is sensitive to the infant's developmental needs, can provide regular vocal and verbal responses in a timely manner. For example, a mother's imitation behaviour stimulates the infant's vocalisations and smiles (Field, Guy, & Umbel, 1985). The mother's responsive speech and infant-directed speech conveys the information of patience, interest and the willingness to facilitate reciprocal interaction. This encourages the infant to express in the 'conversation', which supports building security and competence in mastering language (Baumwell, TamisLeMonda, & Bornstein, 1997; Leigh, Nievar, & Nathans, 2011; Olson, Bayles, & Bates, 1986).

In the first year of life, verbal and non-verbal communicative skills develop rapidly, and the mother's sensitivity during reciprocal interaction offers not only the scaffold for the infant's future language skills, but also for their social and communicative development more generally (Landry et al., 1997; Sylvestre & Merette, 2010). Tamis-LeMonda and colleagues' small longitudinal study of 40 healthy mother-child pairs reported a positive relationship between maternal responsiveness and children's expressive language development in the first two years. Maternal responsiveness (including mothers' verbalisations, vocalisations and the engagement in the interactions with the infant) predicted the early emergence of children's milestones in imitations, first words, first 50 words, combinatorial speech and speech about the past at 9 months and particularly at 13 months (Tamis-LeMonda, Bornstein, &

Baumwell, 2001). Similar findings from Leigh et al. showed that maternal sensitivity has a positive influence and predicted later child language development at the ages of 2 and 3 years (Leigh et al., 2011). At the vocabulary acquisition stage in late infancy, it was a mother's responsive speech (verbal stimulation) rather than maternal vocabulary that predicted the infant's later vocabulary (Bornstein, Tamis-LeMonda, & Haynes, 1999). A larger study of full-term (N = 112) and very low birth weight infant groups (N = 187, including high- and low-risk subgroups) found that faster acquisition of cognitive-language and social initiative behaviours were associated with mothers' sensitive responsiveness to maintain infants' attention and interests in both groups. This correlation was even stronger in the high-risk children (Landry et al., 1997) and was consistent with earlier findings (Ruddy & Bornstein, 1982). In contrast, mothers' control and restrictive behaviour has been negatively associated with children's cognitive and language development (Landry et al., 1997).

According to Interactive Specialisation theory (Johnson, Grossmann, & Kadosh, 2009), the infant brain undergoes a "fine-tuning" process, dependent on experience, into specialised functions. Maternal caregiving provides one of the most important social-communication environments to the infant and may have an influence on the infant language and social-emotional neural development. However, little is known about how or if poor caregiving affects the neural networks that underlie healthy language and social-emotional development in the infant and growing child. Limited research has focused on the association between maternal sensitivity and infant outcomes. An EEG study reported an association between maternal positive affect during the interaction with the 5-month-old infant and the later frontal resting EEG power in the infant at 10 and 24 months of age (Bernier, Calkins, & Bell, 2016). This finding suggests that warm maternal behaviour may have an effect on the

development of the infant's brain structure and function (Bernier et al., 2016). However, another study reported an non-significant correlation between maternal positive affect and attentional cortical activity related EEG power, although maternal positive affect was positively associated with infants' attention behaviour at 10 months (Swingler, Perry, Calkins, & Bell, 2017). Moreover, Blasi et al, using fMRI, attempted to link maternal behaviour (intrusiveness) with 3- to 7-month-old infant neural vocal response in infants with high and low risk of autism, and found no significant linear relationship (Blasi et al., 2015).

Previous studies call for evidence of the role of maternal behaviour in infant and child language and social-emotional neural development. In addition, there remains a lack of longitudinal, objective psychophysiological evidence linking early parenting sensitivity and mothers' affective-cognitive styles to early child language acquisition.

1.5 The application of fNIRS in infant cognitive neuroscience

Jöbsis found that the near infrared light could transmit the bone and skin in 1977, after which, near infrared light showed the potential as the non-invasive measurement in observing brain responses (Jobsis, 1977, 1999). Similar to fMRI, fNIRS indirectly measures neural responses indicated by haemodynamic changes in the blood. Near infrared light travels like a banana-shaped path from a source, passes through the scalp, skull, cerebrospinal fluid, grey matter, white matter, and is finally collected by a detector (Gervain et al., 2011; Lloyd-Fox, Blasi, & Elwell, 2010; Villringer & Chance, 1997).

fNIRS is a relatively new neuroscience tool compared to EEG and fMRI, with the advantage of portability, low cost, quietness, and relatively less requirement of head stabilisation, it has been widely developed in functional brain imaging studies. Studies have proved fNIRS as an effective tool in observing infant cognitive neural development, such as infant neural processing of language, action, facial expression and voice sound (Gervain et al., 2008; Grossmann, Lloyd-Fox, & Johnson, 2013; Grossmann, Parise, & Friederici, 2010; Lloyd-Fox et al., 2017; Lloyd-Fox et al., 2012; Lloyd-Fox et al., 2014; Lloyd-Fox, Wu, Richards, Elwell, & Johnson, 2013; Minagawa-Kawai, van der Lely, et al., 2011; Pena et al., 2003). These studies provided valuable pioneering work in infant fNIRS.

Although fNIRS does not have a strict requirement of body stabilisation, the need to reduce artefacts in infant fNIRS data becomes the priority in observing awake infant haemodynamic responses. Even in well-designed infant fNIRS studies, there were large amounts of data rejection with an average exclusion rate of around 40 percent (Lloyd-Fox et al., 2010). Researchers have applied various algorithms to improve signal-to-noise ratio in fNIRS data (Fox, Wagner, Shrock, Tager-Flusberg, & Nelson, 2013; Hocke, Oni, Duszynski, & Corrigan, 2018; Tak & Ye, 2014; Wilcox, Bortfeld, Woods, Wruck, & Boas, 2005), however, there is no golden rule that applicable to all fNIRS data (Hocke et al., 2018; Pinti et al., 2018).

Independent Component Analysis (ICA) has been known as a powerful tool in identifying artefacts and sources in EEG and fMRI data. This data-driven approach does not involve a prior information/model, and is proficient in solving ‘cocktail party’ problems - separating sounds of independent speakers from collections of mixed sounds. In fNIRS data, each recording site (channel) collected a mixture of

signals and noises, which are independent of each other and may be effectively identified by ICA algorithm (Fox et al., 2013; Yuan, 2013).

1.6 Chapter summary

(a) Language and social-emotional developments are key elements in effective communication and mental well-being. The delay and impairment in language and social-emotion development are found associated with developmental disorders and mental health problems in childhood, adolescence and adulthood. Atypically developed language and social-emotion cognition in children and adults may have their roots established in infancy.

Language and social-emotion cognition develop at a fast speed from birth throughout the first year of life, research has increasingly moved towards the infant's early language and social-emotional development to find early indicators of risk. However, there is a wide variation in normal language and social-emotional development, which leads to the difficulty in the accurate early diagnosis with behavioural assessments.

(b) Contemporary psychophysiological methods (i.e. EEG, fMRI, fNIRS) support the view that the infant brain has formed measurable responses from a young age. Identifiable differences in neural responses to social stimuli were found between typically developing infants and infants at risk of developmental disorder from as young as 4 months.

The use of human vocalisations (voice and vocal emotional sounds) as stimuli supported the identification of typical and atypical language and social-emotional

brain development in infants and children. Infant neural responses to human voice and vocal emotional stimuli are recognised as indicators of language and social-emotional development.

(c) Even though infants showed neural responses representing the preference to human vocalisations over non-vocal sounds, studies involving stimuli with elements of speech and emotion obtained overlap but different results. It is unclear when infants develop specialised neural responses to human vocalisations, and whether speech and emotion influences the neural responses.

The infant brain activation patterns in processing emotional vocalisations are somewhat inconsistent across studies. It is less clear whether the processing of voice and emotional vocalisations share the same neural basis, because only few studies have compared these processing mechanisms.

(d) There is limited knowledge of the developmental trajectory of human vocalisations processing in infants. Considering the speed of development in infancy, it is crucial for us to be able to pinpoint the activation patterns at each age point associated with human vocalisation processing. Longitudinal studies that allow observations in the developmental process at different ages may help to understand infant vocal emotion developmental trajectory and indicate the best observation time period for obtaining robust, more consistent findings across studies.

(e) As the primary carers, the quality of maternal care provides so-called scaffolding for the language and social-emotional development of the infants. Mothers who have difficulties in providing sensitive caregiving to their infants may negatively influence the language and social-emotional development of their infant. Despite behavioural evidence showing the association between maternal sensitive caregiving and the

infant language and social-emotional development, there is a lack of neural mechanism evidence which underlies how caregiving behaviour might mediate or influence infant language and social-emotional brain development.

(f) fNIRS has advantages over EEG and fMRI in observing infant auditory neural responses, however, fNIRS is hindered by large data loss and a lack of widely accepted analysis method. ICA may an effective tool to identify signals from collected fNIRS data.

1.7 The overall study aims and objectives

1.7.1 Aims:

(a) to develop and pilot a fNIRS protocol that could be applied acceptably and feasibly in a clinical setting in order potentially to monitor and evaluate the neurological underpinnings of early language and social-emotional development;

(b) to track longitudinal changes in neural correlates of human voice and emotional vocalisations processing in typically developing infants at the age of 6, 9 and 12 months;

(c) to explore the role of early maternal caregiving behaviour as a possible mediator in the healthy development of voice and vocal emotion processing in infants.

1.7.2 Objectives:

(i) To set up an fNIRS system and develop a protocol for reliable use in a routine clinical setting with infants aged between 6 and 12 months;

(ii) To pilot the feasibility and acceptability of testing longitudinally on 3 occasions in 40 healthy infants;

(iii) To measure brain responses to human voice and vocal emotional stimuli longitudinally between and within infants;

(iv) To develop a method for improving the signal-to-noise ratio in optical imaging data with both standardised method and state-of-art methodology such as Independent Components Analysis (ICA);

(v) To examine the evidence for laterality or regional specificity of the emerging responses to voice and vocal emotion processing in infant brain;

(vi) To explore the correlation between maternal sensitivity and infant voice processing fNIRS outcomes.

1.8 Overview of thesis

This PhD thesis is presented in the journal format with five papers. At the time of writing, two papers are under review (Chapter 5 and Chapter 6), and the other three papers are under preparation for submission (Chapter 3, Chapter 4, and Chapter 7). Chapter 1 in this thesis reviewed the literature on early infant language and social-emotional development as well as providing an overview of previous research on infant voice and vocal emotion processing. Furthermore, research on the mother's caregiving role in language and social-emotion development is also reviewed. Chapter 2 describes methodological considerations in the longitudinal assessment of infant voice and vocal emotion neural processing. Chapter 3 (paper 1) considers fNIRS data analysis methods which are compared between standardised analysis and Independent Component Analysis (ICA) implanted emotion-sensitive components identification procedures. Chapter 4 (paper 2) discusses the acceptability and

feasibility of the longitudinal use of fNIRS in infants between the ages of 6 to 12 months. Chapter 5 shows 6-month-old infants' neural responses to emotional vocalisations and the exploration of the associations with maternal caregiving behaviours (sensitive responsiveness and directiveness). Longitudinal infant neural responses to emotional vocalisations and to human voice versus non-vocal sounds at the age of 6, 9 and 12 were reported in Chapter 6 and Chapter 7 respectively. Chapter 8 presents a summary of findings, discusses research implications, strengths and limitations, as well as considering suggestions for future work.

Chapter 2 Longitudinal infant voice and vocal emotion fNIRS study methodology

2.1 Longitudinal infant voice and vocal emotion brain development study design

This PhD study involved developing an acceptable and reliable method to observe changing voice and vocal emotion processing in the infant brain in the first year of life. The longitudinal measurement of infant brain responses both to human voice versus non-vocal sounds, and neutral vocal versus vocal emotions (happy and angry) in the same testing frame allowed for a powerful within-subject and an exploratory between-subject design of 3 time periods at the age of 6, 9, and 12 months. A key challenge is that infants have short recording time tolerance and tend to generate a large proportion/ percentage of motion artefacts. The analysis method package was developed to improve the signal-to-noise ratio in infant neuroimaging data.

Given the central role of maternal care quality in early speech and language development, the study also comprised a cross-sectional correlational association analysis between the infant's vocal emotion-sensitive brain responses and the infant's early social experience (from the mother) which derived from observations of mother-infant interaction at 6 months.

Acceptability was tested by asking mothers for their feedback after each measurement session and by the participant retention rate over the three time points. Feasibility was based on both the practicability of undertaking the developed fNIRS paradigm in a non-research, clinical setting and an assessment of the quality of fNIRS data (i.e. valid fNIRS data) derived within this setting. In addition, in order to make sure participated infants' language was developing typically, and there was no difference in the language competence between the included and excluded infants,

their language development was monitored using the language subset of using the language subset of the Bayley-III at 6, 9 and 12 months of age.

2.2 Participants

Three National Health Service Sure Start Children's Centres in Manchester, UK, were visited to recruit healthy mother-infant pairs with the inclusion and exclusion criteria as follows:

Inclusion Criteria:

(1) Mother was aged 18 – 40, healthy and did not have mental or major physical illness;

(2) Mother was a white fluent English speaker (in order to decrease the cultural influence, as well as the influence of skin pigment blockage of the near infrared light on the study results);

(3) Mother was pregnant or the biological mother of a healthy infant aged 1 to 6 months;

(4) The infant was full-term, normal birth weight and not diagnosed with medical, sensory or developmental disorder;

(5) If the mother has an infant she was taking care of her own infant;

Exclusion Criteria:

(1) Woman has a diagnosed psychiatric illness or is on psychotropic medication;

(2) Mother with infant diagnosed with medical, sensory, or developmental disorder;

(3) Woman whose infant is not under her own care.

Participant information sheet is attached to the Appendix 1 and Appendix 2; participant consent form is attached to the Appendix 3, study advert is attached to Appendix 4 in the thesis.

In total, 42 mothers accepted the invitation to take part in the study and signed the consent forms before their infants' age of 6 months (21 girls and 21 boys). When these infants reached 6 months, 40 out of 42 mother-infant pairs participated in the assessment (2 mother-infant pairs were lost to follow-up, 1 girl and 1 boy); 39 mother-infant pairs participated the 9 months assessment (1 mother-infant pair were lost to follow-up because of the illness of the infant, 1 girl); and 38 mother-infant pairs participated in the 12 months assessment (1 mother-infant pair were lost to follow-up because of the unavailability of the mother, 1 girl) The recruitment and participation information is shown in figure 2.1, mothers' demographical information at 6 months is shown in the Appendix 5.

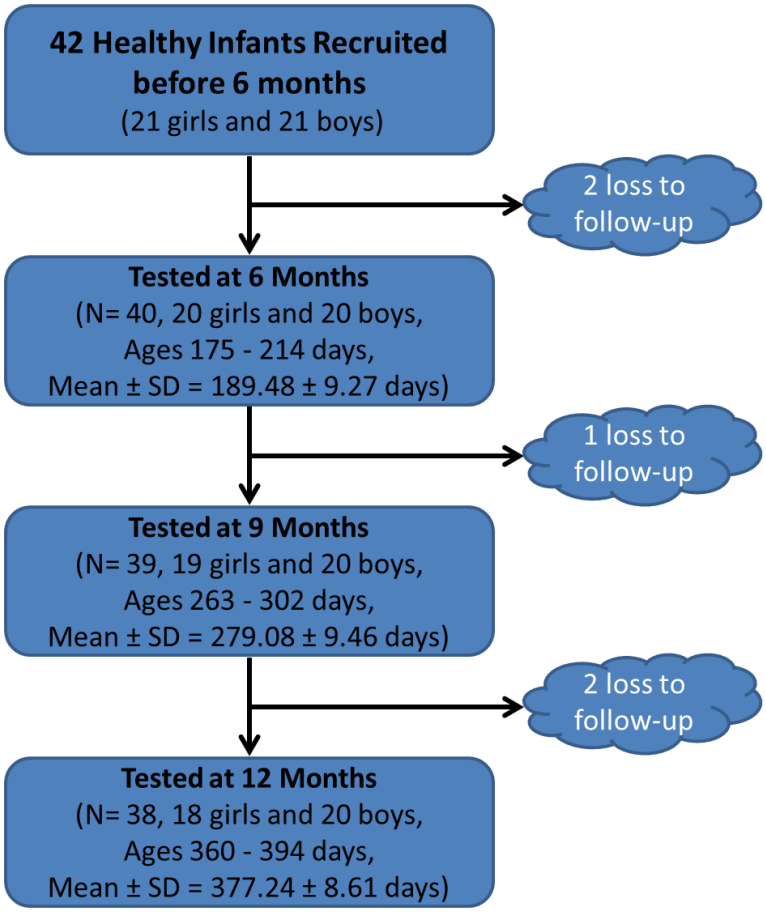


Figure 2.1 Recruitment participation flow chart.

2.3 Longitudinal infant voice and vocal emotion measurement

Four procedures were carried out in the present longitudinal study: (1) infant voice and vocal emotion processing brain responses measurement for 11 minutes; (2) infant language behaviour assessment with Bayley-III language subscale for 10-20 minutes; (3) mothers' feedback for about 5-10 minutes. (4) A six-minute videotaped mother-infant play session at 6 months only.

2.3.1 Infant voice and vocal emotion processing brain responses measurement

2.3.1.1 Experimental paradigm and procedure

Figure 2.2 presents the experimental paradigm used at all three time points. Infants wore the NIRS headband, sitting in their mothers' laps in front of a laptop during the task and listened to the acoustic stimuli. The task started with a 20 sec rest period, followed by a 5 sec trial presented through loudspeakers (SPL = 70dB). Consistent with previous research (Grossmann et al., 2010), a 5 sec silent cartoon was played along with each trial to attract infants' attention and reduce motion artefacts. After each trial, a 10 sec silent baseline along with the blurred cartoon was presented. The same video was played for both voice and non-vocal conditions. The task was presented with PsychoPy software (Peirce, 2007).

Two tasks were included in the fNIRS session: the voice versus non-vocal task; and the vocal emotion task. The paradigm and procedure were the same for both tasks and the only difference was the stimuli. At 6 months, all infants were given a break between the two tasks; at 9 and 12 months, a break was optional and most infants completed both tasks without a break. In cases where infants were fussy or reluctant to have the headgear on or to stay seated, the task was paused to allow a break and the decision to resume rested with the mother. If the infant remained unsettled on the second attempt, the researcher discussed whether to arrange another visit to do the uncompleted part(s).

The infant's behaviour during the fNIRS procedure was video recorded and later reviewed to check the number of trials in which the infant attended to the screen without large motion artefacts. Infants who did not attend for 4 or more of 8 trials per condition were excluded from data analysis. This data selection method is consistent with previous studies (Lloyd-Fox et al., 2017; Lloyd-Fox et al., 2012; Lloyd-Fox et al., 2009).

2.3.1.2 Sound stimuli

Sound stimuli in the voice and non-vocal task and vocal emotion task were selected from two sound batteries separately:

In the voice and non-vocal task, two types of stimuli from the validated battery (Capilla, Belin, & Gross, 2013) were presented to infants: (a) in the voice condition there were 5 female, non-speech, emotionally neutral vocalisations (e.g. interjection 'ah', yawn and etc.); (b) in the non-vocal condition there were 5 environmental sounds (e.g. sounds of running water, a bell, a horn and etc.).

In the vocal emotion task, three types of stimuli from a well-validated battery (Maurage, Joassin, Philippot, & Campanella, 2007) were presented to infants: 15 female non-speech vocalisations of angry, happy and neutral prosody (interjection 'ah').

Stimuli within the task were normalised to the same duration and mean intensity (with Praat sound-analysis software, (Boersma & van Heuven, 2001). Each condition was repeated 8 times; and the length of the testing session for the voice versus non-vocal task was 260 seconds; and for the vocal emotion task was 380 seconds. The stimuli, paradigm and measurement procedures were the same at all three measurement time points.

2.3.1.3 Infant fNIRS headgear

Infant age appropriate headgear development:

In the present study, infants' cerebral responses were recorded with a multichannel NIRS data collection system. The system was previously applied in an adult study (Trevithick et al., 2015), which was approved by local ethical committee (REC Ref

No. 12/NW/0021). The system was built by Biomedical Optics Research Laboratory, (Dept. of Medical Physics and Bioengineering, University College London) and applied with 780nm and 850nm continuous wavelengths and 10 Hz sampling rate (Everdell et al., 2005).

In consideration of infants' small head sizes and limited tolerance compared to adults, the headband was changed in terms of the size and distances between source-detector pairs. Limited by the lack of individual infant cerebral anatomy information in the present study, we were not able to offer accurate localisations information with fNIRS. However, great effort was put into designing infant age-appropriate headgear. Three procedures were followed in designing the headgear and source-detector distributions: (1) get regions of interest from previous fMRI, EEG and fNIRS studies; (2) summarise and project these cerebral regions to the scalp corresponding to EEG 10-20 system; (3) design the distribution of sources and detectors based on the anatomy of infant brain, and to cover the regions of interest. The number of sources and detectors offered by the system was also taken into consideration.

Infant neuroimaging studies in the auditory domain were reviewed and most of these studies found the significant responses in the frontal and temporal regions (Grossmann, Oberecker, et al., 2010; Grossmann et al., 2005; Lloyd-Fox et al., 2012; Lloyd-Fox et al., 2009; Minagawa-Kawai, Cristia, Vendelin, Cabrol, & Dupoux, 2011; Ortiz-Mantilla, Hamalainen, & Benasich, 2012; Pena et al., 2003; Taga & Asakawa, 2007; Teinonen, Fellman, Naatanen, Alku, & Huotilainen, 2009; Watanabe et al., 2013; Zhang et al., 2014). Regions of interest from fMRI and fNIRS studies were transformed to EEG 10-20 system with both online converter (Münster T2T-Converter, https://archive.is/YhQl6) and visual convert with indications from

the studies (Kabdebon et al., 2014; Koessler et al., 2009; Tsuzuki & Dan, 2014). Source-detector distances were decided both based on the suggested infant spatial sensitive profile (Fukui, Ajichi, & Okada, 2003), as well as the coverage of the regions of interest.

Infant headgear description:

Two detectors and six sources formed 12 source-detector pairs and have formed a doughnut shaped panel in each hemisphere. To achieve the best spatial sensitivity profile for infants (Fukui et al., 2003), the distances between source and detectors were fixed between 1.5 and 2.5 cm. Channels were distributed according to the 10-20 system and attached to a custom-made Velcro headband. According to the infants' head growth standards from World Health Organisation (World Health Organization (WHO), 2003) and previous longitudinal study (Lloyd-Fox et al., 2017), the head circumferences of 6 to 12 months old infants do not change significantly. Therefore, the application of a fixed source-detector array across three age time points was considered to be a reasonable and practical approach (Lloyd-Fox et al., 2017). The headband was adjusted by calculating the distance between the glabella and the ear of each infant, and then ensuring that T3 and T4 are between the two bottom sources in each hemisphere. This procedure has been carried out for all the infants at each time point. The locations of the channels and the channel positions with respect to the 10-20 system are presented in Figure 2.2.

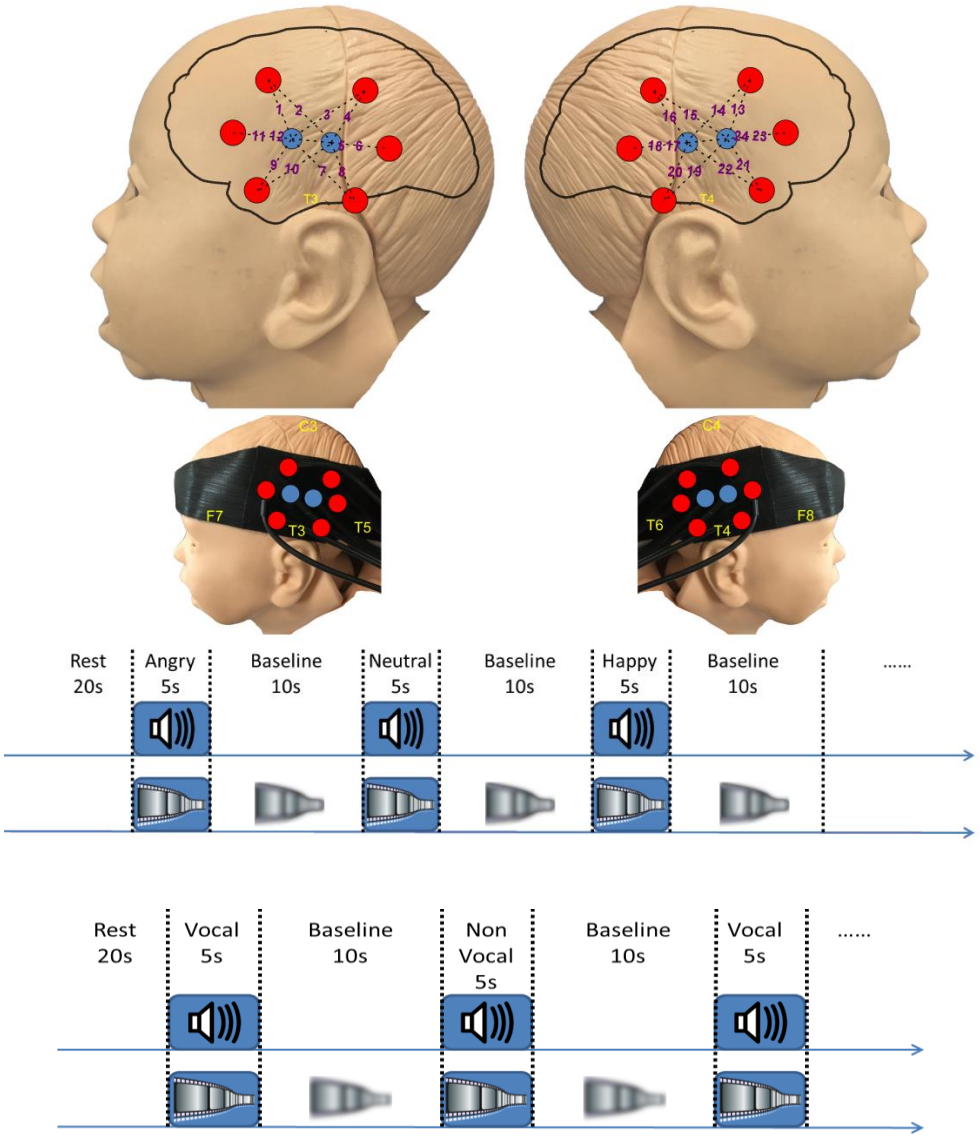


Figure 2.2. Study experimental task design and channel distribution. The head model illustrates the source-detector distribution where red dots represent sources (6 on each hemisphere) and blue dots represent detectors (2 on each hemisphere). Sources and detectors form 12 recording channels on each hemisphere, which are marked in purple numbers (upper head models), and are held by Velcro headband. The channel locations with respect to 10-20 system are marked in yellow (middle head models). The bottom two streamlines demonstrate the timeline of the experimental tasks.

2.3.2 Infant language behaviour assessment – Bayley-III language subscale

At each age point, infant language competence was assessed with the language subset of the Bayley Scales of Infant and Toddler Development, Third Edition (Bayley-III). The Bayley-III language subset consists of Receptive and Expressive Communication subscales, and was administered by a trained researcher and it took about 10 to 20 minutes to complete The raw scores were converted to composite scores and were normalised for age.

2.3.3 Mothers' feedback

After each age point session, mothers were asked to fill in the Feedback Questionnaire about their satisfaction as well as their infants' feeling during the fNIRS measurement session (Appendix 6). The Feedback Questionnaire has eight items and each with four choices (strongly disagree, disagree, agree and strongly agree): 1. The length of time of the measurement procedure was reasonable; 2. The fNIRS equipment looks good to me (e.g. the equipment didn't look too ugly or daunting); 3. The headgear fitted my baby's head; 4. The measurement session was interesting; 5. My baby concentrated on the sounds, images and videos at most of the time; 6. My baby felt comfortable during the measurement; 7. My baby liked the measurement session; 8. I feel happy to take part in the next measurement session.

2.3.4 Maternal sensitivity assessment - MACI

A 6-min mother-infant free play interaction session was video recorded after fNIRS data collection. Mothers were instructed to sit on a floor mat and play with their infant as they would normally do at home with optional (supplied) toys. Recording commenced once mother and infant were settled into play. The videos were later

rated by a trained rater (blind to participant information and study aims) using the Manchester Assessment of Caregiver-Infant Interaction (MACI; (Wan, 2015; Wan, Brooks, Green, Abel, & Elmadih, 2017), a coding system comprising eight 7-point scales validated for use in normative and at-risk groups (Wan et al., 2012, 2013).

Seven scales (two caregiver scales, three infant scales and two dyadic scales) were rated in the current study: (1) caregiver sensitive responsiveness (prompt, appropriate and attuned responses to infant behaviour and state to meet the infant's immediate and developmental needs, including an attentive attitude, appropriate engagement, and the provision of support and structuring in response to infant behaviour and a lack of behaviour); (2) caregiver nondirectiveness (a behavioural and mental focus on the infant's experience rather than the caregiver's own desires and agenda that are not at the service of the infant's experience; thus, directiveness (i.e. lower ratings) includes demanding, intrusive, negative/critical and other controlling behaviours directed at the infant); (3) infant attentiveness to caregiver (the infant's interest to the caregiver, through direct eye contact or joint activity, acceptance of and interest in caregiver, face/body orientation, and other references to caregiver activity, such as imitation); (4) positive affect (the overall amount and degree of positive emotional affective display by the infant, in their vocal, facial and gestural/bodily expression), negative affect (the overall amount and degree of negative emotional affective display by the infant, in their vocal, facial and gestural/bodily expression); (5) infant liveliness (amount and level of physical activity, particularly those behaviours initiated by the infant spontaneously); (6) dyadic mutuality (Amount and level of reciprocity, attunement and ''togetherness'', including shared attention, infant acceptance of caregiver involvement, playing together, flow and body orientation); (7) dyadic engagement intensity (Degree of

intensity of engagement by both parties at its optimal point, directly or through mutual object focus, including the degree of interest, arousal and positivity/excitement).

Inter-rater agreement based on independent blind ratings of 12 (30%) video recorded interactions was high (intraclass correlation using single measures, absolute agreement: caregiver sensitive responsiveness $r = 0.84$; caregiver nondirectiveness $r = 0.70$; infant attentiveness $r = 0.57$; infant negative affect $r = 0.82$; infant positive affect $r = 0.81$;.; infant liveliness $r = 0.65$; dyadic mutuality $r = 0.80$; dyadic engagement intensity $r = 0.30$; all $p < 0.05$).

Chapter 3 fNIRS data analysis: comparisons of standardised analysis procedure against ICA procedure

3.1 Abstract:

fNIRS is a widely used tool in cognitive neuroscience studies especially in infants, children and patients. Despite the wide application, there is a lack of a universal method that deals with artefacts. This is important in order that data capture is maximised for the technique to be deployed in hard to reach patient groups and in routine clinical settings. The present study has applied both standardised analysis and Independent Component Analysis (ICA) implanted emotion-sensitive components identification procedures. ICA has significantly improved the quality of emotion vocalisation related responses both in the waveforms and in the signal-to-noise ratio (SNR). Averaged SNR of emotion-sensitive ICs has significantly improved for both emotion conditions. All the analysis steps were based on user-friendly, easy- and ready-to-use toolboxes, aiming to offer clinicians, psychologists, and psychiatrists with a simple set of directions to undertake fNIRS data capture and potentially analysis.

Keywords: infant, fNIRS, ICA, SNR

3.2 Introduction

Over the past few decades, a new form of optical imaging known as functional Near Infrared Spectroscopy (fNIRS) has been developed and applied in neurosciences, and its feasibility has been tested within various population groups (Grossmann, Oberecker, et al., 2010; Herrmann, Walter, Ehlis, & Fallgatter, 2006; Lloyd-Fox et al., 2012; Moriguchi & Hiraki, 2009; Moser, Cutini, Weber, & Schroeter, 2009; Suto, Fukuda, Ito, Uehara, & Mikuni, 2004; Zweifel et al., 2010). fNIRS has advantages within the clinical domain of studies for its ease and portability, low cost and, thus, its feasibility for delivering in a wide range of clinical community settings. Technical advantages include a combination of relatively high spatial (but not as high as in fMRI) and temporal resolution (not as high as in EEG). For these reasons, fNIRS has gained traction over the past decade as a clinical research tool and has been increasingly applied in clinical studies in special populations such as young and old participants (Grossmann, Oberecker, et al., 2010; Herrmann et al., 2006; Lloyd-Fox et al., 2012; Moriguchi & Hiraki, 2009; Moser et al., 2009; Suto et al., 2004; Zweifel et al., 2010). These studies suggest that fNIRS provides an elegant and inexpensive solution to examine localised cortical activity with which to measure and monitor cognitive development. Its safety and lack of radiation mean it can be applied longitudinally and in children and they are not hindered by mechanical noise when applying auditory stimuli. Furthermore, it can be used in many more environments, including those as natural as walking (Miyai et al., 2001; Piper et al., 2014) and at the bedside (Zweifel et al., 2010).

When neurons are activated, blood flow increases in the cerebral region, accompanied by the increased consumption of oxygen. Theoretically, this process generates an increase in oxy-haemoglobin, a decrease in deoxy-haemoglobin in the

blood. fNIRS measures dynamic changes of haemoglobin in the volume of blood flow associated with neuronal activation. fNIRS is a non-invasive neuroimaging tool, using infrared light (wavelength ranging from 680 to 1000 nm) to detect the oxy-haemoglobin and deoxy-haemoglobin concentration changes in the cerebral blood volume to indirectly measure neural activities. Near infrared light travels from a source, passes through the scalp, skull, cerebrospinal fluid, grey matter, white matter, and is finally collected by a detector located within 2 to 4.5 centimetres from the source (Gervain et al., 2011; Lloyd-Fox et al., 2010; Villringer & Chance, 1997). The detected light forms a banana-shaped random scattering path between a source-detector pair (Gervain et al., 2011). In addition to the collected light, some light scatters in all directions and some light is absorbed in the blood within the region. With the assumption that the light scattering is constant, the remaining loss of the infrared light is due to the absorption. Oxy- and deoxy-haemoglobin concentrations have different absorption characteristics to near infrared light (Delpy & Cope, 1997). With modified Beer-Lambert law, oxy-haemoglobin and deoxy-haemoglobin can be calculated (Gervain et al., 2011; Lloyd-Fox et al., 2010; Villringer & Chance, 1997). The continuous wave (CW) system is frequently used along with two wavelengths fall in the spectrum range that can best transmit to intervening tissues, such as skull, skin, etc.(Villringer & Chance, 1997).

Despite its broad application in the fNIRS literature, there is an urgent need to separate signal from artefacts to improve the signal to noise ratio (Aslin & Mehler, 2005). Artefacts such as movements, respiration and heartbeats affect the fNIRS data significantly. Such contamination is worse in clinical trials when the sample size is small, and/or with limited measurement time (Zweifel et al., 2010). Even in well-designed infant fNIRS studies there were large amounts of data rejection with an

average exclusion rate around 40 percent (Lloyd-Fox et al., 2010): Lloyd-Fox excluded 17 out of 50 infants and 18 out of 42 infants (Lloyd-Fox et al., 2012; Lloyd-Fox et al., 2014); Grossmann has excluded 26 out of 76 infants (Grossmann, Oberecker, et al., 2010); while Nakano excluded 63 of 143 participants (Nakano, Watanabe, Homae, & Taga, 2009); all as a result of artefact or insufficient trials.

Neuroscience researchers usually adopt two approaches to improve the the signal-to-noise ratio (SNR) in outputs. One relatively simple strategy is to extend the length of the recording time (increase the number of repetitions) in order to increase the number of repetitions. Once the number of repetitions is increased, the analysis can tolerate exclusion of all the 'bad data' i.e. that which exceeds a certain value, based on visual observation or rejection criteria (Grossmann, Oberecker, et al., 2010; Lloyd-Fox et al., 2012; Lloyd-Fox et al., 2009; Piper et al., 2014). However, if the researchers choose this strategy to optimise data capture, they need to take into account the tolerance of their population group. Thus, it is well known that the longer the recording time, the fussier and more fatigued the participants will become (Lloyd-Fox et al., 2010). For example, in awake baby studies, the acceptable recording length is 10 to 15 minutes (Aslin & Mehler, 2005); similar to that of adult patients' tolerance (Miyai et al., 2001; Suto et al., 2004). Instead of rejecting channels or trials directly, the other strategy that can improve SNR in infant studies is to correct the artefacts with an external apparatus (Izzetoglu, Devaraj, Bunce, & Onaral, 2005; Robertson, Douglas, & Meintjes, 2010; Virtanen, Noponen, Kotilahti, Virtanen, & Ilmoniemi, 2011) or using analysis algorithms (Cooper et al., 2012; Cui, Bray, & Reiss, 2010; Grossmann, Oberecker, et al., 2010; Markham, White, Zeff, & Culver, 2009; Medvedev, Kainerstorfer, Borisov, Barbour, & VanMeter, 2008; Schiessl, Wang, & McLoughlin, 2008; Scholkmann, Spichtig, Muehlemann, &

Wolf, 2010; Zhang et al., 2010). Studies have proved the efficacy of algorithms artefacts exclusion or correction in fNIRS data (Brigadoi et al., 2014; Cooper et al., 2012; Cui et al., 2010; Izzetoglu et al., 2005; Robertson et al., 2010; Scholkmann et al., 2010).

However, among frequently used algorithms, there is lack of a universal method that deals with all kinds of artefacts. Each analysis algorithm is specifically efficient in handling certain a type(s) of artefacts (Brigadoi et al., 2014). Unfortunately, various kinds of artefacts appear in specific neuroimaging data and there is a large inter- and intra- individual difference in artefacts in individual data. Independent Component Analysis (ICA), on the other hand, identifies component(s) that related to the stimulus without prior model or hypothesis. It is sensitive to subtle changes in infant fNIRS data (Fox et al., 2013; Katura et al., 2008), especially when amplitudes of signal are embedded in motion artefacts, physiological and system noises, or when the frequencies of signal and noise are within the same range. The present study has applied one widely used data processing procedure and introduced a data-driven method to deal with individual differences in neuroimaging data.

Reliable, commercialised fNIRS systems are available for researchers to use within clinical trials. However, in order for them to become widely used within a routine clinical setting or trial population, it is vital that we can establish more reliable analysis procedures which can deal with large amounts of imaging data as well as minimise artefact in the signal. The present study aimed to deal with the artefacts in real-life clinical populations by using a simple methodology with readily available, low cost online analysis toolboxes. The overarching objective of the study is to enable neuroimaging to become part of routine clinical and research assessment of at risk children and to translate research to clinical settings.

3.3 Method

3.3.1 Ethics and Participants

The National Health Service ethics committee approved the study (REF: 15/NW/0684). Thirty-eight healthy infants of healthy mothers participated in the study from community health centres in Greater Manchester. Mothers consented on behalf of their infants. All infants were born with normal birth weight (>2500 g); 37 were full term (37–42 weeks gestation); and one was born at 36 weeks gestation (corrected for gestational age). Infants had no hearing difficulties according to parental report. Nine infants did not meet the minimum 4 out of 8 trials per experimental condition as a result of motion artefacts (N = 4), and fussiness (N = 5) were excluded from analysis. This attrition rate is within the standard range for infant fNIRS studies (see Grossmann et al. 2010; review by Lloyd-Fox et al., 2010). Twenty-nine typically developing 12-month-old infants were included in the final sample (15 girls and 14 boys, between 366 and 389 days, M = 376 days, SD = 6.2).

3.3.2 Data acquisition

The infants' cerebral responses were recorded with a multichannel NIRS data collection system. The system was built by Biomedical Optics Research Laboratory, (Dept. of Medical Physics and Bioengineering, University College London) and applied with 780nm and 850nm continuous wavelengths and 10 Hz sampling rate, the mean power emitted by each laser diode is approximately 2mW (Everdell et al., 2005). Two detectors and six sources formed 12 source-detector pairs in each hemisphere, and were distributed at temporal regions, which have been shown to be voice sensitive in previous research in infants (Grossmann, Oberecker, et al., 2010; Lloyd-Fox et al., 2012; Pena et al., 2003; Taga & Asakawa, 2007) and adults (Belin

et al., 2000; Ethofer, Anders, Erb, et al., 2006; Grandjean et al., 2005). To achieve the best spatial sensitivity profile for infants (Fukui et al., 2003), the distances between source and detectors were fixed between 1.5 and 2.5 cm. Channels were distributed according to the 10-20 system and attached to a custom-made Velcro headband. According to the infants' head growth standards from the World Health Organisation (World Health Organization (WHO), 2003) and from one previous longitudinal pilot study (Lloyd-Fox et al., 2017), there are no significant changes in the head circumference of 6- to 12-months-old infants. Therefore, the application of a fixed source-detector array across three age time points was reasonable and practical (Lloyd-Fox et al., 2017). In each occasion, the headband was adjusted by calculating the distance between the glabella and the ear, ensuring that T3 and T4 were between the two bottom sources in each hemisphere. This procedure was carried out for all the infants at each time point. The locations of the channels and the channel positions with respect to the 10-20 system are presented in Figure 2.2.

3.3.3 Experimental paradigm and procedure

The stimulus material consisted of 15 female non-linguistic vocalisations of angry, happy and neutral prosody (interjection 'ah') from a well-validated battery of vocal emotional expressions (Maurage et al., 2007). This battery has high internal consistency for each emotion set and high levels of specificity (independence between the ratings in the different emotion sets (Maurage et al., 2007). These stimuli have been validated in previous research in children of different ages (Chronaki, Benikos, et al., 2015). Five normalised stimuli, each lasting 1 sec, from the same condition were selected and then combined to form a 5 sec trial. All vocal stimuli were normalised with Praat sound-analysis software (Boersma & van Heuven, 2001) to the same duration of 1 000 ms and mean intensity of 73dB (see

Appendix 7 for details on stimuli acoustic properties). Vocal emotional stimuli were the same for all the three time points' data collection.

Infants wore the fNIRS headband, sat in their mother's lap in front of a laptop during the experiment and listened to the vocal stimuli. The task started with a 20 sec rest period, followed by a 5 sec trial presented through loudspeakers (SPL = 70 dB). Consistent with previous research (Grossmann, Oberecker, et al., 2010), a 5 sec silent cartoon was played along with each trial to attract infants' attention and reduce motion artefacts. After each trial, a 10 sec silent baseline along with the blurred cartoon was presented. The task was presented with PsychoPy software (Peirce, 2007). The same emotional expression did not occur consecutively. Each condition (angry, happy and neutral) was presented 8 times amounting to a total number of 24 trials. The total length of the testing session was 6 minutes and 20 seconds.

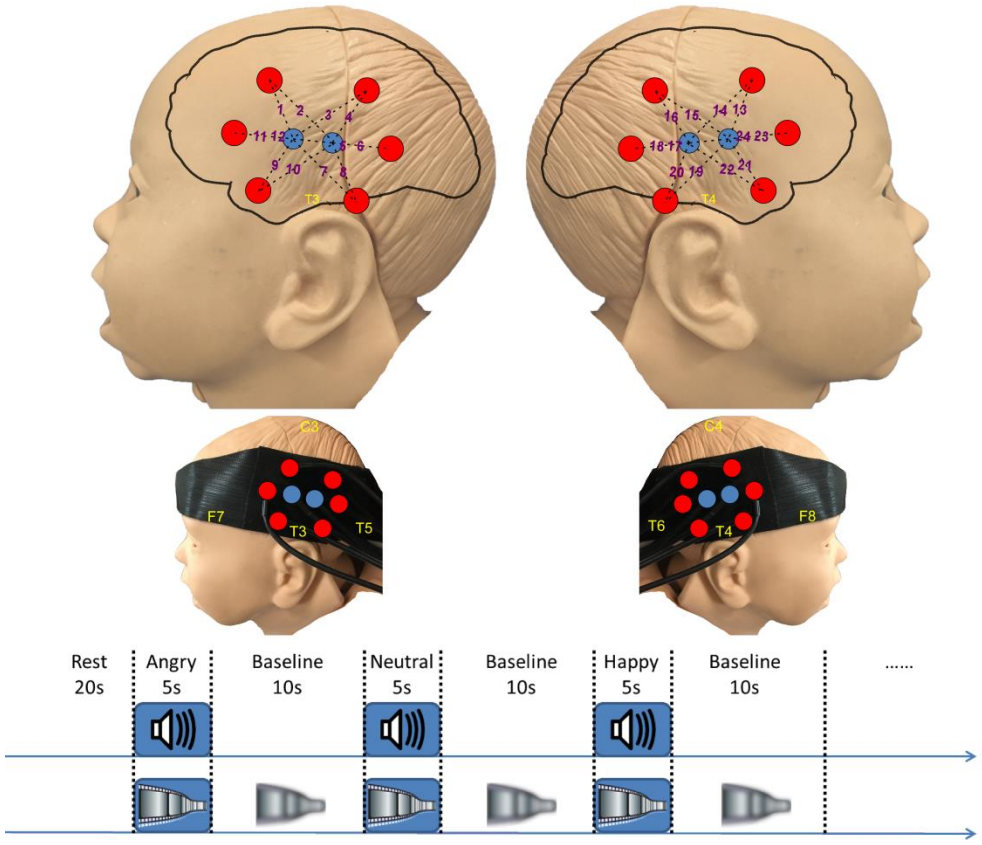


Figure 3.1. Study experimental task design and channel distribution. The head model illustrates the source-detector distribution where red dots represent sources (6 on each hemisphere) and blue dots represent detectors (2 on each hemisphere). Sources and detectors form 12 recording channels on each hemisphere, which are marked in purple numbers (upper head models), and are held by Velcro headband. The channel locations with respect to 10-20 system are marked in yellow (middle head models). The bottom two streamlines demonstrate the timeline of the experimental tasks.

3.3.4 Data processing

Behavioural assessment was carried out before data analysis. Infants' behaviour was video-recorded during the experimental task; this allowed us to code the number of trials each infant has attended to (i.e. while looking at the screen) without large motion artefacts. At least four out of eight trials per condition (i.e. 50% of trials per

condition) were set as the criterion for inclusion of each infant dataset. Twenty-nine out of thirty-eight datasets were included for further analysis using this restriction. This inclusion rate was consistent with criteria for inclusions in previous infant fNIRS studies (Lloyd-Fox et al., 2010).

Data that survived the restriction criterion were processed in two processing streamlines. The first streamline implemented commonly used procedure to reject and correct artefacts (from here on referred to as standardised procedure), which was consistent with our previous study (under review), and similar to other published studies (Brigadoi et al., 2014; Cooper et al., 2012; Grossmann, Oberecker, et al., 2010; Lloyd-Fox et al., 2012; Scholkmann et al., 2010). Emotion-sensitive channels were identified with the standardised procedure. In the second streamline, Independent Component Analysis (Markham et al., 2009; Medvedev et al., 2008; Schiessl et al., 2008; Zhang et al., 2010) was introduced to find emotion-sensitive responses (components).

Standardised analysis procedure

All the datasets included in the analysis were filtered at 0.01 to 0.5Hz with 3rd order Butterworth filter to eliminate slow drifts, instrument noise and physiological artefacts such as heartbeats (Cooper et al., 2012; Fox et al., 2013; Grossmann, Oberecker, et al., 2010; Zhang et al., 2010). The remaining artefacts were identified on a channel by channel basis with the algorithm 'hmrMotionArtifactByChannel' implemented in the HOMER2 fNIRS toolbox (version 2.1, http://homer-fnirs.org/, Huppert et al., 2009). Within the time interval (tMotion), if the change of the signal amplitude exceeded the threshold (AMPthresh) or the standard deviation changes were greater than a factor (STDEVthresh) multiplied by the original channel

standard deviation, the time period (tMask time before and after the motion artefact) was marked as artefacts. The time period of motion artefact within the channel was corrected with a cubic spline interpolation algorithm with p set to 0.99 as recommended (Cooper et al., 2012; Scholkmann et al., 2010). Since the algorithm works on a channel by channel basis, the actual standard deviation threshold for the motion artefact varies according to the standard deviation of the original channel; the setting of the STDEVthresh is the multiplication factor rather than a fixed threshold (i.e. in the current study the standard deviation threshold is 20*standard deviation of the channel). This means that the standard deviation threshold varies from channel to channel and subject to subject. All the values were set as follows: tMotion=5s; tMask=1s; STDEVthresh=20; AMPthresh=5.

After pre-processing, Oxy- and Deoxy-Haemoglobin concentration changes (ΔHbO^2 and ΔHbR) were averaged across trials in the same emotion condition within each dataset in HOMER2, with the time window of 1 sec before and 15s after the stimulation onset. The averaged time course of each channel was corrected by subtracting the mean of the 1 sec before the stimulation. The analysis focused on Oxy-Haemoglobin concentration changes, because it is the most sensitive indicator of changes in cerebral blood flow. Kolmogorov-Smirnov tests were conducted to test the normality of the haemodynamic time courses, results showed that p values were larger than 0.05, which indicate that haemodynamic data were normally distributed. Based on the earlier work, we targeted a time window of 2 sec to 9 sec after stimulus onset (in preparation). Mean amplitudes of cortical Oxygen haemodynamic responses were averaged over the time window of 2 sec to 9 sec after stimulus onset. The averaged ΔHbO^2 to the expression conditions (angry, happy and neutral) were evaluated with repeated measures ANOVA and post-hoc pairwise comparisons.

Significant channels that showed emotion-sensitive responses were extracted for later analysis procedures.

Task-correlated Independent Component Analysis

After pre-processing with a standardised correction procedure, the second analysis pipeline started with the full length of artefact-corrected oxygenated haemoglobin concentration matrices (channels*time points).

Independent Components extraction

In the present optical imaging study, each recording site (channel) collected a mixture of haemodynamic responses to stimuli, along with different kinds of noise (i.e. physiological noise, and other noises). Even after the standardised analysis processing, task-unrelated noises remained. Since the underlying haemodynamic responses of neurones are independent from each other, Independent Component Analysis has been carried out to separate task-related signal (i.e. emotion-sensitive responses) and task-unrelated noise that mixed in haemodynamic concentration data.

FastICA v2.5 toolbox (http://research.ics.aalto.fi/ica/fastica/, Hyvärinen and Oja, 2000) was used to decompose independent components in the current study. FastICA is implanted in Matlab (The MathWorks, Inc., Natick, MA, USA). The algorithm is based on fixed-point iteration to find the maximum non-Gaussianity of the estimated source (Hyvärinen & Oja, 2000). FastICA has a friendly interface which enables beginners to set parameters and import data. The following parameters were set to extract components: Approach: deflation; Number of ICs: 12; Nonlinearity: pow3; Stabilization: off; Fine-tune: off; Maximum number of iterations: 1000; Initial state: random. In the current study, we have focused on the left hemisphere ΔHbO^2 from

two emotional conditions (angry and happy); each subject has a matrix of 12 channels * time points for each emotional condition was put in FastICA.

Quality assessment

The quality of both channel data and IC data were assessed on their time courses. Two steps were involved in the assessment: (1) Fast Fourier Transformation (FFT); (2) SNR calculation (Schiessl et al., 2008). Signal-to-noise ratio (in decibel) was calculated by comparing the maximum signal power with the maximum noise power (Schiessl et al., 2008). According to the experimental paradigm and the physiological development of infants, task-related responses (signal) should have been seen at the frequency range of 0.06 to 0.08 Hz, whereas physiological artefacts (noise) such as respiration and heartbeat should be at the frequency range of 0.5 to 1 Hz and 1.3 to 2.7 Hz, respectively. In the equation below, P_S represents the power of the signal, which was defined as the maximum power within the frequency range of 0.06 to 0.08 Hz; P_N represents the power of noise, which was identified as the maximum power at the frequency range larger than 0.08 Hz.

$$SNR = 10\log_{10}(P_S/ P_N)$$

Emotion-sensitive components identification

The obtained independent components (ICs) were averaged across repetitions within each condition and dataset, and correlated with the emotion-sensitive channel. Those ICs with high correlation coefficients (> 0.40) to the emotion-sensitive channel (channel 9) were identified as the emotion-sensitive ICs.

Identification of task-related time courses

Since task-related channel was identified through statistical analysis, there were chances that channel data in some datasets might not have captured the task-related responses but rather noise. Therefore, we observed all the time courses (both task-related channel and IC time courses) in a qualitative way, to ensure the capture of task-related responses. In order to be objective in the qualitative observation, a blind time course selection procedure was also carried out. Irrespective of SNR, all the channel and IC time courses were averaged over the repetition trials and randomised in the presentation order. Task-related time courses were selected based on the criterion that an identifiable peak (which related to the task within the time window of 2s after the stimulus onset and 4s after the stimulus offset) would be identified.

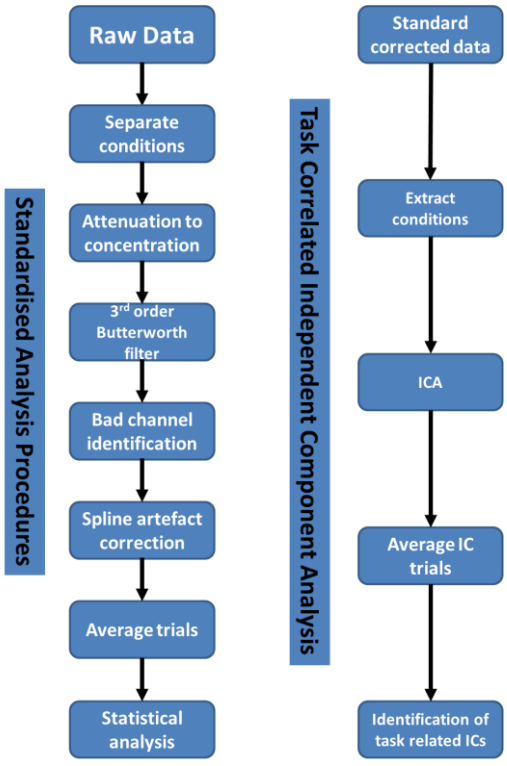


Figure 3.2 Data processing streams

Two streamlines in identifying task-related channels and task-related components.

3.4 Results

The results suggested that, to some extent, the standardised analysis procedure excluded noise and identified task-related channels. Channel 9 in the left hemisphere was sensitive to emotional prosody in ΔHbO^2 ($F(2, 56) = 4.17$, $p=.021$, $\eta_p^2=.13$). Happy voices had significantly greater Oxy-Haemoglobin concentration changes when compared to angry voices ($F(1, 28) = 10.53$, $p = .003$, $\eta_p^2 = .27$). The averaged waveform for each condition is shown in figure 3.4.

One IC was selected for each subject; there was no difference in the total number of ICs ($t = 0.85$, $p = 0.401$), or correlation coefficients ($t = 0.45$, $p = 0.654$) between the angry and happy conditions. In summary, the coefficient of determination (R^2) for two conditions ranged between 17.6% and 92.2%; the averaged coefficient of determination was 42.3% for the angry condition and 41% for the happy condition. This result indicates that emotion-sensitive channel and IC shared 41% of the variance on average. The detailed correlation coefficient results are shown in Table S3.1 in Supplementary Information.

Signal-to-noise ratio improvement in decibel

The quality of task-related channel time courses and task-related ICs was evaluated by the FFT-SNR quality assessment procedure. Paired-sample t-tests were conducted to examine SNR after applying two analysis methods in each condition. In summary, SNR of task-related ICs were larger in both conditions ($t > 3.2$, $p < 0.01$, Table 3.1). The average FFT-SNR for task-related channel time courses across the 29 datasets was -0.75 dB; the mean FFT-SNR for the corresponding task-related ICs was 0.91 dB and the improvement of SNR was 1.66 dB on average for the angry condition. In the happy condition, the average FFT-SNR for task-related channel data was -0.71

dB, for the corresponding task-related ICs was 0.70 dB and the average SNR improvement was 1.41 dB (Table 3.1, and Table S2 in Supplementary Information).

Table 3.1 the comparison of quality in task-related channel and task-related IC time courses

Condition	Task-related Channel time courses (29 in total)	Task-related IC time courses (29 in total)	Channel data Mean±SD	IC Mean±SD	t	p
Angry	13/29	25/29	0.75±1.79	0.91±1.96	4.31	**<0.001****
Happy	10/29	24/29	0.71±2.39	0.70±1.15	3.22	**0.003***

*p < 0.05; ** p < 0.001.

Time course improvement

Task-related time courses were identified, after time courses from the sensitive channel and IC were assessed blindly. From the time course perspective, the improvement seen in the data using the ICA procedure was obvious: in the angry condition, out of a total of 29 time courses, 13 emotion-sensitive channel time courses showed task-related responses, compared with 25 emotion-sensitive IC time courses being task-related. This was even more striking in the happy condition where only 10 out of 29 emotion-sensitive channel time courses showed task-related responses, while 24 out of 29 emotion-sensitive IC time courses were task-related. Overall, the improvement seen in time courses using ICA was doubled for both emotion conditions (Table 3.1). Figures 3.3 and 3.4 provide a more visual version of the improvement in time courses, demonstrating the averaged emotion-sensitive IC time courses compared with the corresponding emotion-sensitive channel time course.

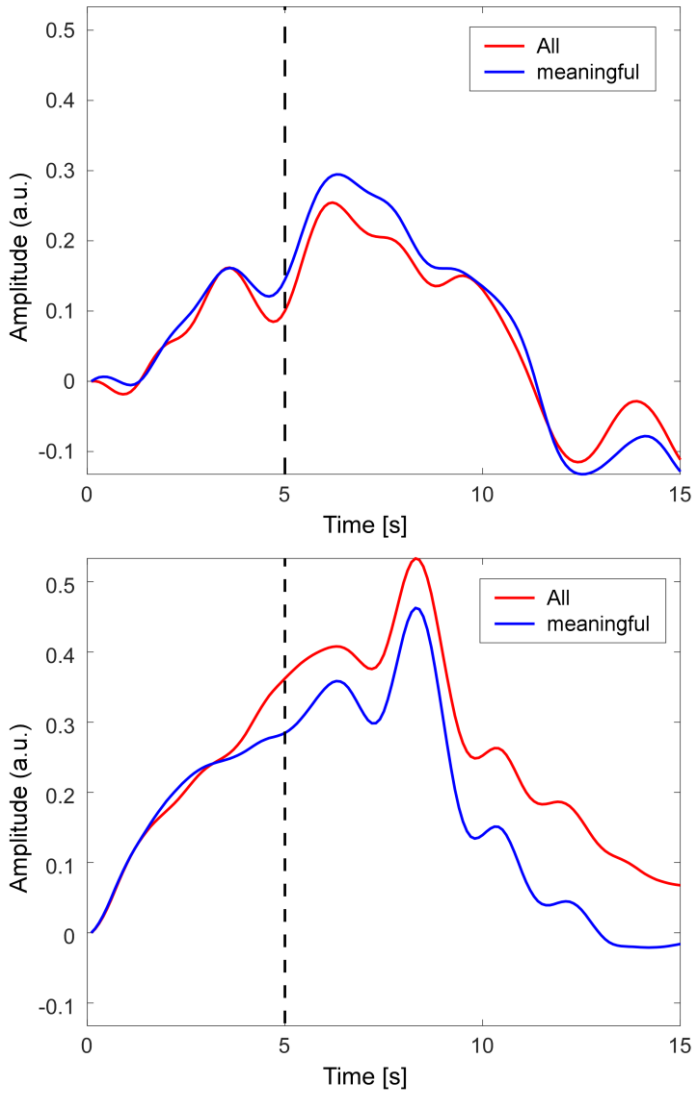


Figure 3.3 shows the averaged emotion-sensitive IC time courses for the angry (top figure) and happy conditions (bottom figure). Blue waveforms represent the average task-related time courses after blind identification, and red time courses represent all averaged time courses.

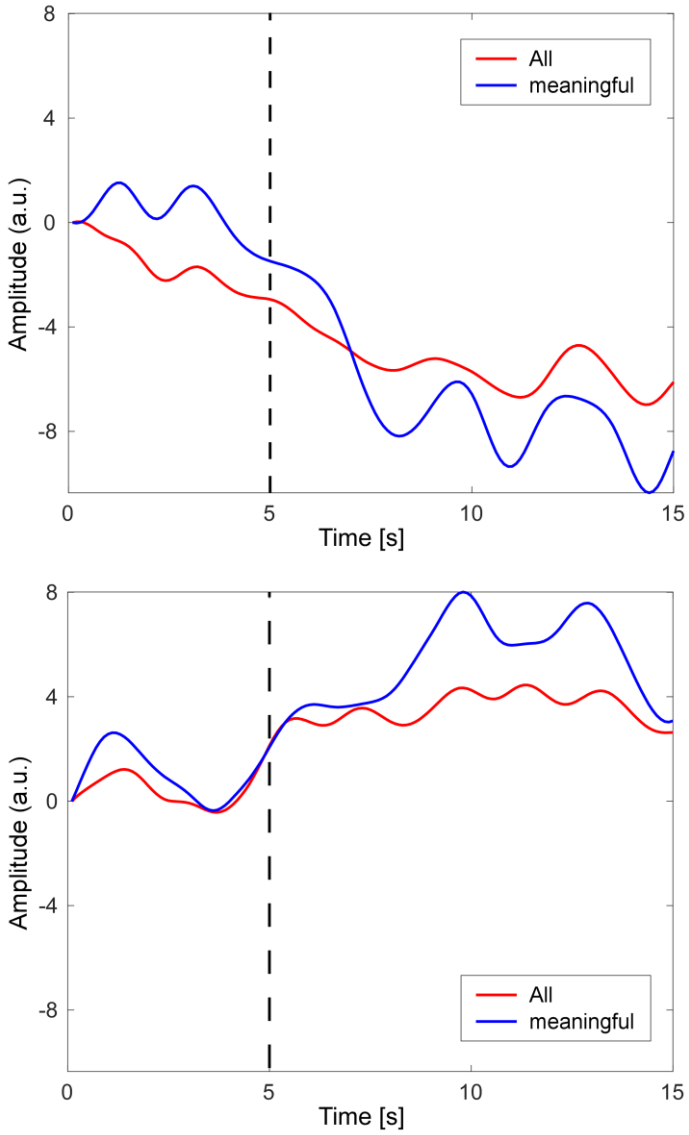


Figure 3.4 shows the averaged emotion-sensitive channel time courses for the angry (top figure) and happy conditions (bottom figure). Blue waveforms represent the average task-related time courses after blind identification, and red time courses represent all averaged time courses.

3.5 Discussion

The present study is part of a project piloting the development of a clinically acceptable and feasible infant fNIRS-paradigm for use in routine clinical community settings. As such our primary aim was to improve the SNR of infant fNIRS data in order to optimise data capture in high-risk infant groups. To do this, we applied both standardised analysis and ICA involved emotion-sensitive components identification procedures. There were two main findings. First, the standardised analysis procedure excluded common large artefacts in the data, but overlooked the individual artefact differences in infants. Secondly, we report that the alternative analysis route compensated for the shortcomings of the standardised analysis procedure by applying individual ICA. As a result, we found that the identified emotion-sensitive ICs showed significant improvement both in the waveforms (qualitative improvement) and in SNR (quantitative improvement) compared to results from the standardised analysis procedures. Time courses from emotion-sensitive ICs were less noisy and clearly related to the stimulation condition than time courses from the emotion-sensitive channel. Averaged SNR of emotion-sensitive ICs was significantly improved for both emotional conditions.

Using the standardised analysis procedure, initially we achieved a relatively crude measure of emotion-sensitive brain activation. Channel 9 in the left hemisphere was identified showing emotion-sensitive responses. It is assumed that fNIRS ΔHbO^2 should increase from the onset of stimulation until reaching a peak at 2 - 4s after stimulation offset (Brigadoi et al., 2014). The averaged emotion-sensitive channel time course for each condition is shown in figure 3.4. In the angry condition, ΔHbO^2 started from zero and decreased over time; while in the happy condition, ΔHbO^2 had a small peak in the stimulation condition and a greater increase in the baseline

period. These results were not consistent with the assumption or with typical haemodynamic response waveform shapes. One plausible explanation for this is insufficient artefact exclusion alongside greater individual differences. The former explanation is in line with the limitation of spline interpolation.

Spline interpolation implanted standardised analysis subtracts artefacts based on accurately identified noise standard deviation. It corrects artefacts with sudden deviant jumps in the light attenuation time courses (Scholkmann et al., 2010). This technique is valuable for correcting identified artefacts without affecting the un-identified signal. The disadvantage of this technique is the heavy dependence on the variance of artefacts, which need to be significantly greater than the signal, to ensure the correct detection and removal (Scholkmann et al., 2010). Therefore, noises that have similar or smaller amplitude compare with the signal are not identified or corrected by spline interpolation.

The standardised analysis procedure directed us towards the stimulation correlated location in brain (channel 9). Further emotion-sensitive independent component identification was carried out, to detect individual sensitive ΔHbO^2 in both emotional stimuli conditions. ICA separates a series of components from different sources within the brain without necessarily knowing where the sources derive from. In EEG and optical imaging studies, ICA functions well in decomposing artefact components such as eye-movements, heartbeats and respirations (Makeig, Bell, Jung, & Sejnowski, 1996; Markham et al., 2009; Medvedev et al., 2008; Schiessl et al., 2008; Vigario, Jousmaki, Hamalainen, Hari, & Oja, 1998; Zhang et al., 2010; Zhao, Valentini, & Hu, 2015). By transforming pre-processed data into zero mean and unit variance independent components, pre-processed ΔHbO^2 were first decomposed into stimulation-related and unknown physiological responses. Then further correlational

analysis ensured identification of emotion-sensitive responses. In each dataset, one emotion-sensitive independent component was obtained after performing a correlational analysis between the task-related channel and ICs (the corresponding averaged time course for each condition was shown in Figure 3.3). In both conditions, emotion-sensitive IC waveforms showed gradually increased responses immediately following onset of the stimulus and continued to a peak around 2 – 4s after the stimulation offset. In addition, there was a significant improvement of SNR after applying emotion-sensitive components identification procedures. All these elements suggest that the signal we were capturing was indeed infant vocal-emotional brain responses.

Results from the standardised analysis procedure suggest that this well-recognised problem in automatic correction procedure, i.e. the insufficiency in artefact identification and correction and lack of information about individual difference were responsible for the differences in findings between analysis procedures. Other studies have suggested that the most frequently used analysis techniques are, indeed, fit for a certain type(s) of artefacts (Brigadoi et al., 2014; Cooper et al., 2012), and that it is not possible to eliminate all the noise embedded in fNIRS data. The application of ICA in the present study compensated for these shortcomings. The identified emotional sensitive ICs represent infant neural responses to emotional stimuli without the interference of artefacts. Individual differences were also taken into consideration by applying individual ICA. The improvement of SNR, as well as the waveforms, suggested the efficacy of ICA in infant fNIRS data analysis. In addition, the decomposition process did not involve any model which avoids the problem of a lack of infant model, and the potential mismatch between a model and actual infant data.

Functioning brain imaging studies on special population groups, such as infants, children and patients, generally require relatively short recording time with a small number of channels. As a result of these constraints, data from such populations is more likely to contain a relatively high proportion of artefact compared to that from healthy adult populations. This means that clinical researchers should be very cautious when dealing with imaging data. In order to make the maximum use of the collected data, the present study used standardised analysis procedures which were relatively conservative in correcting artefact-contaminated time periods. So to overcome this shortcoming i.e. leaving artefact in the data, we applied a more novel task-related components identification procedure.

Our findings suggest that such an approach results in greater accuracy for the observation of true stimulation-related brain responses. The overarching intention of this study was to develop an ecologically valid analysis method suitable for future clinical use and possibly for monitoring responses to interventions within clinical trials. All the analysis steps were based on user-friendly, easy- and ready-to-use toolboxes, aiming to offer a simple means with which to undertake fNIRS data analysis in a routine clinical setting.

3.6 Supplementary Information

Table S3.1 Identification of emotion-sensitive IC in each subject

Subject ID	Angry			Happy		
	Total IC No	Correlation Coefficient (R^2)	Task-related IC location	Total IC No	Correlation Coefficient (R^2)	Task-related IC location
1	12	0.71 (0.50)	9th	12	0.54 (0.29)	10th
2	9	0.43 (0.18)	9th	12	0.86 (0.74)	4th
3	12	0.71 (0.50)	9th	7	0.67 (0.45)	1st
4	12	0.51 (0.26)	4th	9	0.69 (0.48)	9th
5	8	0.50 (0.25)	6th	9	0.58 (0.34)	3rd
6	12	0.58 (0.34)	1st	10	0.51 (0.26)	3rd
7	8	0.60 (0.36)	3rd	8	0.56 (0.31)	8th
8	10	0.78 (0.61)	3rd	10	0.88 (0.77)	8th
9	12	0.54 (0.29)	5th	9	0.75 (0.56)	8th
10	12	0.60 (0.36)	9th	8	0.64 (0.41)	8th
11	9	0.74 (0.55)	5th	8	0.82 (0.67)	7th
12	10	0.69 (0.48)	5th	10	0.71 (0.50)	10th
13	12	0.74 (0.55)	5th	12	0.55 (0.30)	12th
14	12	0.82 (0.67)	11th	9	0.77 (0.59)	4th
15	10	0.76 (0.58)	3rd	12	0.76 (0.58)	3rd
16	9	0.83 (0.69)	3rd	9	0.69 (0.48)	5th
17	8	0.73 (0.53)	2nd	12	0.91 (0.83)	5th
18	9	0.81 (0.66)	6th	10	0.52 (0.27)	8th
19	9	0.64 (0.41)	5th	10	0.66 (0.44)	8th
20	10	0.58 (0.34)	10th	10	0.42 (0.18)	10th
21	12	0.48 (0.23)	6th	10	0.45 (0.20)	7th
22	12	0.84 (0.71)	9th	12	0.60 (0.36)	1st
23	8	0.47 (0.22)	2nd	8	0.55 (0.30)	7th
24	12	0.56 (0.31)	6th	10	0.51 (0.26)	7th
25	12	0.52 (0.27)	1st	10	0.79 (0.62)	5th
26	8	0.96 (0.92)	2nd	12	0.60 (0.36)	4th
27	12	0.55 (0.30)	1st	8	0.52 (0.27)	6th
28	12	0.80 (0.64)	2nd	12	0.60 (0.36)	10th
29	8	0.50 (0.25)	1st	12	0.45 (0.20)	11th
Average mean±SD	10.4±1.7	0.65±0.14		10±1.6	0.64±0.14	

Table S3.2 Signal-to-noise ratio of emotion-sensitive channel and IC in each subject

Subject ID	Angry			Happy		
	Channel 9 SNR	IC SNR	Improvement	Channel 9 SNR	IC SNR	Improvement
1	1.62	4.05	2.43	-1.32	0.87	2.19
2	-0.58	0.75	1.32	0.02	0.74	0.71
3	-3.96	-1.37	2.59	1.73	2.50	0.77
4	-2.75	0.46	3.21	0.71	0.00	-0.71
5	-1.58	3.60	5.18	0.95	2.45	1.50
6	0.02	0.62	0.59	3.15	1.15	-2.00
7	0.39	0.75	0.36	-3.28	2.88	6.16
8	-4.53	0.27	4.79	-0.38	1.67	2.05
9	-0.46	3.16	3.62	-3.95	0.23	4.19
10	0.85	2.01	1.16	-0.12	-2.14	-2.02
11	0.19	-0.42	-0.62	-0.45	0.67	1.12
12	-0.59	1.95	2.54	-5.62	0.99	6.61
13	0.35	1.00	0.65	0.51	2.23	1.71
14	-1.92	-4.90	-2.98	0.08	0.69	0.62
15	0.56	1.85	1.29	3.69	1.80	-1.88
16	-0.09	2.54	2.63	-0.05	-0.26	-0.21
17	2.70	1.41	-1.29	-0.59	0.01	0.60
18	-4.07	0.33	4.39	-1.95	-0.01	1.94
19	-0.31	-2.22	-1.91	0.87	0.51	-0.36
20	-2.06	-1.51	0.55	0.28	1.14	0.85
21	-1.36	0.75	2.11	0.04	0.65	0.61
22	-1.88	1.09	2.96	-2.50	-1.20	1.30
23	-3.05	-0.10	2.95	0.68	0.39	-0.29
24	-0.29	2.79	3.08	-1.59	-0.41	1.18
25	0.29	2.51	2.22	-2.88	-0.79	2.09
26	-1.49	-1.22	0.27	-2.61	0.94	3.54
27	-0.44	3.48	3.92	-6.87	-0.38	6.50
28	2.22	0.16	-2.06	2.63	0.97	-1.66
29	0.53	2.58	2.05	-1.72	1.98	3.70

Chapter 4 The feasibility and acceptability of longitudinal functional imaging in infants of healthy mothers

4.1 Abstract:

Early delays or impairments in language and social-emotional functioning may represent early markers of developmental atypicality. Advances in psychophysiological tools have led to an understanding of the neural correlates that underpin infant voice and vocal emotion processing as the precursors to language and social-emotional brain development. Because of the little evidence for the longitudinal developmental trajectory of infant voice and vocal emotion processing, there is an urgent need for longitudinal studies. However, infant neuroimaging studies carry several challenges which hinder the exploration of infant longitudinal research. The present study piloted a longitudinal fNIRS study of infant neural activation in response to voice and social-emotional stimuli at 6, 9 and 12 months of age, and tested its acceptability and feasibility. The high satisfaction and retention rate, as well as low loss of data, indicate a promising outlook for the clinical use of fNIRS even in future routine assessment of infant neurodevelopment.

Keywords: fNIRS, infant, feasibility, acceptability, clinic

4.2 Background

Language and social-emotional atypicalities may act as important early markers of potential developmental problems in children (Beitchman et al., 1996; Chronaki, Benikos, et al., 2015; Chronaki, Garner, et al., 2015) and, subsequently, in adolescents (Beitchman et al., 2001) and adults (Baum & Nowicki, 1998; Beitchman et al., 2008). Because of the wide variation in normal language and social-emotional development (Bates et al., 1992; Cole et al., 2006; Fenson et al., 1994; Kitayama et al., 1995; Määttä et al., 2016), many, if not most children with language or social-emotion developmental problems are not diagnosed until around the age of 3 - 6 years old (Palfrey et al., 1987; Pinto-Martin et al., 2005). In order to target intervention and social resources to those at highest risk within a vulnerable population at an early enough age to optimise the clinical and cost effectiveness of interventions, research has increasingly moved towards identifying early indicators of risk.

A recent study reported that language and social-emotional developmental delays/deficits are identifiable at a group level in emergent developmental disorders at around 9-13 months of age from behavioural measures (Wan et al., 2018). With contemporary psychophysiological methods (e.g. EEG, fMRI, fNIRS and etc.), identifiable difference in neural responses to social stimuli were found between typically developing infants and infants at risk of developmental disorder from as young as 4 months (Blasi et al., 2015; Elsabbagh et al., 2012; Lloyd-Fox, Blasi, et al., 2013). The use of human vocalisations as stimuli supported the identification of typical and atypical language and social-emotional brain development in infants (Blasi et al., 2015; Grossmann, Oberecker, et al., 2010; Kuhl et al., 2014; Lloyd-Fox, Blasi, et al., 2013) and children (Chronaki, Benikos, et al., 2015; Tsao et al., 2004).

Studies have provided ground breaking findings, yet there remains little evidence for the longitudinal developmental trajectory of infant voice and vocal emotion processing. Only one previous infant study has examined the development of the infant auditory 'social brain' longitudinally in a rural Gambian cohort (Lloyd-Fox et al., 2017).

It is emphasised that there is a need for longitudinal studies in children to improve wellbeing and prevent mental illness, as well as for longitudinal functional imaging research (Wykes et al., 2015). Albeit the urgent need for infants and children longitudinal imaging research, infant neuroimaging studies carry several challenges which hinder the exploration of longitudinal research. These include infants' low tolerance of being seated, relatively short attention span, low-level of cooperation in setting up the system, and there are more artefacts in the data compared to the data from healthy adult populations. Even well-designed infant fNIRS studies suffer from large amounts of data loss, with an average exclusion rate of around 40 percent (reviewed by (Lloyd-Fox et al., 2010). The overall goal, therefore, was to create an acceptable and feasible fNIRS paradigm for an empirical approach to be taken in routine clinical assessment and monitoring of early infant language and social-emotional development. The aims of this study were i) to develop a method to assess infants' neural processing of human non-speech voice and vocal emotion; ii) to pilot the acceptability and feasibility of this method for mothers and infants in a routine clinical setting; iii) to test the feasibility of collecting reliable longitudinal data using the fNIRS imaging paradigm. We were particularly concerned to ensure that the fNIRS measurement paradigm should be acceptable to mother-infant pairs, and could feasibly be applied in non-research a clinical setting while obtaining reliable

imaging data; and, finally, that we could recruit women and their infants to attend for testing on three consecutive occasions approximately 3 months apart.

Acceptability was tested by asking mothers for their feedback after each measurement session and by the participant retention rate over the three time points. Feasibility was based on both the practicability of undertaking the developed fNIRS paradigm in a non-research, clinical setting and the quality of fNIRS data (i.e. valid fNIRS data). To ensure that the excluded datasets were not due to biases within the sample characteristics, for example infant liveliness, affect and etc., infant behavioural characteristics were evaluated and compared between infants whose data were included in and whose data were excluded from the final analysis. Infant liveliness, social attentiveness and affect were evaluated from observed mother-infant interactions. In addition, in order to make sure participated infants' language were developing typically, and that there was no difference in the language competence between the included and excluded infants, their language development was monitored using the language subset of the Bayley-III at 6, 9 and 12 months of age.

4.3 Materials and methods

Considerations for the Experimental Design

We chose fNIRS as our infant neuroimaging measurement for its non-intrusiveness and advantages over fMRI and EEG because of a) its portability and flexibility for delivery in a wide range of clinical and community settings; b) its silence provides the possibility of applying auditory stimuli to participants; c) it is arguably less intimidating/scary looking (compared to an MRI scanner tube as well as being considerably quieter for the infant); the infant can sit in the mother's lap during measurement; d) it has relatively high temporal resolution (compared to fMRI), and high spatial resolution (compared to EEG); (f) importantly, it is low cost compared to other neuroimaging tools which facilitates repeated testing and the potential for clinical scalability.

On account of an infant's small head size, we created an adjustable headset fit for infants. In addition, to ensure successful measurement in an awake infant, we adopted a relatively small number of channels on the infant's head. The localisations of the channels were chosen carefully in order to cover the temporal regions, which have been suggested to respond sensitively to voice stimuli (Belin et al., 2000; Ethofer, Anders, Wiethoff, et al., 2006; Grandjean et al., 2005; Grossmann, Oberecker, et al., 2010; Lloyd-Fox et al., 2012; Pena et al., 2003; Taga & Asakawa, 2007).

Stimuli were selected from validated batteries (Capilla et al., 2013; Maurage et al., 2007), according the infant's language and social-emotional developmental level, which was suggested by previous behaviour studies that infants are capable of distinguishing basic emotions such as happy and angry by the age of 6 months

(Mastropieri & Turkewitz, 1999; Singh et al., 2002; Walker-Andrews & Grolnick, 1983).

We were aware from the previous literature as well as from our own early piloting that infants tend to have a rather low tolerance for sitting still and a relatively short attention span (Aslin & Mehler, 2005). Therefore, we created as short a measurement session as we felt was possible, i.e. 11 minutes in total; and we configured the programme with easy exit (and re-start) buttons. Following previous studies, we also embedded a silent cartoon in the experiment for the infant to watch while the sounds were playing so as to make the test more interesting to infants; and to attract their attention and keep them as still as possible on their mother's lap with no artificial restraints.

Lastly, to make it as accessible as possible to new mothers, we carried out the test in local clinics generally within walking distance of the mothers. This was particularly important as part of our piloting of the method for a future routine clinical use. To optimise success rates of follow up of mothers and their infants longitudinally, to give as much flexibility to participants as possible, and to encourage the participation in the three measurements, each infant was given a one month time window for each session at per age point (i.e. tested between 6-7 months, 9-10 months and 12-13 months).

Experimental paradigm and procedure

Figure 4.1 displays the experimental paradigm of the study. The infant wore the NIRS headband whilst s/he sat in the mother's lap, in front of a laptop during the experiment and listened to the sound stimuli. The task started with a 20 sec rest period, followed by a 5 sec trial presented through loudspeakers (SPL = 70dB).

Consistent with previous research (Grossmann, Oberecker, et al., 2010), a 5 sec silent cartoon was played along with each trial to attract infants' attention and reduce motion artefacts. After each trial, a 10 sec silent baseline, along with the blurred cartoon was presented. The task was presented with PsychoPy software (Peirce, 2007).

Two tasks were included in the fNIRS session: the voice versus non-vocal task; and the vocal emotion task. The paradigm and procedure were the same for both tasks and the only difference was the stimuli. At 6 months, all infants were given a break between the two tasks; at 9 and 12 months, a break was optional and most infants completed both tasks without a break. In cases where infants were fussy or reluctant to have the headgear on or to stay seated, the task was paused to allow a break and the decision to resume rested with the mother. If the infant remained unsettled on the second attempt, the researcher discussed whether to arrange another visit to do the uncompleted part(s).

The infant's behaviour during the fNIRS procedure was video recorded and later reviewed to check the number of trials in which the infant attended to the screen without large motion artefacts. Infants who did not attend for 4 or more of 8 trials per condition were excluded from data analysis. This data selection method is consistent with previous studies (Lloyd-Fox et al., 2017; Lloyd-Fox et al., 2012; Lloyd-Fox et al., 2009).

Sound stimuli

Two types of stimuli from a validated battery (Capilla et al., 2013) were presented to infants for the voice versus non-vocal task: (a) in the voice condition there were 5 female, non-speech, neutral vocalisations (e.g. interjection 'ah', yawn and etc.); (b)

in the non-vocal condition there were 5 environmental sounds (e.g. sounds of the bell, horn and etc.).

In the vocal emotion task, stimuli consisted of 15 female, non-speech vocalisations of angry, happy and neutral prosody (interjection 'ah') from a well-validated battery of vocal emotional expressions (Maurage et al., 2007).

Stimuli within the task were normalised to the same duration and mean intensity (with Praat sound-analysis software, (Boersma & van Heuven, 2001). Each condition was repeated 8 times; and the length of the testing session for the voice versus non-vocal task was 260 seconds; and for the vocal emotion task was 380 seconds. The stimuli, paradigm and measurement procedures were the same at all three measurement time points.

fNIRS Data acquisition

In the present study, infants' cerebral responses were recorded with a multichannel NIRS data collection system. The system was built by Biomedical Optics Research Laboratory, (Dept. of Medical Physics and Bioengineering, University College London) and applied with 780nm and 850nm continuous wavelengths and 10 Hz sampling rate (Everdell et al., 2005). Two detectors and six sources formed 12 source-detector pairs on each hemisphere and were distributed at temporal regions, which have been shown to be voice sensitive in previous research in infants (Grossmann, Oberecker, et al., 2010; Lloyd-Fox et al., 2012; Pena et al., 2003; Taga & Asakawa, 2007) and adults (Belin et al., 2000; Ethofer, Anders, Wiethoff, et al., 2006; Grandjean et al., 2005). To achieve the best spatial sensitivity profile for infants (Fukui et al., 2003), the distances between source and detectors were fixed between 1.5 and 2.5 cm. Channels were distributed according to the 10-20 system

and attached to a custom-made Velcro headband. According to the infants' head growth standards from World Health Organisation (World Health Organization (WHO), 2003), and previous longitudinal study (Lloyd-Fox et al., 2016), head circumferences of 6 to 12 month old infants do not change significantly. Therefore, the application of a fixed source-detector array across three age time points was considered to be a reasonable and practical approach (Lloyd-Fox et al., 2016). The headband was adjusted by calculating the distance between the glabella and the ear of each infant, and then ensuring that T3 and T4 were between the two bottom sources on each hemisphere. This procedure was carried out for all the infants at each time point. The locations of the channels and the channel positions with respect to the 10-20 system are presented in Figure 4.1.

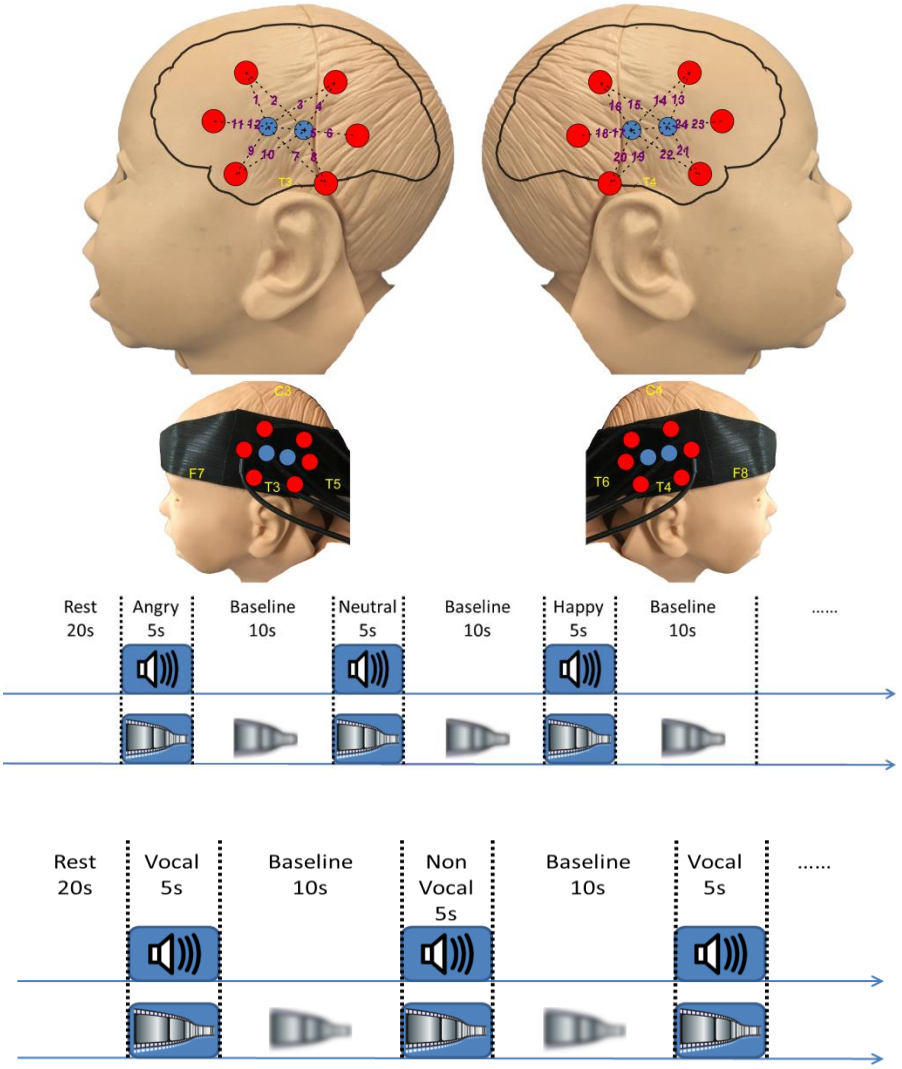


Figure 4.1. Study experimental task design and channel distribution. The head model illustrates the source-detector distribution where red dots represent sources (6 on each hemisphere) and blue dots represent detectors (2 on each hemisphere). Sources and detectors form 12 recording channels on each hemisphere, which are marked in purple numbers (upper head models), and are held by Velcro headband. The channel locations with respect to 10-20 system are marked in yellow (middle head models). The bottom two streamlines demonstrate the timeline of the experimental tasks.

Bayley Scales of Infant and Toddler Development III - Language subset (Bayley, 2006)

A test of receptive and expressive communication items from the language subset was administered by a trained researcher and it took about 10 to 20 minutes to complete. The raw scores were converted to composite scores and were normalised for age.

Infant behavioural and affective tendencies based on mother-infant interaction

The loss of infant data in psychophysiology studies is mainly due to the infant fussiness and motion artefacts during data collection. The first step of data analysis is normally observation based exclusion procedure, which may be subjective and biased by the participants' behavioural and affective tendencies. To test whether there was a bias in the data exclusion, as part of the feasibility, we also measured 6-month-old infant behavioural and affective tendencies, from a mother-infant free play interaction session, which was video recorded following the fNIRS procedure. The interaction sessions were rated by a trained rater (blind to participant information and study aims) using the validated Manchester Assessment of Caregiver-Infant Interaction (MACI-Infant; (Wan et al., 2017; Wan et al., 2013)). As per a standard procedure, mothers were instructed to sit on a floor mat and to play with their infant as they would usually do at home with optional (supplied) toys for around 6 minutes. Recording commenced once mother and infant were settled into play. The infant scales of interest were: (1) Attentiveness to caregiver (infant's acceptance of and interest in their mother through eye contact or joint activity, face/body orientation and references to mother's activity; e.g. imitation); (2) positive affect (the overall amount and degree of positive emotional affective display by the

infant, in their vocal, facial and gestural/bodily expression), negative affect (the overall amount and degree of negative emotional affective display by the infant, in their vocal, facial and gestural/bodily expression); (3) Liveliness (the amount and level of physical activity, particularly if initiated by the infant).

Inter-rater agreement was determined based on the independent blind ratings of 14 (35%) randomly selected videos using intraclass correlation (single measures, absolute agreement definition). Rating agreement was moderate to excellent, based on thresholds by Koo & Lee, 2016 (infant attentiveness $r = 0.57$; $p = 0.009$; infant liveliness $r = 0.65$, $p = 0.001$; infant negative affect $r = 0.82$, $p < 0.001$; infant positive affect $r = 0.81$, $p < 0.001$).

Mothers' feedback questionnaire

To test our aim regarding acceptability, mothers completed an 8-item questionnaire after each measurement session about their satisfaction as well as their infants' feelings during the fNIRS procedure, as outlined in Table 4.2.

Participants

Ethical approval for the study was obtained from the UK National Health Service ethics committee (ref: 15/NW/0684).

Forty-two healthy mothers (without mental or major physical illness) with a healthy infant (20 girls and 20 boys) were recruited from 3 National Health Service Sure Start Children's Centres in Greater Manchester, UK. Mothers were aged 18 – 40 years, white fluent English speakers (to increase cultural homogeneity and reduce this as a possible confounder). Infants were born full-term, normal birth weight and were not diagnosed with medical, sensory or developmental disorder.

Strategies were carried out to familiarise mothers with fNIRS measurement sessions longitudinally. When approaching mothers, the researcher familiarised mothers with all study procedures as well as fNIRS data collection system. Mothers were talked through the measurement procedures on a number of occasions e.g. by phone during each arrangement contact. On the measurement day, mothers were also re-introduced to the fNIRS equipment before data collection. The mothers' queries were answered on each occasion.

To optimise retention of participants in the study, a three-step action plan was carried out. First, we arranged a testing time directly with the mother not earlier than a month in advance; second we followed this up a week before the agreed appointment; lastly, a reminder was sent to the mother the day before the test. We were also aware that this approach should not make participants feel overburdened or hassled by the researcher. Therefore, the second and third steps were accomplished by short messages which stated 'please reply to this message if you would like to make any change to our appointment'. In addition, before the study began, a time window of a month was set for each infant measurement to accommodate any re-arrangement request.

4.4 Results

The acceptability of the longitudinal measurement sessions to mothers and infants

A summary of participant retention over the three-time points of the longitudinal fNIRS measurements

Flow charts of participant retention are shown in Figure 4.2 and participant information in Table 4.1. Overall, study retention was high, as 38 out of 40 infants (95%) completed all 3 testing sessions. At 6 months, 40 infants were tested. At 9 months, N = 1 was lost to follow-up. At 12 months, a further N = 1 was lost to follow-up because the mother was unavailable.

Table 4.1 Participants' information

		Voice vs non-vocal task			Vocal emotion task		
		TOTAL	INC	EXC	TOTAL	INC	EXC
6 Months	No. of infants	40	31 (77.50%)	9 (22.50%)	40	29 (72.50%)	11 (27.50%)
	Sex	F:20; M:20	F:13; M:18	F:7; M: 2	F:20; M:20	F:15; M: 14	F:5; M: 6
	Mean age (SD) (days)	189.48 (9.27)	190.58 (9.59)	185.7 (7.30)	189.48 (9.27)	189.3 (9.66)	189.9 (8.58)
	Age range (days)	175-214	180-214	175-195	175-214	175-214	176-207
9 Months	No. of infants	39	35 (89.74%)	4 (10.26%)	39	30 (76.92%)	9 (23.08%)
	Sex	F:19; M:20	F:16; M:19	F:3; M: 1	F:19; M:20	F:11; M: 19	F:8; M: 1
	Mean age (SD) (days)	279.08 (9.46)	278.91 (8.70)	280.5 (16.50)	279.08 (9.46)	279.7 (8.81)	277.1 (11.74)
	Age range (days)	263-302	264-296	263-302	263-302	268-302	263-295
12 Months	No. of infants	38	33 (86.84%)	5 (13.16%)	38	29 (76.32%)	9 (23.68%)
	Sex	F:18; M:20	F:16; M:17	F:2; M: 3	F:18; M:20	F:14; M: 15	F:4; M: 5
	Mean age (SD) (days)	377.24 (8.61)	376.33 (8.64)	380.2 (6.61)	377.24 (8.61)	375.6 (7.84)	378.1 (8.58)
	Age range (days)	360-394	360-394	371-389	360-394	360-394	365-391

Note: TOTAL = total number of participants attended; INC = included participants; EXC = excluded participants; F = female; M = male.

Mothers' feedback

All the participated mothers filled the feedback form after the session (Table 4.2), except for 3 occasions (1 at 9 months, 2 at 12 months), which were due to the infant fussiness at the end of the session.

All the participating mothers were happy with the length of measurement; nearly 90% of mothers were satisfied with the appearance of the equipment; more than 95% of mothers agreed that the headgear was well-fitted for their infants' head. More than 95% of mothers thought the measurement session was interesting, and their infants were attracted by the sound and video played during the measurement. More than

77% of mothers agreed that their infants felt comfortable during the measurement. More than 75% of mothers thought their baby liked the measurement session. All the mothers filled in the feedback form and showed their willingness to take part in the follow up test session, including the two participants who were lost of follow-up have expressed their willingness to take part.

Despite the satisfaction of mothers with the fNIRS equipment, some mothers made recommendations for improving the testing environment: a more spacious room, which should allow the fNIRS system, a play mat, and the infant's stroller along with the infant's necessities to be stored. Mothers also suggested improvements of the fNIRS headgear settings: such as softer and lighter headgear, wireless headsets or a wire holder to help share the weight of the headset and they suggested we use headsets with a more colourful appearance.

Table 4.2 Summary of three time-points' Feedback Questionnaire

Mother's feedback	Age points	Strongly Disagree	Disagree	Agree	Strongly Agree
1. The length of time of the measurement procedure was reasonable	6m(N=40)	0	0	24(60%)	16(40%)
	9m(N=38)	0	0	19(50%)	19(50%)
	12m(N=36)	0	0	15(41.7%)	21(58.3%)
2. The fNIRS equipment looks good to me (e.g. the equipment didn't look too ugly or daunting)	6m(N=40)	1(2.5%)	4(10%)	24(60%)	11(27.5%)
	9m(N=38)	0	2(5.3%)	19(50%)	17(44.7%)
	12m(N=36)	0	2(5.6%)	21(58.3%)	13(36.1%)
3. The headgear fitted for my baby's head	6m(N=40)	0	2(5%)	20(50%)	18(45%)
	9m(N=38)	0	1(2.6%)	16(42.1%)	21(55.3%)
	12m(N=36)	0	0	19(52.8%)	17(47.2%)
4. The measurement session was interesting	6m(N=40)	0	2(5%)	24(60%)	14(35%)
	9m(N=38)	0	1(2.6%)	21(55.3%)	16(42.1%)
	12m(N=36)	0	0	21(58.3%)	15(41.7%)
5. My baby concentrated on the sounds, images and videos at most of the time	6m(N=40)	0	1(2.5%)	29(72.5%)	10(25%)
	9m(N=38)	0	1(2.6%)	30(78.9%)	7(18.4%)
	12m(N=36)	0	2(5.6%)	24(66.7%)	10(27.8%)
6. My baby felt comfortable during the measurement	6m(N=40)	0	4(10%)	25(62.5%)	11(27.5%)
	9m(N=38)	0	5(13.2%)	21(55.3%)	12(31.6%)
	12m(N=36)	0	8(22.2%)	17(47.2%)	11(30.6%)
7. My baby liked the measurement session	6m(N=40)	2(5%)	7(17.5%)	25(62.5%)	6(15%)
	9m(N=38)	0	8(21%)	25(65.8%)	5(13.2%)
	12m(N=36)	0	9(25%)	22(61.1%)	5(13.9%)
8. I feel happy to take part in the next measurement	6m(N=40)	0	0	17(42.5%)	23(57.5%)
	9m(N=38)	0	0	13(34.2%)	24(63.2%)

session

Feasibility of the longitudinal measurement sessions

The success in undertaking the developed fNRIS paradigm in the community clinic settings longitudinally in 38 healthy infants suggests that the experimental paradigm and data collection approach are practical and feasible.

Valid fNIRS data

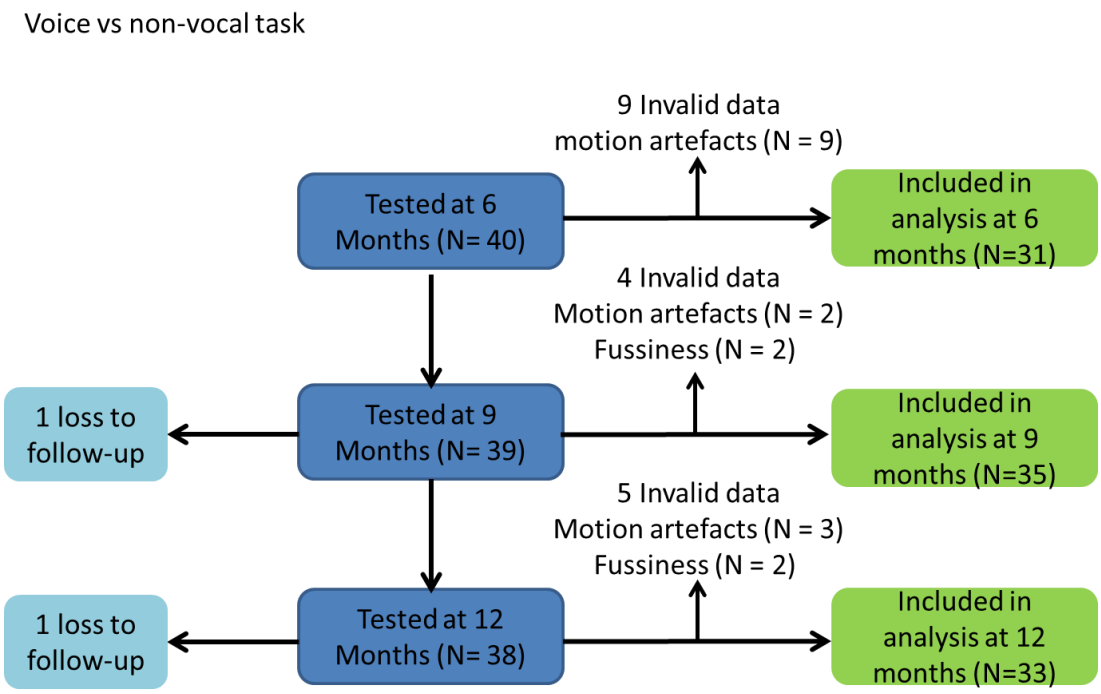
Infants' behaviours during fNIRS data collection were viewed and coded off-line for the voice versus non-vocal task and the vocal emotion task. According to the inclusion criterion, a minimum of 4 out of 8 trials per condition was needed for the inclusion of a valid dataset. This inclusion rate is consistent with criteria for inclusions in previous infant NIRS studies (Lloyd-Fox et al., 2010). Two flow charts in Figure 4.2 show the valid data included in the two tasks, the included and excluded participants' information is shown in Table 4.3.

In the present study, we tested 40, 39 and 38 participants at each age point respectively. The data loss was 15.3% for the voice versus non-vocal task, and 24.8% for the emotion task; both of them were smaller than the average data loss of 40% ((Lloyd-Fox et al., 2010), shown in Figure 4.2, Table 4.1).

In the voice versus non-vocal task, 31 datasets were included in the 6 months' time point; 35 datasets were included in the 9 months' time point; and 33 datasets were included in the 12 months' time point; 25 datasets consist the longitudinal sample (25 infants provided valid data at all time points). In the vocal emotion task, 29 datasets were included in the 6 months' time point; 30 datasets were included in the

9 months' time point; and 29 datasets were included in the 12 months' time point; 21 datasets consist the longitudinal sample (21 infants provided valid data at all time points).

We tested whether excluded infants differed from included infants in terms of behavioural characteristics (infant attentiveness to mother, positive affect, negative affect and liveliness, as evaluated from mother-infant play interaction) and language competence using independent sample t-tests, which found no significant differences (all the $p > 0.10$ and $t < 1.6$), suggesting the exclusion was due to the lack of attention to the task, but not seem to be the result of inherent or characteristic difference in the excluded group's liveliness, affect, attentional skills (as indicated by attentiveness to caregiver), language competence.



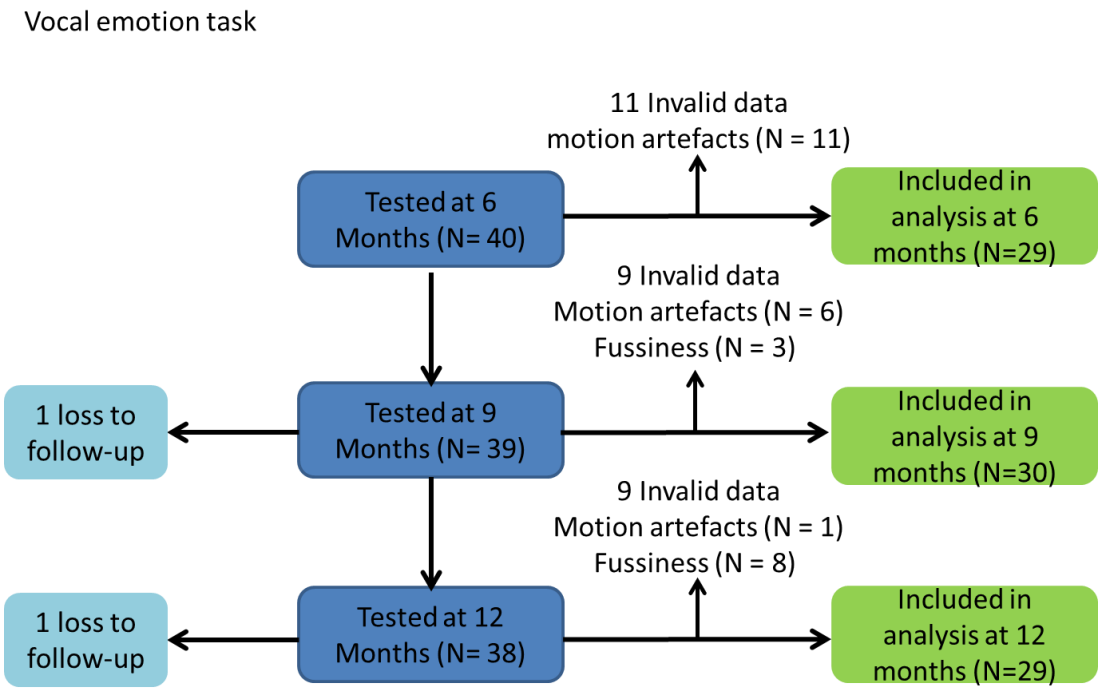


Figure 4.2 Flow charts of participant retention and infants included and excluded in the analysis at 6, 9, and 12 months.

Table 4.3 Participant characteristics of included and excluded infants at 6 months

	Voice vs non-vocal task		Vocal emotion task	
	Included	Excluded	Included	Excluded
No. of infants	31	9	29	11
Age (days, Mean ± SD)	190.5 ± 9.59	185.7 ± 7.30	189.3 ± 9.66	189.9 ± 8.58
Sex (Female: F; Male: M)	F: 13; M: 18	F: 7; M: 2	F: 15; M: 14	F: 5; M: 6
Infant attentiveness	3.94 ± 1.34	3.89 ± 1.96	4.07 ± 1.51	3.55 ± 1.37
Infant positive affect	2.39 ± 1.20	3.00 ± 1.00	2.52 ± 1.12	2.55 ± 1.37
Infant negative affect	2.90 ± 1.74	2.89 ± 2.03	2.62 ± 1.78	3.64 ± 1.63
Infant liveliness	4.39 ± 0.84	4.22 ± 1.09	4.48 ± 0.87	4.00 ± 0.89
Language score	101.29 ± 6.84	103.00 ± 4.24	102.21 ± 6.42	100.27 ± 6.21

Infant language development

Infants' language scores at each age point (6 months Mean ± SEM: 101.53 ± 1.05; 9 months Mean ± SEM: 98.90 ± 1.06; 12 months Mean ± SEM: 92.92 ± 0.93) were within the normal development score range (90 composite scores), suggesting the participants were normally developing infants.

4.5 Discussion

This study aimed to develop and subsequently to pilot the acceptability and feasibility of longitudinal fNIRS method in observing infants' neural development of voice and vocal emotion processing between 6 and 12 months. The developed fNIRS paradigm was accepted by the mothers and infants with high retention and satisfaction rate; and was piloted at three time points in 38 healthy infants in the community clinics with low data loss. The retention rate of the infants and their mothers across the three-time periods was 95%; this suggests that mothers were successfully engaged in the study and that the procedures were well tolerated. Mothers also reported a high satisfaction rate with the fNIRS measurement for their infants and were happy to engage and return for the longitudinal test. Our carefully developed experimental paradigm and measurement procedures allowed for very low data loss and secured 84.7% and 75.2% of fNIRS data for the voice and vocal emotion task respectively. This provides a strong case for the feasibility of our fNIRS data collection and analysis methods. By observing 6-month-infant behavioural characteristics (attentiveness to mother, infant affect and infant liveliness during the free play with their mothers), as well as language competence, no bias was found in excluding infants other than their attention to the task. In our view, this suggests that our data exclusion procedure is less likely to be influenced by the infant behaviour or language competence observed outside of the testing session. Furthermore, infants' language scores from the Bayley's scale suggested that participating infants did not have any language delays, at least in the first year.

Our study has shown its acceptability and feasibility in repeated testing of healthy infants at local community health centres. Strategies used both to engage infants and mothers, as well as in the design of the experimental paradigm and the headgear

settings likely contributed to the success of the longitudinal fNIRS data collection. Mothers became very familiar with the measurement procedure from the beginning; they expressed willingness and were happy to help with the study and to participate in the research. From the mothers' feedback, over 90% of them were satisfied with the experimental paradigm agreeing that the length of the measurement was reasonable; that the measurement session was interesting, and that their infants were concentrated on the sounds or videos most of the time. The feedback on these items suggests that the experimental design was acceptable to mothers and infants. The feedback items on the headgear settings also received a high satisfaction rate. More than 90% of mothers believed the headgear fitted their infants' heads well; fewer mothers thought their infants felt comfortable during the measurement, especially when infants were older (i.e. 9 and 12 months). This was because of infant perceived greater consciousness and awareness; as well as greater autonomy of the more mature infants (9 and 12 months) compared to the young infants (6 months).

In the present study, we have applied rigorous data analysis methods alongside a carefully constructed fNIRS design for a community setting. Together, this has allowed for low loss of data for the fNIRS infant data collection of 15.3% for voice versus non-vocal task and 24.8% for vocal emotion task, suggesting the feasible collection of sufficient high quality data. However, the infant data exclusion rate was high. Infants' low tolerance of being seated, and/ or wearing the headband (especially in infants who didn't like to wear a hat in daily life) led to incompletion of the tasks, and/or significant infant movement during data collection. This is also reflected in the higher data loss rate in the vocal emotion task than in the voice versus non-vocal task. Given that there were three conditions in the vocal emotion task, it required a longer data collection length to obtain sufficient repetition for each

condition, which means it was harder to fulfil the criteria (4 out of 8 trials in each condition) to include a dataset. This is a very common problem and challenge for the use of any functional imaging tool with awake infants and children. A suggestion to secure infant neuroimaging data is to design the study with a short recording time and use a relatively small number of channels (Aslin, 2012; Lloyd-Fox et al., 2010). Another ethical way to improve the data quality and to reduce infants' movements and fussiness is to design interesting experimental paradigms. An effective infant fNIRS experimental paradigm attracts infants' attention, reduces movements and has a short data collection time.

fNIRS is a relatively new neuroscience tool compared to EEG and fMRI. With advantages of relatively high temporal and spatial resolution, portability, low cost, quietness, and less requirement of head stabilisation, it has been widely used in observing infant neural responses to auditory, visual, and olfactory stimuli. An eye tracker is also a widely used psychophysiological tool in unrevealing infant cognitive processing by observing infants' looking behaviours, however, the interpretations of infant looking behaviours suffer from lack of accuracy (Aslin, 2007; Gredeback, Johnson, & von Hofsten, 2010; Lloyd-Fox, Blasi, et al., 2013). Future studies could make more use of the compatibility of fNIRS with eye-tracker, which will combine infants' looking behaviours with corresponding stimuli evoked neural responses, thereafter, to interpret infant underlying cognitive processes (Lloyd-Fox, Wu, et al., 2013; Urakawa, Takamoto, Ishikawa, Ono, & Nishijo, 2015).

As a result of this feasibility pilot's findings, fNIRS showed great potential and a promising future in infant studies of typical and atypical neural development. The present study tested the longitudinal application of fNIRS measurement in healthy infants on three occasions between the ages of 6 - 12 months. Our infant fNIRS

voice and vocal emotion paradigm reliably and feasibly attained adequate retention of mother-infant participants and was short enough to allow little loss of infant data as a result of inattention and intolerance to the paradigm. In addition, mothers reported the paradigm was interesting and acceptable to them. With such a satisfactory retention rate and with acceptability of the paradigm from infants and caregivers, it is expected that future studies in larger samples using broader infants groups using various types of auditory stimuli should be tested. It is also expected that this paradigm to be tested in vulnerable and hard to reach populations of mothers such as mothers with mental illness and their infants. Finally, findings from the present study call for some improvements in the customised headgear and especially development of infant-/ child-friendly, comfortable headgear as well as testing environments which are easier for mothers and infants. Future studies may also consider the use of the fNIRS paradigm in home settings.

Chapter 5 Is infant neural sensitivity to vocal emotion associated with mother-infant relational experience?

5.1 Abstract

An early understanding of others' vocal emotions provides infants with a distinct advantage for eliciting appropriate care from caregivers and for navigating their social world. Consistent with this notion, an emerging literature suggests that a temporal cortical response to the prosody of emotional speech can be observed in the first year of life. Furthermore, neural specialisation to vocal emotion in infancy may vary according to early experience. Neural sensitivity to emotional non-speech vocalisations was investigated in 29 six-month-old infants using near-infrared spectroscopy (fNIRS). Angry vocalisations evoked a stronger response in the left temporal region compared to neutral vocalisations, and happy vocalisations evoked increased activation in right temporal cortex compared to angry vocalisations; both effects survived statistical correction. Furthermore, the strength of the former effect (anger minus neutral) was positively correlated with the degree of directiveness used by mothers during independently videotaped play interaction with their infant. This first fNIRS study of infant vocal emotion processing implicates left and right temporal mechanisms similar to those found in adults and suggests that infants who experience more directive caregiving or social play may process vocal anger more strongly or preferentially by six months of age.

Keywords: emotional prosody, NIRS, mother-infant interaction, vocal emotion, vocal perception, superior temporal cortex

5.2 Introduction

The perception of vocal emotion has a special role in the formation of infant attachment and social development. From early infancy, infants rely heavily on emotional prosodic information from their caregivers, such as affective warmth or fear, to help guide behaviour to elicit care and, ultimately, to maintain safety from threat (Lohaus, Keller, Ball, Elben, & Voelker, 2001; Mumme et al., 1996). Such auditory processing is likely to be rudimentary in the early months (Trevarthen, 2017), but at around 5 months, experimental studies suggest that infants are able to discriminate vocal affective expressions from a non-caregiver woman's voice (Caron, Caron, & MacLean, 1988; Fernald, 1993; Flom & Bahrick, 2007; Walker-Andrews & Grolnick, 1983). Soon after, infants develop the ability to 'social reference' known adults to gain vocal and facial information on how to react to ambiguous, potentially threatening situations (Striano & Rochat, 2000; Vaish & Striano, 2004). Infants also seem to use the voice more than facial expression as a reliable source of emotional information (Caron et al., 1988; Vaish & Striano, 2004). This would logically follow given that young infants cannot always see others' facial cues because of their relative immobility.

Several functional neuroimaging studies of infant voice processing have been conducted and suggest that the temporal and/or frontal cortical regions are sensitive to voice from between 3 and 7 months of age (Blasi et al., 2015; Blasi et al., 2011; Dehaene-Lambertz, Dehaene, & Hertz-Pannier, 2002; Grossmann, Oberecker, et al., 2010; Lloyd-Fox et al., 2012). Two of these studies further report that emotional prosody elicited a stronger response in voice-sensitive regions (Grossmann, Oberecker, et al., 2010; Missana et al., 2017). These findings broadly mirror the timeline suggested by looking time studies (Fernald, 1993; Flom & Bahrick, 2007;

Walker-Andrews & Grolnick, 1983), and may reveal an early version of the adult temporo-frontal vocal emotion processing pathway (Alba-Ferrara, Ellison, & Mitchell, 2012; Alba-Ferrara, Hausmann, Mitchell, & Weis, 2011; Ross & Monnot, 2011; Zhang et al., 2018) which prioritises the processing of emotional (Ethofer et al., 2012; Pell et al., 2015) (especially negative (Frühholz & Grandjean, 2013; Grandjean et al., 2005; Pell et al., 2015)) prosody. In adults, the relatively stronger neural response to vocal negativity likely reflects an attentional bias for negative stimuli (Peeters & Czapinski, 1990); furthermore, children also show this negativity bias in a range of socio-communicative skills, such as social referencing and language acquisition (Chronaki et al., 2012; Vaish et al., 2008).

Studies on the infant processing of emotional speech show an increased temporal activation in response to angry and happy speech compared to neutral speech in 7- to 8-month-old infants (Grossmann, Oberecker, et al., 2010; Grossmann et al., 2005; Missana et al., 2017). However, infants may be able to detect or discriminate emotion within non-speech vocalisations earlier than in speech, but infant Near Infrared Spectroscopy (NIRS) studies so far have focused on speech processing (Blasi et al., 2011; Lloyd-Fox et al., 2012) or older infants (Grossmann, Oberecker, et al., 2010; Grossmann et al., 2005; Missana et al., 2017). Two functional magnetic resonance imaging (fMRI) studies of non-speech prosody processing in 3- to 7-month asleep infants reported stronger insular and frontal activation to sad in contrast with neutral vocalisations (Blasi et al., 2015; Blasi et al., 2011). The present study sought to extend our current understanding of the emergence of vocal emotion sensitivity by using non-speech stimuli with six-month-old infants.

From the earliest months of life, infants begin to regulate their own behaviour and emotions according to the quality of care they receive (Trevarthen, 2017). An

infant's emerging ability to process and differentiate vocal emotions may play an important role and may be influenced by the affective tendencies of the infant's primary caregiver (usually the mother) that accompany her caregiving or interactive style. Evidence from elecotroencephalogram (EEG) studies has linked maternal caregiving behaviour to longitudinal changes in infants' neural networks serving attentional processes (Bernier et al., 2016; Swingler et al., 2017). While maternal sensitivity is typically characterised by positive vocal cues from high emotional warmth (Lohaus et al., 2001; Lohaus, Keller, Ball, Voelker, & Elben, 2004), infants with sensitively responsive mothers may prioritise attention to all strong emotional information as they have learned through experience that others' vocalisations (and their own) are meaningful and relevant for understanding and navigating their interpersonal relationships and environment. On the other hand, maternal directiveness, a behavioural tendency typically characterised by vocal and behavioural demands, intrusions and/or critical utterances, often takes emotionally negative forms, acting as an internal guide to acceptable behaviour, therefore, giving rise over time to a possible neural bias towards the processing of negative prosody. One study to date has attempted to link maternal behaviour (intrusiveness) with 3- to 7-month-old infant neural vocal response - in infants at high and low risk of autism, and found no significant linear relationship (Blasi et al., 2015).

The current study investigated 6-month-old infant hemodynamic response to emotional prosody in non-speech vocalisations. The key objective was to test whether there was increased neural activation in the temporal region in response to non-speech emotional (angry, happy) compared to neutral vocalisations, as found in adult studies. Secondly, we wished to explore whether individual variation in neural response to emotional prosody would correlate with infants' real-life caregiving

experience, as measured from independently video-recorded mother-infant play interactions. We predicted that maternal sensitivity would be associated with higher neural responsiveness in the temporal region to emotional prosody (happy minus angry, anger minus neutral) while maternal directiveness would be associated with higher neural responsiveness to anger prosody (anger minus neutral).

5.3 Materials and methods

Participants

Forty healthy 6-month-old infants of healthy mothers were recruited from three community health centres in Manchester, UK. The final sample consisted of 29 infants (15 girls, 14 boys, aged between 175 and 214 days, M = 189 days, SD = 9.66), as 11 infants did not meet the minimum 4 out of 8 trials per experimental condition as a result of motion artefacts. This attrition rate is within the standard range for infant NIRS studies (Lloyd-Fox et al., 2010). A power analysis using the G*power program (Faul, Erdfelder, Lang, & Buchner, 2007) indicated that a sample size of N = 29 would give 92% power to achieve an effect size of 0.59 (which equals to eta-squared of 0.26). All infants were born full term (37–42 weeks gestation) except n=1 born at 36 weeks gestation (corrected age used), at normal birth weight (>2500g), and had no hearing difficulties according to parent report. The UK National Health Service ethics committee approved the study (ref: 15/NW/0684).

Experimental paradigm and procedure

During the fNIRS experimental procedure (Figure 5.1), infants sat on their mother's lap facing a laptop and wearing the NIRS headband. The task started with a 20-sec rest period, followed by a 5-sec trial presented through loudspeakers (SPL = 70 dB). A 5-sec silent cartoon video was shown during each trial to attract infant attention and reduce motion artefact, as consistent with previous research (Grossmann, Oberecker, et al., 2010). After each trial, a 10-sec silent blurred cartoon baseline was presented. The task was presented with PsychoPy software (Peirce, 2007). Each condition (angry, happy and neutral) was presented 8 times amounting to a total

number of 24 trials. The same emotional expression did not occur consecutively. The testing session lasted 6 minutes, 20 seconds.

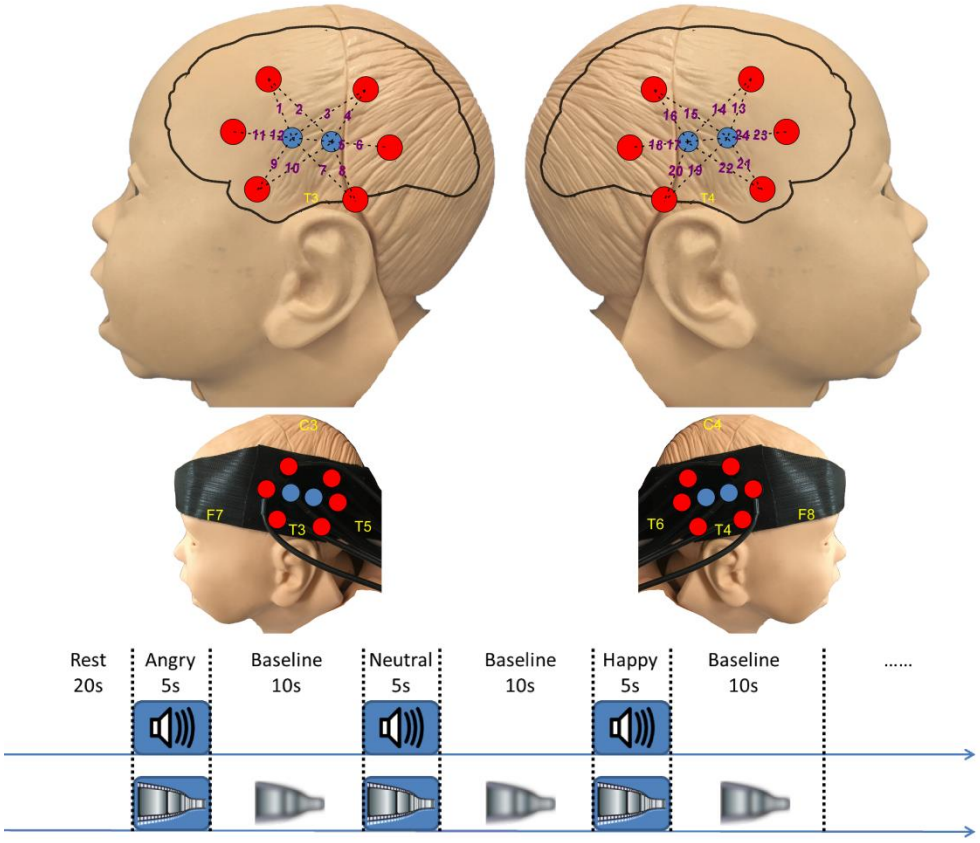


Figure 5.1. Study experimental task design and channel distribution. The head model illustrates the source-detector distribution where red dots represent sources (6 in each hemisphere) and blue dots represent detectors (2 in each hemisphere). Sources and detectors form 12 recording channels in each hemisphere, which are marked in red numbers (upper head models) and are held by Velcro head band. The channel locations with respect to the 10-20 system are marked in red (middle head models). The bottom streamline demonstrates the timeline of the experimental task stimulus presentation and baseline.

Vocal stimuli

The stimulus material consisted of 15 adult female, non-speech vocalisations of angry, happy and neutral prosody (interjection 'ah') from a well-validated battery of vocal emotional expressions (Maurage et al., 2007). This battery has high internal consistency for each emotion set as well as high levels of specificity (independence between the ratings in the different emotion sets (Maurage et al., 2007). These stimuli have been validated in previous research in children (Chronaki et al., 2012). Five normalised stimuli, each lasting 1 sec, from the same expression category were selected and combined to form a 5-sec trial. All vocal stimuli were normalised with Praat sound-analysis software (Boersma & van Heuven, 2001) to the same duration of 1 000 ms and mean intensity of 73 dB.

fNIRS data acquisition

During functional cerebral activation, the NIRS setting measures the attenuation of light that corresponds to changes in haemoglobin concentration (Villringer & Chance, 1997). In the present study, infants' cerebral responses were recorded with a multichannel NIRS data collection system. The system was built by Biomedical Optics Research Laboratory (Dept. of Medical Physics and Bioengineering, University College London) and applied with 780nm and 850nm continuous wavelengths and 10Hz sampling rate (Everdell et al., 2005). Two detectors and six sources formed 12 source-detector pairs in each hemisphere and were distributed at temporal regions, which have been shown to be voice sensitive in previous research in infants (Grossmann, Oberecker, et al., 2010; Lloyd-Fox et al., 2012; Pena et al., 2003; Taga & Asakawa, 2007) and adults (Belin et al., 2000; Ethofer, Anders, Wiethoff, et al., 2006; Grandjean et al., 2005). To achieve the best spatial sensitivity

profile for infants (Fukui et al., 2003), the distances between source and detectors were fixed between 1.5 and 2.5 cm. Channels were distributed according to the 10-20 system and attached to a custom-made Velcro headband. The headband was adjusted by calculating the distance between the glabella and the ear, ensuring that T3 and T4 are between the two bottom sources in each hemisphere. The locations of the channels and the channel positions with respect to the 10-20 system are presented in Figure 5.1.

fNIRS Data analysis

Video-recorded infant behaviour during the task was viewed to code whether the infant attended to the screen without large motion artefacts. Four out of eight trials per condition were set as a criterion for inclusion of each infant dataset.

All the datasets analysed were filtered at 0.01 to 0.5Hz with 3rd order Butterworth filter, to eliminate slow drifts, instrument noise and physiological artefacts, such as heartbeats (Cooper et al., 2012; Fox et al., 2013; Grossmann, Oberecker, et al., 2010). The remaining artefacts were identified on a channel by channel basis with the algorithm 'hmrMotionArtifactByChannel' implemented in the HOMER2 NIRS toolbox (version 2.1, http://homer-fnirs.org/, Huppert et al., 2009(Huppert et al., 2009)). Within the time interval (tMotion), if the change of the signal amplitude exceeded the threshold (AMPthresh) or the standard deviation changes were greater than a factor (STDEVthresh) multiplied by the original channel standard deviation, the time period (tMask time before and after the motion artefact) was marked as artefact. The time period of motion artefact within the channel was corrected with a cubic spline interpolation algorithm with p set to 0.99 as recommended (Cooper et al., 2012; Scholkmann et al., 2010). Since the algorithm works on a channel by channel basis, the actual standard deviation threshold for the motion artefact varies

according to the standard deviation of the original channel; the setting of the STDEVthresh is the multiplication factor rather than a fixed threshold (i.e. in the current study the standard deviation threshold is 20*standard deviation of the channel). This means that the standard deviation threshold varies from channel to channel and subject to subject. All the values were set as follows: tMotion=5s; tMask=1s; STDEVthresh=20; AMPthresh=5.

After pre-processing, data were converted to Oxy- and Deoxy-Haemoglobin concentration changes (ΔHbO^2 and ΔHbR) in HOMER2 and averaged across trials in the same emotion condition within each dataset, with the time window of 1 sec before and 15s after the stimulation onset. The averaged time course of each channel was corrected by subtracting the mean of the 1 sec before the stimulation. The analysis focused on ΔHbO^2 as the most sensitive indicator of changes in cerebral blood flow. Kolmogorov-Smirnov tests were conducted to test the normality of the haemodynamic time courses, results showed that p values were larger than 0.05, which indicate that haemodynamic data were normally distributed. Based on earlier work showing that the haemodynamic response reaches the peak around 2 to 4 sec post stimulus (Brigadoi et al., 2014), we targeted a time window of 2 sec to 9 sec after stimulus onset. Mean amplitudes of cortical haemodynamic responses (ΔHbO^2 and ΔHbR waveforms) were averaged over the time window of 2 sec to 9 sec after stimulus onset. The averaged haemodynamic responses to the expression conditions (angry, happy and neutral) were evaluated with repeated measures ANOVA and post-hoc pairwise comparisons. We calculated partial eta-squared (Cohen, 1973; Kennedy, 1970) to estimate the effect sizes for the main effect of emotion as well as for contrasts. Partial eta-squared takes values between 0 and 1. Values of 0.02, 0.13

and 0.26 are indicative of a small, medium, and large effect size, respectively (Murphy, Myors, & Wolach, 2014).

FDR correction for multiple comparisons was applied (Benjamini & Hochberg, 1995), consistent with other recent infant studies (Blasi et al., 2015; Lloyd-Fox et al., 2017). As the detector array covers a large area of the infant's brain, we do not expect all detectors to cover brain areas that are responding to our stimulation. Therefore, we only include channels that show a response to the stimulus paradigm. Within identified emotional sensitive channels, comparisons were corrected with the following steps: (i) A number of p values obtained from post-hoc comparisons were arranged with ascending order (from the smallest to the largest) with an order number index, (ii) Adjusted α values were calculated with the equation αadjust = (order index/total number of comparisons)*0.05 and (iii) A comparison was deemed to be significant if the pairwise p value is smaller than the adjusted α value (αadjust).

Maternal interaction behaviour

A 6-min mother-infant free play interaction session was video recorded during the same visit following the fNIRS session. Mothers were asked to sit on a floor mat and play with their infant as they would normally do at home optionally using a small set of (supplied) toys. Recording commenced once mother and infant were settled into play. The videos were later rated by a trained rater (blind to family information and study aims) using the Manchester Assessment of Caregiver-Infant Interaction (MACI (Wan, 2015; Wan et al., 2017)), a validated rating scheme comprising eight 7-point scales suitable for use with normative and at-risk groups (Wan et al., 2014; Wan et al., 2012). The current study focused on the two caregiver scales, which are normally distributed in a non-clinical population: (1) sensitivity: the degree to which the infant's behaviour and state were met by prompt, appropriate and attuned

responses to meet the infant's immediate and developmental needs, including an attentive attitude, appropriate engagement and the provision of support and structuring in response to infant behaviour and a lack of behaviour); (2) directiveness (reversed in this study from the 'nondirectiveness' scale for ease of interpretation): the degree of restrictive or controlling behaviour as characterised by demanding, intrusive, critical and/or other controlling behaviours or comments directed at the infant). Inter-rater agreement based on the independent blind ratings of 12 (30%) video recorded interactions was high (intraclass correlation using single measures, absolute agreement definition: sensitivity: $r = 0.84$; directiveness $r = 0.70$; both $p < 0.001$).

5.4 Results

Repeated measures ANOVA with emotion (angry, happy and neutral) as the within-subject factor revealed 3 channels that were sensitive to emotional prosody in ΔHbO^2: Channel 2 in the left hemisphere ($F (2, 56) = 3.38$, $p = .040$, $\eta_p^2 = .11$), channel 14 in the right hemisphere ($F (2, 56) = 3.24$, $p = .047$, $\eta_p^2 = .10$) and channel 16 in the right hemisphere ($F (2, 56) = 4.38$, $p = .017$, $\eta_p^2 = .14$) (Table 5.1). Pairwise comparisons showed significant increased ΔHbO^2 on hearing angry compared to neutral voices (channel 2: $F (1, 28) = 9.76$, $p = .004$, $\eta_p^2 = .26$) and happy compared to angry voices (channel 16: $F (1, 28) = 8.26$, $p = .008$, $\eta_p^2 = .23$) which survived FDR correction (Figure 5.2). Two further pairwise comparisons did not survive FDR correction (Table 5.1): happy compared to neutral voices (channel 14: $F (1, 28) = 5.62$, $p = .025$, $\eta_p^2 = .17$) and happy compared to angry voices (Channel 14: $F (1, 28) = 4.26$, $p = .048$, $\eta_p^2 = .13$).

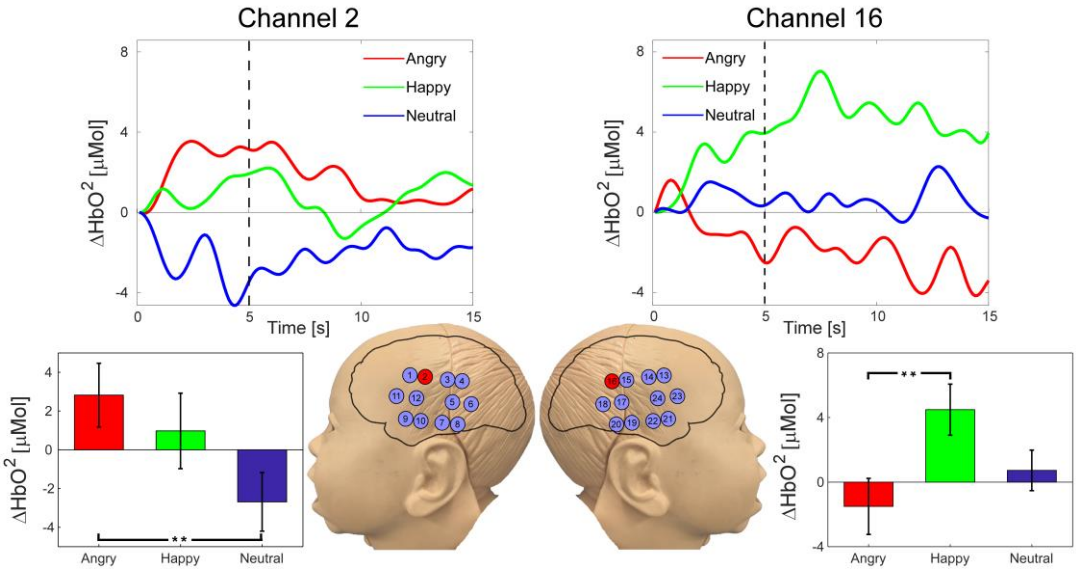


Figure 5.2 Averaged time courses of ΔHbO^2 in channel 2 and channel 16. Averaged time courses of ΔHbO^2 across all datasets in channel 2 and channel 16 per vocal emotion (angry in red, happy in green and neutral in blue) in the time period of 15 sec (5 sec stimulus and 10 sec baseline). The channel location is marked in red in the infant head model. The stimulus offset is marked by the dashed line (at 5 sec). The time (in sec) and change in amplitude (μMol) are in the x and y axis respectively. The mean and SEM value of ΔHbO^2 in each channel per vocal emotion is shown in the bar plot. ** represents the significant ($p < 0.01$) pairwise comparisons after FDR correction.

Table 5.1. Summary of emotion (Angry, Happy, Neutral) ANOVA effects on ΔHbO^2

Channel	Emotion	Mean ± SEM	ANOVA			Pairwise Comparisons				Adjusted α value
			F	p	Partial Eta-squared	Comparison[a] (A, H and N)	F	p	Partial Eta-squared	α_{adjust}
2	Angry	2.82±1.6	3.38	0.040	0.11	A > H	0.56	0.462	0.02	0.044
	Happy	0.97±1.9				A > N	9.76	**0.004***	0.26	0.006
	Neutral	-2.68±1.5				H > N	2.86	0.102	0.10	0.033
14	Angry	0.29±1.34	3.24	0.047	0.10	H > A	4.26	0.048	0.13	0.022
	Happy	4.02±1.67				A > N	0.11	0.746	0.004	0.050
	Neutral	-0.33±1.24				H > N	5.62	0.025	0.17	0.017
16	Angry	-1.51±1.74	4.38	0.017	0.14	H > A	8.26	**0.008***	0.23	0.011
	Happy	4.49±1.58				N > A	1.10	0.300	0.04	0.039
	Neutral	0.73±1.25				H > N	3.80	0.060	0.12	0.028

* Comparison survived FDR correction (comparisons for which the p values were smaller than the adjusted α value).

[a] A=Angry, H=Happy, N=Neutral

DeoxyHb concentration changes complemented the ΔHbO^2: 2 channels were sensitive to emotional prosody and survived FDR correction: a significant effect of emotion (left hemisphere: channel 2: F (2, 56) = 4.04, p = .020, η_p^2 = .13), particularly in response to angry compared to neutral voice (F (1, 28) = 10.26, p = .003, η_p^2 = .27) and a significant effect of emotion in channel 16 in the right hemisphere (F (2, 56) = 3.62, p = .030, η_p^2 = .11) in response to happy compared to angry voice (F (1, 28) = 7.45, p = .010, η_p^2 = .21).

Bivariate correlations tested whether ΔHbO^2 concentration changes (emotion minus neutral ΔHbO^2) in the two significant vocal emotion-sensitive areas that survived FDR correction (angry minus neutral ΔHbO^2 in left hemisphere channel 2; happy minus angry ΔHbO^2 in right hemisphere channel 16) were associated with maternal interactive behaviour ratings (sensitivity, directiveness). Although ΔHbO^2 in neither region was associated with maternal sensitive responsiveness, increased activation to angry minus neutral prosody was negatively correlated with maternal directiveness: r = 0.406, p = 0.029 (Figure 5.3).

There was a lack of significant correlation coefficient, after excluding the data from both ends (i.e. maternal directiveness scored = 1 or 7). However, this does not necessarily make the findings any less valid. As the sample was not selected based on maternal behaviour (i.e. this was a secondary objective), the majority of mothers were – as expected - rated in the middle range for (non)directiveness. The findings suggest that recruiting a (larger) sample specifically to test out associations with maternal behaviour may produce stronger effects.

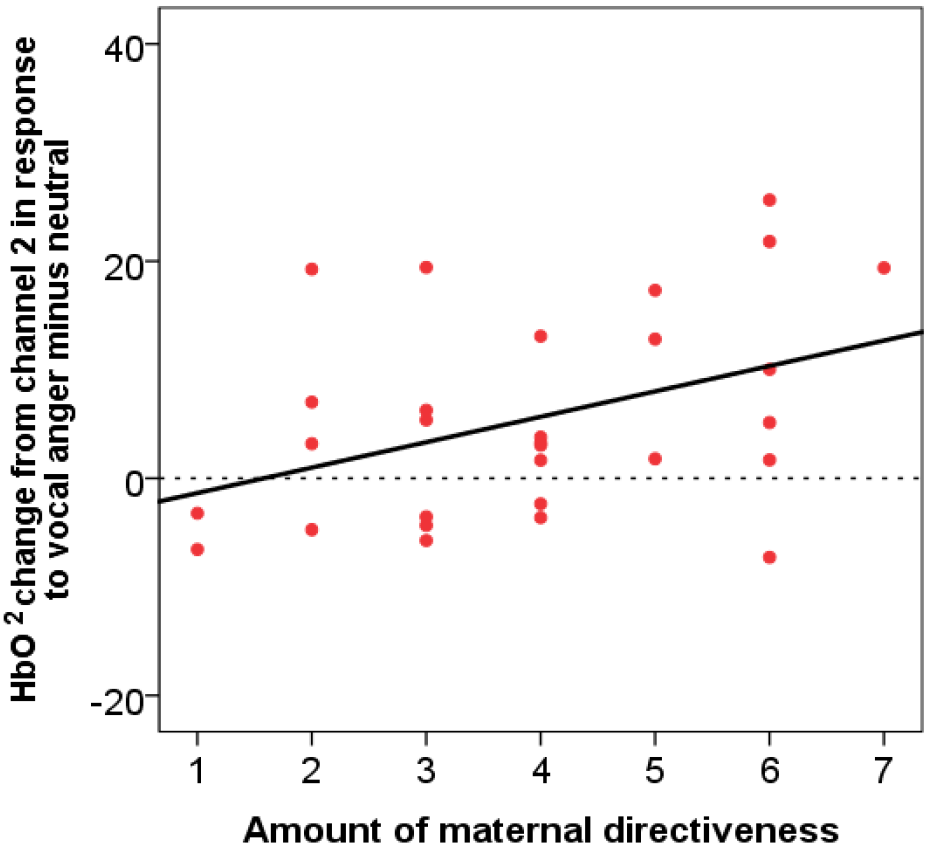


Figure 5.3. Association between neural responses to angry minus neutral prosody and maternal directiveness. Infant neural response to angry minus neutral vocalisations (y axis) increases linearly with independent ratings of how directive mothers were towards the infant during play interaction (x axis). The black hard line represents the best fit of HbO^2 change for each rating on the maternal directiveness scale. The dotted line represents no HbO^2 change.

5.5 Discussion

This is the first study of infant neural processing of emotional non-speech prosody to demonstrate the heightened recruitment of left and right temporal cortex at six months in response to vocal emotion. The findings also offer preliminary evidence of stronger negative vocal discrimination in the temporal region linked with early social or caregiving experience. Specifically, vocal anger evoked stronger responses in the infants' left anterior superior temporal cortex (STC) compared to neutral prosody and, as predicted, infants with stronger activation in this vocal anger-sensitive region experienced more maternal directive social interactions. By contrast, happy prosody evoked increased responses in the right posterior (and possibly anterior) STC compared to angry prosody. However, the strength of this response in the right temporal cortex was not associated with our measures of maternal care behaviour using mother-infant interaction ratings.

Consistent with previous behavioural and neuroimaging research (Blasi et al., 2015; Blasi et al., 2011; Grossmann, Oberecker, et al., 2010; Grossmann et al., 2005), our findings suggest that 6-month-old infants are able to differentiate emotional from neutral sounds and to distinguish basic emotions, irrespective of whether this is contained in speech or not. However, our main findings extend on the literature by demonstrating a left and right superior temporal activation pattern in response to non-speech vocal emotions, suggesting that at least part of the temporo-frontal network that is recruited in adult vocal emotion processing (Alba-Ferrara et al., 2012; Alba-Ferrara et al., 2011; Ross & Monnot, 2011; Zhang et al., 2018) is already developed by 6 months of age. Although vocal emotion processing was informed by adult brain lesion studies to be lateralised to the right hemisphere (Alba-Ferrara et al., 2012; Ross & Monnot, 2011; Zhang et al., 2018), current evidence supports the

crucial role of bilateral superior temporal and inferior frontal regions (Ethofer et al., 2012; Frühholz, Ceravolo, & Grandjean, 2012; Frühholz & Grandjean, 2013; Witteman, Van Heuven, & Schiller, 2012), based on paradigms involving varied stimuli (speech and semantic meanings) and task requirements (implicit and explicit tasks). While other infant brain studies implicate the temporal cortices (Grossmann, Oberecker, et al., 2010; Grossmann et al., 2005; Grossmann, Vaish, et al., 2013), broadly supporting the temporo-frontal network, our study specifically yielded STC activation. STC is involved in voice and speech perception from as young as 3 months (Dehaene-Lambertz et al., 2002; Grossmann, Oberecker, et al., 2010; Kuhl et al., 2014), but previous studies have not demonstrated heightened activation in this area based on emotional valence. By contrast, Blasi et al. (2015) reported right STC activation in their fMRI study of 3- to 7- month infants in response to neutral greater than sad non-speech prosody (Blasi et al., 2015).

According to Interactive Specialisation theory (Johnson et al., 2009), the superior temporal cortices form part of the 'social brain' that undergoes a "fine tuning" process, dependent on experience, into specialised functions. Our findings suggest that infant neural sensitivity to vocal emotions, although present by 6 months, may not yet be stable or specialised. Furthermore, at 6 months, neural responses to emotional prosody are likely to be highly sensitive to how the emotion is carried, i.e. in speech or otherwise. Studies involving speech processing may be confounded by individual differences in the semantic understanding of speech content (i.e. receptive language), since many infants at 6 months show much more understanding of words than previously thought (Bergelson & Swingley, 2012, 2015). The current study focused on non-speech prosody as the medium for carrying emotional stimuli, reflecting how caregivers commonly express themselves to preverbal infants.

In addition, we report that hearing angry vocalisations evoked a response localised to the left anterior STC. Negative vocalisations act as a cue to react to or avoid a dangerous situation (Bowlby, 1969; Dykas & Cassidy, 2011) and may reflect a general negativity attentional bias that is seen in adults (Chronaki, Hadwin, Garner, Maurage, & Sonuga-Barke, 2015; Ito, Larsen, Smith, & Cacioppo, 1998; Peeters & Czapinski, 1990; Schupp et al., 2004; Stifter & Fox, 1987). Vocal anger is a particularly salient signal that may serve a central evolutionary adaptive function (Vaish et al., 2008). However, this brain region also showed a modest response to happy prosody, which could suggest that this region responds to different emotions or varying acoustic properties at six months but is starting to 'favour' vocal anger. Neural sensitivity to angry compared to neutral voice has been reported in other infant studies (Grossmann, Oberecker, et al., 2010; Grossmann et al., 2005; Grossmann, Vaish, et al., 2013), raising the question of whether a prioritised neural response to threatening vocal information may be innate, consistent with evolutionary explanations (Vaish et al., 2008) that would therefore not require infants to have any conceptual understanding of such emotions. Consistent with this possibility, Zhang et al. (2014) studied the automatic brain responses of sleeping neonates and found fronto-central discrimination between angry and fearful vocal syllables but only when the neonates were in an active sleep state (Zhang et al., 2014). Taken together, the findings may suggest early cerebral specialisation for the automatic perception of threat-related emotional voices, which may become a conscious attentional process at 6 months of age, as reflected in the recruitment of the left STC.

The neural response to happy compared with angry prosody in the right posterior STC that we report is consistent with a right lateralised effect found in other infant

and adult studies on vocal emotion (Alba-Ferrara et al., 2012; Grossmann, Oberecker, et al., 2010; Ross & Monnot, 2011; Zhang et al., 2018). We expected a neural sensitivity to happy prosody on the basis that infants prefer infant-directed speech, which is often characterised by exaggerated positive affect (Saint-Georges et al., 2013; Singh et al., 2002), and because of its likely importance for facilitating bonding and secure attachment (Leigh et al., 2011; Lohaus et al., 2001; Tronick, 1989). However, in our study, neural responsiveness to happy prosody was not significant when compared to that elicited in response to neutral prosody, and the effect in channel 14 (anterior STC) did not survive FDR correction, possibly suggesting that this localised happy-specific sensitivity is not yet stable developmentally or may only be present in a subgroup.

With respect to our second objective, we found that infant neural sensitivity to prosodic anger was associated with maternal directive interactions. This was in the direction predicted, although this was not consistent with findings from Blasi et al.'s study (2015). The distinct findings may be derived from three aspects: (1) the age group of infants were different in the two studies. We have assessed infants at the age of 6 months, while infants' ages in Blasi et al.'s study were between 4 and 7 months. Infants are dramatically developing within the first year of life, different age groups of infants may have different developmental level of social-emotion and neural processing of vocalisations. (2) We have assessed infant neural processing of female produced interjection of 'ah' in an emotionally happy, angry and neutral voice, while stimuli in Blasi et al.'s study were crying, laughter sounds. The difference in the stimuli may contribute to the difference in the results. (3) Methodological differences may also contribute to the inconsistent findings. fNIRS has limited spatial resolution compared to fMRI, in addition, the present study only

covered temporal regions, activations in the frontal regions to emotional stimuli were not captured; while Blasi et al.'s study was empowered by whole-head, high spatial resolution fMRI.

Maternal directiveness typically involves vocal and behavioural demands, intrusions and/or critical utterances. Although not all such directiveness carries vocal negativity, being the recipient of high caregiver directiveness is likely to involve appraising negative emotion more often as a guide to acceptable behaviour, giving rise to a neural bias towards the processing of negative prosody. Our findings require replication in a larger sample but provide preliminary evidence that may suggest that early social experience in the form of directive caregiver interactions may promote cortical specialisation in vocal anger. This increased sensitivity to vocal anger could also be a result of prior exposure to anger, or other negative prosody, or may reflect environmental stress (or both). A previous study linked greater infant neural responses to angry relative to neutral speech with environmental stress in the form of high interparental conflict (Graham, Fisher, & Pfeifer, 2013), although these findings do not preclude an explanation of previous vocal anger exposure. Alternatively, given that mother-infant interaction effects are bidirectional (Cohn & Tronick, 1988; Eisenberg et al., 2010), this left STC activation in response to vocal anger may underlie a broader infant behavioural repertoire that might elicit more directive maternal responses. For example, certain groups, such as young children with behavioural difficulties, or those at risk of developmental delay and/or disability may experience more directiveness (Blacher, Baker, & Kaladjian, 2013; Brown, McIntyre, Crnic, Baker, & Blacher, 2011; Green, Caplan, & Baker, 2014; Wan et al., 2012) as mothers may (understandably) try to elicit social engagement and particular kinds of response. This early attentional bias to negative vocalisations may be an

adaptive function of the early social or caregiving experience; however, in the longer-term, high levels of caregiver directiveness may lead to desensitisation to negative emotions, reflected behaviourally in child passivity, which is likely to be detrimental to emotion and attention regulation, as well as language development (Akhtar, Dunham, & Dunham, 1991; Mathis & Bierman, 2015).

We predicted that infants who receive more sensitive caregiving (versus lower sensitivity) may show earlier, or stronger, neural response to emotional prosody (whether positive or otherwise), as they are likely to have experienced positive vocal positivity and warmth and a (Lohaus et al., 2001; Lohaus et al., 2004) and understand vocal information as salient for guiding action. However, we found no support for this in our six month old infants. Since participants in the present study were healthy mothers and infants, it is possible that significant effects are seen only in a clinical or at-risk group (for whom high maternal sensitivity can be a protective factor). Another consideration is that high maternal sensitivity does not always entail obviously high positive vocal affect, but rather affect is attuned (i.e. well-modulated to infant affect) and generally well matched (i.e. high positive maternal affect would not be sensitive in the face of infant distress or neutral affect) (Jonsson et al., 2001). Furthermore, if this vocal negativity processing effect reflects a largely innate 'negativity bias' (as discussed earlier), then it would not be influenced by maternal sensitivity.

The present study made several methodological improvements on most previous infant studies, including a larger overall sample size and statistical correction for multiple comparisons. Notwithstanding this, several methodological considerations must be taken into account. First, we observed emotion effects on brain responses in the temporal cortical regions and did not investigate the involvement of other (e.g.

frontal) regions implicated in vocal processing (Schirmer & Kotz, 2006). Secondly, since we used only angry and happy emotional stimuli, the anger related effects reported may result from emotional negativity in general, rather than being anger-specific. Thirdly, distinctive neural patterns to emotional categories do not necessarily suggest a conceptual understanding of emotions by infants, although experimental findings indicate that discrete emotions are at least paired with different kinds of infant responses or preferences (Caron et al., 1988; Fernald, 1993; Flom & Bahrick, 2007; Walker-Andrews & Grolnick, 1983), suggesting a level of evaluative appraisal rather than solely an acoustic analysis of pitch characteristics by the infant. A combined fNIRS and experimental approach may provide further validity as to whether infant neural responses to vocal emotions correspond to infant behaviours. Fourthly, the infant's task was hearing an unfamiliar adult female voice; neural sensitivity to vocal emotion at six months may not yet have consistently generalised from their regular caregivers. Infants may even perceive unfamiliar voices, including happy ones, as potentially threatening, which would have led to a reduction of effects seen. Finally, since maternal and infant anger were not directly measured in this study, whether more directive caregivers actually used more anger vocal expressions and/or whether their infants experienced more anger (or irritation) as a result of their social interactions is unknown.

In conclusion, consistent with the important function of vocal emotion perception in the first year of life for guiding attachment and social learning processes, we found novel evidence in infancy that prosodic anger (compared with neutral voices) elicited STC activation, which has been implicated in adult vocal emotion perception. However, the lack of discrete, emotion-specific neural responses that withstood statistical correction for multiple comparisons suggests that such cortical

specialisation is developing and not yet robust by 6 months of age. Furthermore, we report the first preliminary evidence of a link between an individual infant brain's neural responsivity to vocal anger and maternal directiveness in a healthy sample. Further investigation of this association may help us understand better the role of early experience on vocal perception as a building block for communicative and socioemotional development. The current paradigm could be developed to evaluate the effectiveness of parenting interventions on neural sensitivity to vocal emotion in healthy and at-risk groups early in infancy. Such interventions may be designed to target caregiver directiveness to help unravel the directionality of effects. Future research employing longitudinal designs could also be useful to follow the developmental trajectories of neural sensitivity to emotional vocalisations in typical development in order to assess its potential as a biomarker of atypical neurodevelopment in at-risk children (Elsabbagh & Johnson, 2007).

Chapter 6 The Development of the Neural Processing of Vocal Emotion during the First Year of Life

6.1 Abstract

The second half of the first year of life is a period characterised by rapidly developing social cognition. Sensitivity to emotion expressed in human vocalisations plays a fundamental role in social interactions. Previous research has shown that responses increased in voice-sensitive regions to emotional prosody in 7-month-old infants. However, the neurocognitive developmental trajectory for vocal emotion processing in the first year of life is little documented and poorly understood. In this study, we used functional near infrared spectroscopy (fNIRS) to measure brain sensitivity to angry, happy and neutral vocalisations in the same infants at 6, 9 and 12 months of age. Our findings support the consistent development of neural sensitivity to vocal emotion in superior temporal cortices between 6- to 12-months old. This is the first study to track the neural development of vocal emotional prosody processing longitudinally during the second half of the first year. Results have implications for monitoring typically developing infant social-emotional development. Findings from the present study highlighted the need for the exploration of developmental trajectories of neural sensitivity to vocal expressions beyond the first year of life.

Key words: fNIRS, infant, social-emotional, voice, prosody

6.2 Introduction

Voice recognition is fundamental to human social interaction and has long been investigated as a foundation of social cognition in infants (Blasi et al., 2015; Lloyd-Fox, Blasi, et al., 2013) and children (Chronaki, Benikos, et al., 2015; Chronaki et al., 2018). Neonates can recognise the voice of their parents (Ockleford et al., 1988) and this ability is present antenatally in fetuses as demonstrated by research measuring changes in heart rate at 38 weeks gestation (Kisilevsky et al., 2003). One-month-old infants could identify their own mother's voice with intonation but not their mother's speech without prosodic aspects (Mehler et al., 1978), suggesting infants' preference for prosodic features in speech. The bilateral superior temporal cortices of 7-, but not 4-month-old infants were reported to be more sensitive to human voices compared to non-vocal sounds, suggesting that voice sensitivity emerges sometime between 4 and 7 months of age (Grossmann, Oberecker, et al., 2010). The activation pattern in the temporal cortex in 7-month-old infants appears to be comparable to voice selective neural responses in children (Rogier et al., 2010) and adults (Belin & Zatorre, 2000; Belin et al., 2000).

Social cognition develops with increasing sensitivity to emotional prosody from a similar age range as the emergence of voice sensitivity. Emotional prosody refers to changes in the intonation of the voice according to the speaker's emotional state (Banse & Scherer, 1996; Hargrove, 1997). Vocal emotion processing is crucial for the early development of attachment and social referencing (Mumme et al., 1996; Trevarthen, 2017). The infant's preference for emotional vocalisations has important developmental implications. A positive vocalisation is a feature of baby talk and may signal maternal warmth and safety (Lohaus et al., 2001). In contrast, a negative vocalisation may serve as a signal to avoid danger in ambiguous or threatening

situations (Striano & Rochat, 2000; Vaish & Striano, 2004). Newborns tend to display more eye-opening behaviours to happy voices than to angry, sad, or neutral voices (Mastropieri & Turkewitz, 1999). Five, but not 3-month-old infants were able to detect a prosody change between happy and sad vocalisations, suggesting that this ability may develop between 3 and 5 months (Walker-Andrews & Grolnick, 1983). Five-month-old infants behaviourally respond differently to approval (containing positive emotion) and prohibition (containing negative emotion) in infant-directed speech (Fernald, 1993). However, positive affective speech seems to be favoured in 6--month-old infants and to elicit longer looking times compared to neutral or negative affect (Singh et al., 2002).

Such sensitivity to vocal emotion in infancy may also reflect cortical specialisation of the developing infant brain shaped by early social environmental experiences. Neural sensitivity to vocal emotion emerges from the first month of life (review by Grossmann and Johnson, 2007). Although ERP, fNIRS and fMRI studies indicate emergence of an enhanced sensitivity to both positive and negative vocal emotions (compared to neutral) in early infancy, the developmental timeline of this neural processing remains unclear. One group has reported insular and bilateral frontal activation in asleep infants using fMRI in infants aged between 3-7 months (Blasi et al., 2015; Blasi et al., 2011). An fNIRS study in asleep neonates found greater right temporal activation to emotional (happy, angry and fearful) compared to neutral pseudo-speech vocalisations (Zhang et al., 2017). Another fNIRS study in awake 7-month-old infants reported increased right inferior frontal and superior temporal cortical responses to happy and angry, but not neutral speech respectively (Grossmann, Oberecker, et al., 2010). Whilst an ERP study in 8-month-old infants showed enhancement of P300 component over central electrodes and enhanced N200

component over temporal electrodes in response to crying compared laughing and neutral vocalisations (Missana et al., 2017). Although, these findings suggest that early functional specialisation for processing vocal emotions may develop between 3 and 7 months, none has examined the same infants longitudinally over time using the same stimulus paradigm. Lloyd-Fox et al. piloted longitudinal assessment in infants from 2 to 24 months; their findings provided initial evidence, suggested a stable neural responses in temporal cortex to auditory social, compared to non-social, stimuli in infants from 9-13 months onwards (Lloyd-Fox et al., 2017).

Vocal emotion perception is a fundamental skill in social cognition and a building block for future language development. Clarity about the developmental framework of infant social-emotional neural sensitivity is extremely important; not only may it help us understand better key neural developmental processes, but crucially may offer a biomarker in at-risk infants whose developmental time course may be measurably different from their typically developing or lower risk peers. In the present study, therefore, we aimed to track infants through a period of rapid development for social cognition and social referencing in the second half of the first year (6, 9 and 12 months). Based on previous studies, first, we hypothesised that, at all ages from 6 months, infants would show increased brain responses to emotional (angry, happy) compared to neutral vocalisations in voice-sensitive temporal regions. Second, we predicted that developmental trajectories of prosody processing would differ as a function of emotion type. Thirdly, we anticipated that infant neural responses to vocal emotional stimuli would increase with age.

6.2 Methods

Participants

Forty infants of healthy mothers were recruited to the present study from three Manchester community health centres. All infants were born at normal birth weight (>2500 g); 39 were full term (37–42 weeks gestation); one was born at 36 weeks gestation (corrected gestational age). Infants were typically developing and had no hearing difficulties according to parental report. Mothers consented on behalf of their infants. The UK National Health Service ethics committee approved the study (REF: 15/NW/0684).

A total of 40 infants were tested at 6 months (20 girls and 20 boys, aged between 175 and 214 days, M = 189.48 days, SD= 9.27). From those 40 infants, a total of 39 were tested again at 9 months after 1 infant dropped out (19 girls and 20 boys, between 263 and 302 days, M = 279.08 days, SD = 9.46). An additional infant dropped out at 12 months, therefore a total of 38 infants were tested at 12 months (18 girls and 20 boys, between 360 and 394 days, M = 377.24 days, SD= 8.61). The flow chart in Figure 6.1 describes the longitudinal participation.

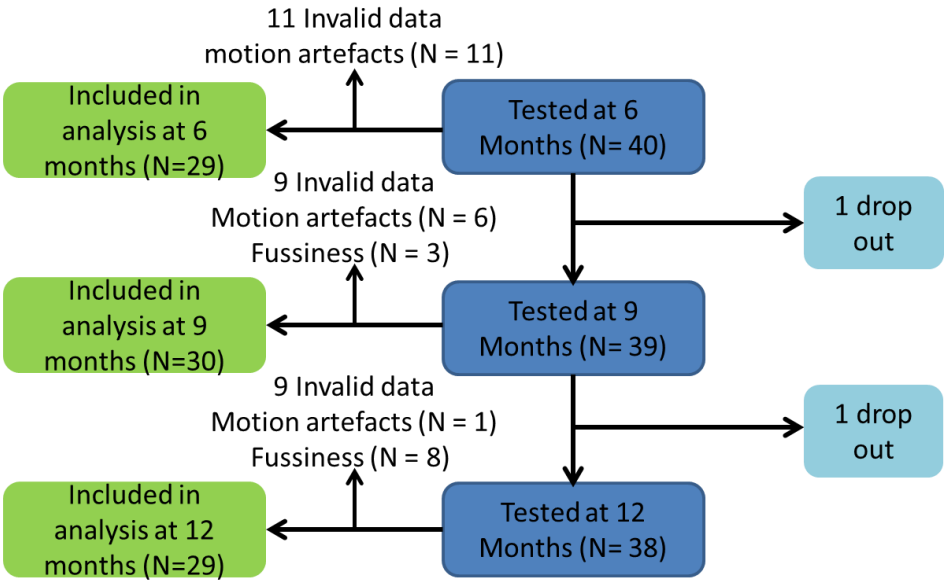


Figure 6.1 Total number of infants tested and included in the analysis at 6, 9, and 12 months.

Experimental paradigm and procedure

Figure 6.2 presents the experimental paradigm used at all three time points. Infants wore the NIRS headband, sat in their mothers' laps in front of a laptop during the task and listened to the vocal stimuli. The task started with a 20 sec rest period, followed by a 5 sec trial presented through loudspeakers (SPL = 70dB). Consistent with previous research (Grossmann, Oberecker, et al., 2010), a 5 sec silent cartoon was played along with each trial to attract infants' attention and reduce motion artefacts. After each trial, a 10 sec silent baseline along with the blurred cartoon was presented. The task was presented with PsychoPy software (Peirce, 2007). The same emotional expression did not occur consecutively. Each condition (angry, happy and neutral) was presented 8 times amounting to a total number of 24 trials. The total length of the testing session was 6 minutes and 20 seconds.

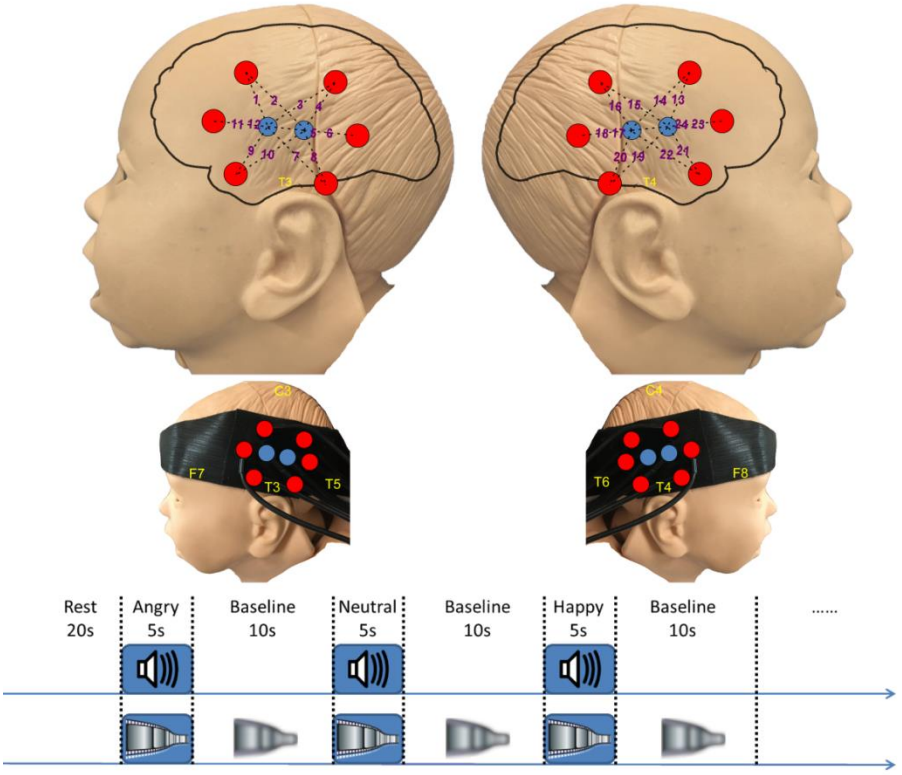


Figure 6.2. Study experimental task design and channel distribution. The head model illustrates the source-detector distribution where red dots represent sources (6 in each hemisphere) and blue dots represent detectors (2 in each hemisphere). Sources and detectors form 12 recording channels in each hemisphere, which are marked in purple numbers (upper head models), and are held by Velcro head band. Scalp landmarks with respect to 10-20 system are marked in yellow (middle head models). The bottom streamline demonstrates the timeline of the experimental task stimulus presentation and baseline.

Vocal emotional stimuli

The stimulus material consisted of 15 female non-linguistic vocalisations of angry, happy and neutral prosody (interjection ‘ah’) from a well-validated battery of vocal emotional expressions (Maurage et al., 2007). This battery has high internal

consistency for each emotion set and high levels of specificity (independence between the ratings in the different emotion sets (Maurage et al., 2007). These stimuli have been validated in previous research in children of different ages (Chronaki, Benikos, et al., 2015). Five normalised stimuli, each lasting 1 sec, from the same condition were selected and then combined to form a 5 sec trial. All vocal stimuli were normalised with Praat sound-analysis software (Boersma & van Heuven, 2001) to the same duration of 1 000 ms and mean intensity of 73dB (see Supplementary Information for details on stimuli acoustic properties). Vocal emotional stimuli were the same for all the three time points' data collection.

Data acquisition

During functional cerebral activation, the fNIRS system measures the changes in attenuation of near infrared light. These changes in attenuation are caused by changes in blood volume and the ratio of oxygenated and deoxygenated blood caused by the haemodynamic response (Villringer & Chance, 1997). In the present study, infants' cerebral responses were recorded with a multichannel NIRS data collection system. The system was built by Biomedical Optics Research Laboratory, (Dept. of Medical Physics and Bioengineering, University College London) and applied with 780nm and 850nm continuous wavelengths and 10 Hz sampling rate (Everdell et al., 2005). Two detectors and six sources formed 12 source-detector pairs in each hemisphere and were distributed at temporal regions which have been shown to be voice sensitive in previous research in infants (Grossmann, Oberecker, et al., 2010; Lloyd-Fox et al., 2012; Pena et al., 2003; Taga & Asakawa, 2007) and adults (Belin et al., 2000; Ethofer, Anders, Wiethoff, et al., 2006; Grandjean et al., 2005). To achieve the best spatial sensitivity profile for infants (Fukui et al., 2003), the distances between source and detectors were fixed between 1.5 and 2.5 cm.

Channels were distributed according to the 10-20 system and attached to a custom-made Velcro headband. According to the infants' head growth standards from the World Health Organisation (World Health Organization (WHO), 2003), and from one previous longitudinal pilot study (Lloyd-Fox et al., 2017), the head circumference of 6 to 12 months old infants does not change significantly. Therefore, the application of a fixed source-detector array across three age time points is reasonable and practical (Lloyd-Fox et al., 2017). The headband was adjusted by calculating the distance between the glabella and the ear, ensuring that T3 and T4 are between the two bottom sources in each hemisphere. This procedure was carried out for all the infants at each time point. The locations of the channels and the channel positions with respect to the 10-20 system are presented in Figure 6.2.

Data analysis

Infant behaviour was video-recorded during the experimental task. A member of the research team viewed the videos and coded the number of trials in which an infant attended to the screen without large motion artefacts. Four out of eight trials per condition were set as a criterion for inclusion of each infant dataset. As a result, 29 datasets were included in the 6 months' time point, 30 datasets were included in the 9 months' time point, and 29 datasets were included in the 12 months' time point. This inclusion rate is consistent with criteria for inclusions in previous infant NIRS studies (Lloyd-Fox et al., 2010).

All the datasets included in analysis were filtered at 0.01 to 0.5Hz with a 3rd order Butterworth filter, to eliminate slow drifts, instrument noise and physiological artefacts such as heartbeats (Cooper et al., 2012; Fox et al., 2013; Grossmann, Oberecker, et al., 2010). The datasets were then converted to optical density data in

HOMER2 NIRS toolbox (version 2.1, http://homer-fnirs.org/, Huppert et al., 2009). The remaining artefacts were identified on a channel by channel basis with the algorithm 'hmrMotionArtifactByChannel' implemented in the HOMER2. Within the time interval (tMotion), if the change of the signal amplitude exceeded the threshold (AMPthresh) or the standard deviation changes were greater than a factor (STDEVthresh) multiplied by the original channel standard deviation, the time period (tMask time before and after the motion artefact) was marked as artefacts. The time period of motion artefact within the channel was corrected with a cubic spline interpolation algorithm with p set to 0.99 as recommended (Cooper et al., 2012; Scholkmann et al., 2010). Since the algorithm works on a channel by channel basis, the actual standard deviation threshold for the motion artefact varies according to the standard deviation of the original channel; the setting of the STDEVthresh is the multiplication factor rather than a fixed threshold (i.e. in the current study the standard deviation threshold is 20*standard deviation of the channel). This means that the standard deviation threshold varies from channel to channel and subject to subject. All the values were set as follows: tMotion=5s; tMask=1s; STDEVthresh=20; AMPthresh=5.

After pre-processing, data were converted to Oxy- and Deoxy-Haemoglobin concentration changes in HOMER2 and averaged across trials in the same emotional condition within each dataset, with the time window of 1 sec before and 15s after the stimulation onset. The averaged time course of each channel was corrected by subtracting the mean of the 1 sec before the stimulation. The analysis focused on Oxy-Haemoglobin concentration changes as the most sensitive indicator of changes in cerebral blood flow. Kolmogorov-Smirnov tests were conducted to test the normality of the haemodynamic time courses, results showed that p values were

larger than 0.05, which indicate that haemodynamic data were normally distributed. Based on earlier work showing that the haemodynamic response reaches the peak around 2 to 4 sec post stimulus (Brigadoi et al., 2014), we targeted a time window of 2 sec to 9 sec after stimulus onset. Mean amplitudes of cortical haemodynamic responses (Oxy- and Deoxy-Haemoglobin waveforms) were averaged over the time window of 2 sec to 9 sec after stimulus onset.

The averaged haemodynamic responses to the expression conditions (angry, happy and neutral) were evaluated with repeated measures Analysis of Variance (ANOVA) and post-hoc pairwise comparisons at each age point (6, 9, and 12 months). We calculated partial eta-squared (Cohen, 1973; Kennedy, 1970) to estimate the effect sizes for the main effect of emotion as well as for contrasts. Partial eta-squared takes values between 0 and 1. Values of 0.02, 0.13, and 0.26 are indicative of a small, medium and large effect size, respectively (Murphy et al., 2014). Consistent with other infant imaging studies (Blasi et al., 2015; Lloyd-Fox et al., 2017), a false discovery rate (FDR, Benjamini and Hochberg, 1995) correction was applied to resolve the issue of multiple statistical comparisons. For channels that showed differential response in Oxy-Haemoglobin concentration changes between emotional conditions, post-hoc comparison p values arranged in ascending order with an order number index allowed us to calculate adjusted α values: $\alpha adjust$ = (order index / total number of comparisons)*0.05. Pairwise p value < adjusted α value remained significant.

To examine age effects and age x emotion interaction effects on the Oxy-Haemoglobin concentration changes to the vocal emotional expressions, we conducted repeated measures ANOVA with emotion (angry, happy, neutral) and age (6 months, 9 months, 12 months) as within subject factors and the Oxy-

Haemoglobin concentration changes as the dependent measure. This analysis included the 21 infants, who had provided valid data at all time points (6, 9 and 12 months). This analysis included those channels identified as showing sensitivity to emotion in the previous analysis (see Emotion effects on neural response to voices section). In addition, Pearson's bivariate correlations examined whether Oxy-Haemoglobin concentration changes for each emotion were associated with age in the longitudinal sample.

6.3 Results

Of the tested 40 six-month-old infants, 29 were included in analysis (15 girls and 14 boys, between 175 and 214 days, M = 189.31 days, SD=9.66). Eleven 6-month-old infants were excluded from analyses as they failed to meet the minimum 4 out of 8 trials per experimental condition (see inclusion criterion in the data analysis section) as a result of motion artefact (N = 11). From the tested 39 infants at 9 months, 30 infants were included in analyses (11 girls and 19 boys, between 268 and 302 days, M = 279.67 days, SD = 8.81). Nine infants were excluded from analyses as they failed to meet the criterion, as a result of motion artefacts (n = 6) or fussiness (n = 3) during data collection. Out of the tested 38 infants at 12 months, 29 infants were included in analyses (14 girls and 15 boys, between 360 and 394 days, M = 375.59 days, SD = 7.84). Nine infants failed to meet the inclusion criterion due to motion artefacts (N = 1) or fussiness (N = 8) during data collection. Derivation of the final full analysis sample is described in Figure 6.1.

The attrition rate for all the three time points was within the standard range for infant NIRS studies (40% on average is an accepted rejection rate from previous studies, see Grossmann et al., 2010; review by Lloyd-Fox et al., 2010). The sample size for each time point was determined by a power analysis using G*power (Faul et al., 2007). This indicated that a sample size of 29 subjects would give 92% power to achieve an effect size of 0.59. Our sample size is consistent with previous fNIRS studies in infants of similar ages (Grossmann, Oberecker, et al., 2010; Zhang et al., 2017).

Emotion effects on neural response to voices

6-month-olds

Our analysis revealed three channels that were sensitive to emotional prosody in Oxy-Haemoglobin concentration (see Figures 6.3 and Table 6.1). In particular, channel 2 in the left hemisphere ($F(2, 56) = 3.38$, $p = .040$, $\eta_p^2 = .11$), channel 14 ($F(2, 56) = 3.24$, $p = .047$, $\eta_p^2 = .10$) and channel 16 ($F(2, 56) = 4.38$, $p = .017$, $\eta_p^2 = .14$) (both in the right hemisphere) showed significant differences in Oxy-Haemoglobin concentration changes when emotion (angry, happy and neutral) was entered as a within-subject factor in repeated measures ANOVAs. Pairwise comparisons showed that Oxy-Haemoglobin concentration changes in channel 2 were significantly larger to angry compared to neutral voices ($F(1, 28) = 9.76$, $p = .004$, $\eta_p^2 = .26$). This effect survived FDR correction (see Figure 6.3 and Table 6.1). In channel 16, Oxy-Haemoglobin concentration changes were significantly increased to happy compared to angry voices ($F(1, 28) = 8.26$, $p = .008$, $\eta_p^2 = .23$). This effect survived FDR correction (see Figure 6.3 and Table 6.1). In channel 14, Oxy-Haemoglobin concentration changes were larger to happy compared to neutral voices ($F(1, 28) = 5.62$, $p = .025$, $\eta_p^2 = .17$) and marginally larger to happy compared to angry voices ($F(1, 28) = 4.26$, $p = .048$, $\eta_p^2 = .13$). This effect did not survive FDR correction (see Table 6.1). In summary, bilateral temporal regions were activated to vocal emotion stimuli.

The analysis of Deoxy-Haemoglobin concentration changes complemented the Oxy-Haemoglobin changes and survived FDR correction. Specifically, there was a significant effect of emotion on Deoxy-Haemoglobin concentration changes in channel 2 in the left hemisphere ($F(2, 56) = 4.04$, $p = .020$, $\eta_p^2 = .13$), with decreased

Deoxy-Haemoglobin concentration changes in response to angry voice compared to neutral voice (F (1, 28) = 10.26, p = .003, η_p^2 = .27). This effect survived FDR correction. Furthermore, there was a significant effect of emotion in channel 16 in the right hemisphere (F (2, 56) = 3.62, p = .030, η_p^2 = .11), with decreased Deoxy-Haemoglobin concentration changes in response to happy voice compared to angry voice (F (1, 28) = 7.45, p = .010, η_p^2 = .21). This effect also survived FDR correction.

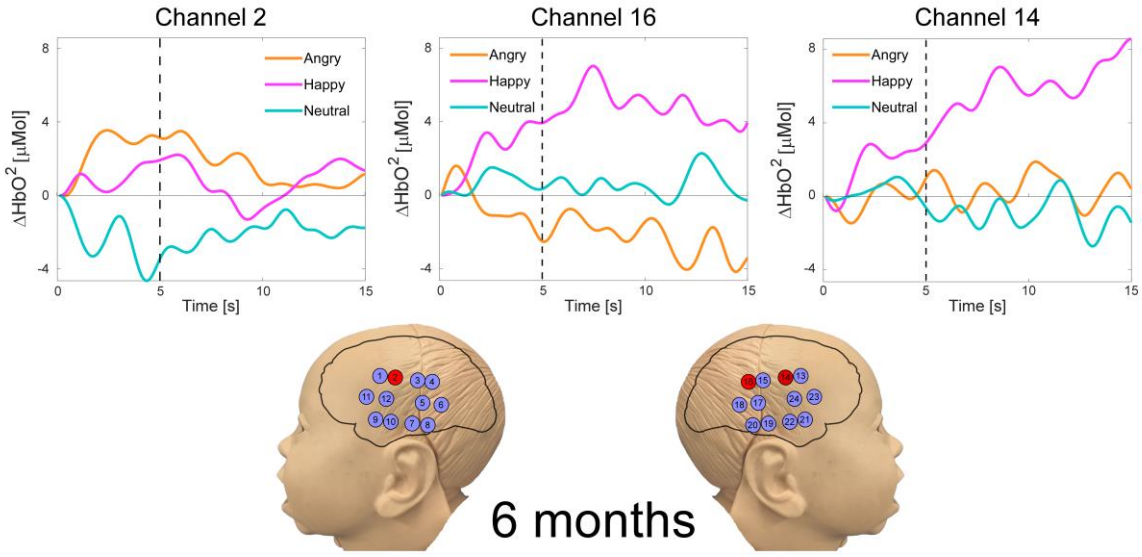


Figure 6.3 Averaged time courses of Oxy-Haemoglobin concentration changes across all datasets in 6-month-olds in channel 2 (in the left hemisphere), channel 14 and channel 16 (in the right hemisphere) per vocal emotion (Angry in orange, Happy in pink and Neutral in light blue) in the time period of 15 sec (5 sec stimulus and 10 sec baseline). The channel location is marked in the lower panel. The stimulus end time is marked by the dashed line. The time (in sec) and change in amplitude (μMol) are in the x and y axis respectively.

9-month-olds

The analysis for 9 months Oxy-Haemoglobin concentration data showed that channel 21 in the right hemisphere was significantly modulated by emotional prosody ($F(2, 58) = 3.45$, $p = .038$, $\eta_p^2 = .11$). Oxy-Haemoglobin concentration changes were significantly greater for the happy compared to angry voices in channel 21. This effect held after FDR correction ($F(1, 29) = 9.59$, $p = .004$, $\eta_p^2 = .25$; See Figure 6.4 and Table 6.1).

Two channels showed a significant effect of emotion in Deoxy-Haemoglobin concentration changes: channel 9 in the left hemisphere ($F(2, 58) = 4.23$, $p = .019$, $\eta_p^2 = .13$), and channel 16 in the right hemisphere ($F(2, 58) = 3.45$, $p = .038$, $\eta_p^2 = .11$). However, none of the comparisons survived FDR correction.

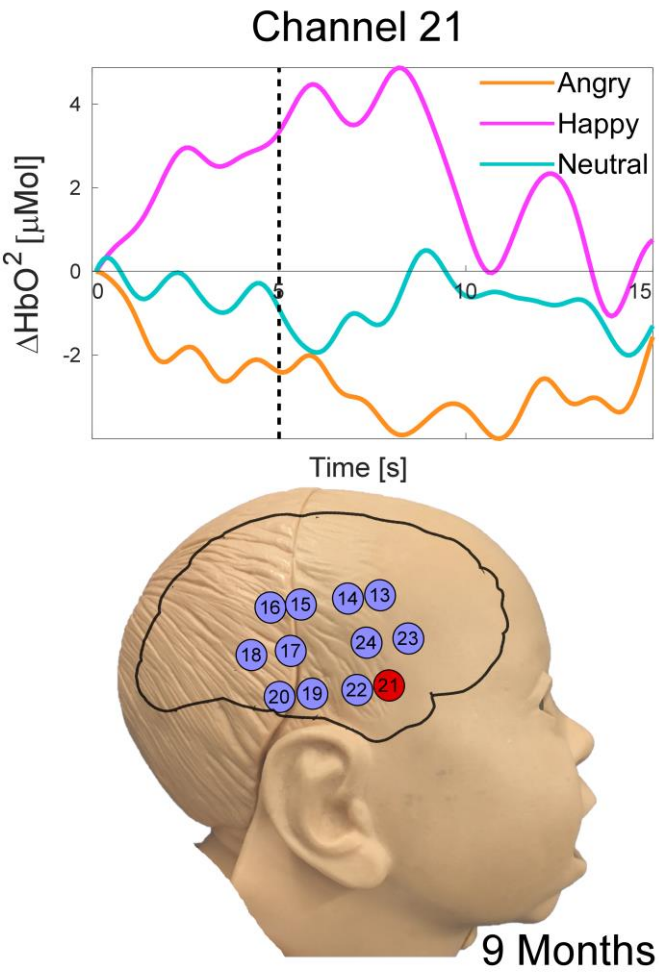


Figure 6.4 Averaged time courses of Oxy-Haemoglobin concentration changes across all datasets in 9-month-olds in channel 21 (in the right hemisphere) per vocal emotion (Angry in orange, Happy in pink and Neutral in light blue) in the time period of 15 sec (5 sec stimulus and 10 sec baseline). The channel location is marked in the lower panel. The stimulus end time is marked by the dashed line. The time (in sec) and change in amplitude (μMol) are in the x and y axis respectively.

12-month-olds

At 12 months, two channels showed significant differences in Oxy- Haemoglobin concentration changes to the emotional prosody: channel 9 in the left hemisphere ($F(2, 56) = 4.17$, $p = .020$, $\eta_p^2 = .13$); and channel 21 in the right hemisphere ($F(2, 56)$

= 3.24, p = .047, η_p^2 = .10). Specifically, pairwise comparisons revealed that happy voices had significantly greater Oxy-Haemoglobin concentration changes when compared to angry voices in channel 9 (F (1, 28) = 10.53, p = .003, η_p^2 = .27). This effect survived FDR correction (see Figure 6.5 and Table 6.1). Neutral voices evoked stronger Oxy-Haemoglobin concentration changes in contrast to angry voices in channel 21 (F (1, 28) = 6.53, p = .016, η_p^2 = .19; See Table 6.1). This effect did not survive FDR correction.

Channel 10 presented a significant effect of emotion in Deoxy-Haemoglobin concentration changes (F (2, 56) = 3.78, p = .030, η_p^2 = .12), with significantly reduced Deoxy-Haemoglobin concentration changes evoked by angry compared to neutral voices (F (1, 28) = 8.41, p = .007, η_p^2 = .23). This effect survived FDR correction.

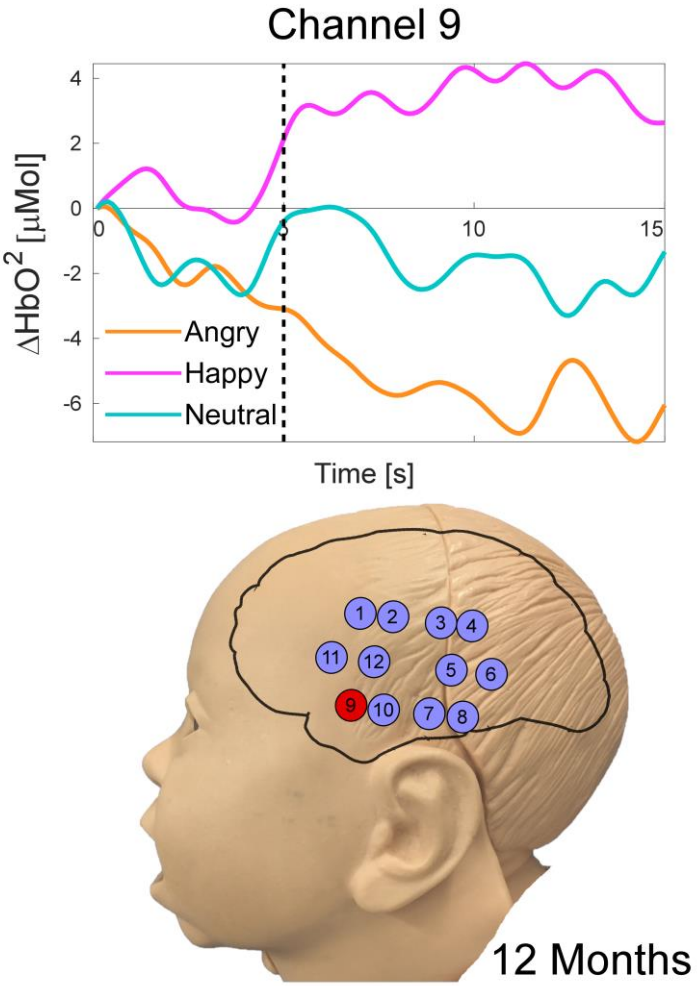


Figure 6.5 Averaged time courses of Oxy-Haemoglobin concentration changes across all datasets in 12-month-olds channel 9 (in the left hemisphere) per vocal emotion (Angry in orange, Happy in pink and Neutral in light blue) in the time period of 15 sec (5 sec stimulus and 10 sec baseline). The channel location is marked in the lower panel. The stimulus end time is marked by the dashed line. The time (in sec) and change in amplitude (μMol) are in the x and y axis respectively.

Table 6.1. Summary of emotion (Angry, Happy, Neutral) ANOVA effects on Oxy-Haemoglobin concentration changes at 6, 9 and 12 months.

Channel	Emotion	Mean ± SEM	ANOVA			Pairwise Comparisons				Adjusted α value
			F	p	Partial Eta-squared	Comparison[a]	F	p	Partial Eta-squared	α_{adjust}
6 Months										
2	Angry	2.82±1.6	3.38	0.040	0.11	A > H	0.56	0.462	0.02	0.044
	Happy	0.97±1.9				A > N	9.76	**0.004***	0.26	0.006
	Neutral	-2.68±1.5				H > N	2.86	0.102	0.10	0.033
14	Angry	0.29±1.34	3.24	0.047	0.10	H > A	4.26	0.048	0.13	0.022
	Happy	4.02±1.67				A > N	0.11	0.746	0.004	0.050
	Neutral	-0.33±1.24				H > N	5.62	0.025	0.17	0.017
16	Angry	-1.51±1.74	4.38	0.017	0.14	H > A	8.26	**0.008***	0.23	0.011
	Happy	4.49±1.58				N > A	1.10	0.300	0.04	0.039
	Neutral	0.73±1.25				H > N	3.80	0.060	0.12	0.028
9 Months										
21	Angry	-2.67±1.68	3.45	0.038	0.11	H > A	9.59	**0.004***	0.25	0.017
	Happy	3.57±1.75				N > A	0.55	0.465	0.02	0.033
	Neutral	-0.84±1.53				H > N	2.52	0.123	0.08	0.050
12 Months										
9	Angry	-3.79±1.40	4.17	0.021	0.13	H > A	10.53	**0.003***	0.27	0.008
	Happy	1.88±1.37				N > A	1.26	0.271	0.04	0.042
	Neutral	-1.32±1.72				H > N	2.74	0.109	0.09	0.033
21	Angry	-2.62±1.79	3.24	0.047	0.10	H > A	0.78	0.385	0.03	0.050
	Happy	-0.43±1.65				N > A	6.53	0.016	0.19	0.017
	Neutral	3.16±1.55				N > H	2.87	0.101	0.09	0.025

*Comparison survived FDR correction. [a]A = Angry, H = Happy, N = Neutral

Age effects on neural response to emotional voices

Results showed a significant main effect of age on the Oxy-Haemoglobin concentration changes in channel 16 ($F (2, 40) = 3.40$, $p = 0.040$, $\eta_p^2 = .29$). In particular, pairwise comparisons showed significantly increased Oxy-Haemoglobin concentration changes at 12 months compared to 9 months ($F (1, 20) = 8.02$, $p = 0.010$, $\eta_p^2 = .29$). A significant main effect of emotion on the Oxy-Haemoglobin concentration changes was found in channel 9 ($F (2, 40) = 4.39$, $p = 0.019$, $\eta_p^2 = .18$). In particular, pairwise comparisons showed significantly increased Oxy-Haemoglobin concentration changes on hearing happy compared to angry voices ($F (1, 20) = 7.83$, $p = 0.011$, $\eta_p^2 = .28$). In addition, results showed a significant age x emotion interaction effect on the Oxy-Haemoglobin concentration changes in channel 16 ($F (4, 80) = 0.024$, $p = 0.024$, $\eta_p^2 = .13$). Specifically, the contrast in Oxy-Haemoglobin concentration changes between 6 and 9 months for happy compared to neutral prosody was significant ($F (1, 20) = 14.80$, $p = 0.001$, $\eta_p^2 = .43$). In particular, Oxy-Haemoglobin concentration changes were larger for happy compared to angry prosody at 6 months and larger for angry compared to happy prosody at 9 months. In addition, the contrast in Oxy-Haemoglobin concentration changes between 6 and 12 months for happy compared to angry prosody was significant ($F (1, 20) = 8.32$, $p = 0.009$, $\eta_p^2 = .29$). In particular, Oxy-Haemoglobin concentration changes were larger for happy compared to angry prosody at 6 months and larger for angry compared to happy prosody at 12 months. The age x emotion interaction effect on Oxy-Haemoglobin concentration changes is shown in Figure 6.6. None of the main effects (age, emotion, or age x emotion interaction) survived FDR correction.

Pearson's bivariate correlation coefficient (r) confirmed the positive correlation between the age and Oxy-Haemoglobin concentration changes for angry prosody in channel 16 ($r = 0.27$, $p = 0.03$). This confirms that Oxy-Haemoglobin concentration changes for angry prosody become larger with increasing age in the first year of life.

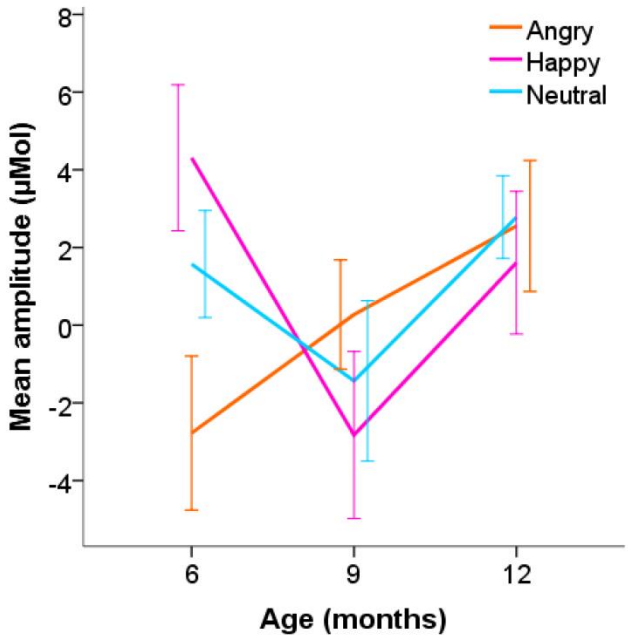


Figure 6.6 Mean amplitudes of Oxy- Haemoglobin concentration changes for each emotion and age in channel 16.

6.5 Discussion

This study tracked infant haemodynamic responses in brain to human emotional vocalisations. We focused on the developmental trajectory of temporal cortical activation to human emotional (angry, happy) non-speech vocalisations, across three time points in 6 to 12 month old infants. There are three main findings. First, we report significant activation in superior temporal cortex in response to vocal emotional stimuli in 6 to 12 month infants. Second, we demonstrate, for the first time, the consistency of activations in infant temporal regions, particularly in response to happy voices. Finally, we found an age effect in the infant neural processing of vocal emotion stimuli, right temporal neural responses to angry voices increased in the second half of the first year of life.

Temporal cortical activation to vocal emotions was consistently elicited in infants aged 6, 9 and 12 months, suggesting that this brain region may underlie prosody processing from the age of 6 months. The pattern of neural response to emotional vocalisations in our study is comparable to that identified previously as voice-sensitive brain regions (Grossmann, Oberecker, et al., 2010; Lloyd-Fox et al., 2012); and in regions within the temporo-frontal network. The temporo-frontal network has been implicated in the neural processing of vocal emotions in adults (Ethofer et al., 2012; Frühholz et al., 2012; Frühholz & Grandjean, 2013; Witteman et al., 2012), as well as in 3-8 months-old infants (Blasi et al., 2015; Grossmann, Oberecker, et al., 2010; Missana et al., 2017).

However, the activation locations within the superior temporal cortices in response to vocal emotion stimuli differed between 6 to 12 months. Specifically, whereas

superior temporal activation was bilateral at 6 months, it was unilaterally localised inferiorly to the right by 9 months and to the left by 12 months. The changing brain activation regions across infant ages may suggest that neural responses to vocal emotion is not yet specialised across this time period and is continuously being 'fine-tuned' across the first year of life (Johnson et al., 2009; Kolb & Gibb, 2011; Leppänen & Nelson, 2009). This is supported by our finding of the age effect, which showed right temporal activation to angry prosody increased with age between 6 and 12 months. Furthermore, although it did not survive multiple comparison corrections, right temporal cortex also showed an effect of age with increased cortical activation to emotional vocalisations; especially greater neural responses in 12-month-old infants compared to 9-month-old infants. The age effect also extends previous findings of neural responses to voice and non-vocal sounds become more distinct with increasing age in 3-to 7-month-old infants (Blasi et al., 2011; Lloyd-Fox et al., 2012). Additional evidence of age effects is apparent within Lloyd-Fox and colleagues' longitudinal study (Lloyd-Fox et al., 2017). They recruited a cohort of 27 0-2-month-old infants and a second cohort of 42 4-8 month-old infants. This second cohort included longitudinal data collection at 9 months and 12 months. Infants listened to human non-speech vocalisations (e.g., cough, yawn) and environmental sounds (e.g., running water, bells). The study reported that, in the first two months of life infants exhibited selectivity to non-vocal sounds, which persisted until 4–8 months when a transition to greater social stimulus selectivity was observed with more stable neural responses to auditory social stimuli by 9 months (Lloyd-Fox et al., 2017). Considering our findings, alongside existing evidence, we would suggest that the neural systems supporting vocal emotion processing in infants

emerges at 6 months of age and continues to be shaped and to develop between 6 to 12 months.

Our study found enhanced neural responses to angry compared to neutral voices at 6 months consistent with previous research (Grossmann, Oberecker, et al., 2010), possibly suggesting an enhanced attention to, and processing of threatening stimuli early in the infant life. Interestingly, this pattern was not observed at 9 or 12 months. This may suggest prioritising of threatening signals by superior temporal cortices at 6 months but that subsequently, priority is given to sensitivity for happy vocalisations, which persists through the first year. However, we observed a significant trend of increasing neural responses with age, suggesting that the infant neural responses to vocal anger are continuously developing over the second half of the first year. This developmental trend is consistent with infant development of social referencing by the end of the first year, by which time the infant shows less behavioural exploration when hearing negative vocalisations in ambiguous situations (See Walker-Andrews, 1997; Mumme et al., 1996). These findings support the notion of a continuously developing neural system for processing vocal emotion in infancy and confirm our prediction that the developmental time course of prosody processing may be different for different types of emotion (i.e. positive and negative).

The developmental pattern observed in the present study suggests consistent neural sensitivity differences between happy and angry vocalisations across three time points. This finding may suggest that there is a temporal sensitivity for processing positive affect in conjunction with an aversion to angry affect. Our findings are supported by recent research showing enhanced attention allocation to happy

vocalisations, as reflected by increased P3 component amplitudes to happy vocalisations compared to neutral and angry vocalisations in adults (Pinheiro, Barros, Vasconcelos, Obermeier, & Kotz, 2017). Enhanced brain responses to happy prosody are consistent with the emotional tones that characterise infant-directed speech. Infant-directed speech is typically used by parents to soothe their infants and infants typically attend to infant-directed speech more than to usual adult speech (Saint-Georges et al., 2013; Singh et al., 2002). This preference for positive affect may assist in facilitating language acquisition (Kuhl, 2004, 2007). Although infants are also exposed to expressions of anger, these expressions are not usually as frequent as expressions of happiness, at least in non-adverse environments (Malatesta, Grigoryev, Lamb, Albin, & Culver, 1986). For instance, positive social interactions with caregivers in healthy mother-infant dyads are typically characterised by smiling faces and happy voices. During interactions with infants, caregivers not only express positive affect (e.g. joy, interest) more frequently than negative (e.g. anger), but also they respond more contingently to infants' positive than negative emotional signals (Eisenberg, Cumberland, & Spinrad, 1998; Malatesta et al., 1986). It is possible that the enhancement of sensory processing of positive signals is a fundamental and early developing neural mechanism, engaged to prioritise the processing of motivationally significant stimuli. It is also possible that while the processing of happy voices is supported by activation of superior temporal regions, other sub-cortical structures (e.g., amygdala) will underlie the processing of negative stimuli (anger, fear).

Importantly, our vocal stimuli were carefully selected and standardised for acoustic parameters. This careful stimulus selection has allowed us to isolate emotion-specific

effects. In our study, the emotion effects which survived correction included larger brain activity to angry compared to neutral voices at 6 months, and happy compared to angry voices at 6, 9 and 12 months. We are confident these effects reflect emotion-specific effects in infants because analyses of acoustic properties of the vocal stimuli showed no significant differences in low-level acoustic features (e.g. fundamental frequency) between angry and neutral voices or between angry and happy voices.

The present study does, however, have some limitations. The lack of a main effect of age on vocal emotion processing that withstood multiple comparison corrections, suggests that our findings require further replication in larger samples of infants. We did not examine brain responses to vocal compared to non-vocal signals which would have allowed us to establish the selectivity of neural responses to vocal signals. Vocal stimuli included in the current study consisted of one type of negative (i.e. angry) and positive (i.e. happy) valence. Future studies should include other types of positive and negative vocalisations to explore developmental trajectories of vocal emotion processing. In addition, all vocal stimuli in this study were from women who were unfamiliar to the infants. Infants are sensitive to their mother's voices (Dehaene-Lambertz et al., 2010; Walker-Andrews et al., 2011). Future research should expand the stimulation paradigm to include emotional voices from own mothers that are meaningful to infants; and which may evoke stronger responses than unfamiliar women's voices.

The present study provides novel evidence for the neural development of vocal emotion processing from 6 to 12 months. Our findings support a pattern of consistent development of superior temporal cortical sensitivity to vocal emotion prosody

between 6 and 12 months of age in typically developing infants. Results have implications for monitoring typically developing infant social-emotional development. Findings from the present study highlight the need to explore the developmental trajectories of neural sensitivity to vocal expressions beyond the first year of life. In addition, this study opens questions up about whether and how early experience or environmental factors influence vocal emotion development.

6.6 Supplementary Information

Table S6.1: Stimuli acoustic properties in terms of duration, intensity and fundamental frequency (pitch)

Emotion	Duration (s)	Mean Intensity (dB)	Max Intensity (dB)	Min Intensity (dB)	Mean Pitch (Hz)	Max Pitch (Hz)	Min Pitch (Hz)
Angry 1	1.00	69.99	75.83	57.59	345.88	492.56	271.00
Angry 2	1.00	73.18	78.39	62.88	142.17	301.61	76.26
Angry 3	1.00	70.67	80.67	52.67	151.38	248.96	97.27
Angry 4	1.00	75.83	80.87	67.67	142.89	346.95	78.54
Angry 5	1.00	73.87	80.24	52.69	210.92	317.19	116.00
Happy 1	1.00	69.99	80.48	61.85	282.41	393.98	180.13
Happy 2	1.00	78.55	81.29	72.01	289.16	312.25	259.06
Happy 3	1.00	73.73	83.60	67.24	347.18	524.56	219.98
Happy 4	1.00	75.31	78.08	70.42	352.31	485.99	224.65
Happy 5	1.00	74.63	79.57	65.49	233.33	327.25	139.94
Neutral 1	1.00	74.04	79.14	58.41	207.20	218.65	199.01
Neutral 2	1.00	71.65	79.37	53.72	190.01	197.30	178.68
Neutral 3	1.00	76.18	78.34	68.43	191.14	194.65	181.69
Neutral 4	1.00	75.17	77.91	59.45	192.91	206.03	176.71
Neutral 5	1.00	75.01	80.27	57.27	215.23	239.25	191.02

The vocal stimuli were evaluated as for the following acoustic parameters: i) mean duration which was identical for all stimuli (5 sec); ii) mean intensity in decibels (angry = 72.70, SD = 2.40; happy = 74.44, SD = 3.08; neutral = 74.40, SD = 1.70); and iii) mean fundamental frequency in Hertz (angry = 198.65, SD = 87.10; happy = 300.87, SD = 49.57; neutral = 199.30, SD = 11.30). These acoustic parameters were used to compare acoustic differences across the three emotions. The means were compared using t-tests. Results showed that stimuli did not significantly differ in their mean intensity ($t < -1.58$, $p > .19$). In addition, there was no significant difference in mean fundamental frequency between angry and neutral stimuli (t =

-.018, $p = .98$) and between angry and happy stimuli ($t = -2.28$, $p > .05$). In line with previous research (Grossmann et al., 2005), fundamental frequency was significantly higher for happy than for neutral stimuli ($t = 4.47$, $p < .05$).

Chapter 7 Longitudinal infant voice and non-vocal sounds processing

7.1 Abstract:

Infant neural activations in response to hearing human vocalisations are thought to be precursors of language and socio-communicative development. However, the human voice stimuli used in previous studies tend to have included emotion or speech within the stimuli, and these may have a specific influence on infant neural activations which were not distinguished. The present study aimed to observe the development of infant neural processing of human voice without the influence of emotion and speech by examining the brain activations to emotionally neutral non-speech vocalisations in the same infants longitudinally on three occasions at 6, 9 and 12 months of age. We found non-vocal stimuli evoked significantly greater responses compared to voice stimuli in temporal cortices at 6 and 9 months of age, but not in 12 months. We did not find greater neural responses to voice than to non-vocal sounds at any time point. Our findings may suggest continuously developing neural responses to neural non-speech voice in infants in the first year of life. The infant's neural responses to human vocalisations may depend on the medium through which voice is carried (e.g. speech or emotional vocalisations). Infants' enhanced neural responses to non-vocal stimuli may represent their heightened attention to novel sounds. Other factors may also influence infant auditory brain responses to voice versus non-vocal stimuli, including the novelty of the stimuli and the infant's familiarity with the sound. Further studies are needed to replicate the present paradigm beyond the first year of life, and to provide comprehensive interpretation to infant neural responses.

Keywords: infant, NIRS, voice, non-vocal, temporal cortex

7.2 Introduction

Language is a critical cognitive skill in human survival which begins developing in infancy. Delay in language development has been found to predict psychiatric and behavioural problems, such as internalisation and externalisation behaviours (Beitchman et al., 1996; Beitchman et al., 1986; Maggio et al., 2014; Toppelberg & Shapiro, 2000). The preference for familiar human vocalisations starts before birth: fetuses show more interest in their own mothers' voices than strange female voices (Kisilevsky et al., 2003). Interest in the maternal voice (and more generally in human voices) becomes evident by 3 months of age in infancy (Blasi et al., 2015; Blasi et al., 2011; Dehaene-Lambertz et al., 2010; Grossmann, Oberecker, et al., 2010; Lloyd-Fox et al., 2012). These early infant social-communication indicators are widely studied with contemporary psychophysiological tools (i.e. EEG, fMRI, fNIRS). The neural network development in response to human voices is recognised as a precursor of early language, as well as social-communication development. Studies using fNIRS and fMRI have reported that brain responses to auditory social stimuli (cough, yawn, laugh and cry) can distinguish between infants with and without risk of neurodevelopmental disorder (Blasi et al., 2015; Lloyd-Fox, Blasi, et al., 2013). In addition, studies also reported the infant ability to process non-linguistic auditory stimuli from the age of 6 months, which predicted language competence at 20, 36, 48 months of age (Cantiani et al., 2016; Choudhury & Benasich, 2011). These studies suggest that the process of human voice stimuli and non-linguistic auditory stimuli are associated with infant neural responses and later language development.

Lloyd-Fox and colleagues provided evidence of neural responses to auditory social stimuli from as early as 2 months up to 24 months of age (Lloyd-Fox et al., 2017;

Lloyd-Fox et al., 2012; Lloyd-Fox et al., 2014). They compared neural responses evoked by human non-speech vocalisations (i.e. laugh, cry, yawn and cough) with those elicited by environmental sounds (i.e. running water, rattles and bells). Their findings suggest that infants aged 4 months and younger had not developed distinct responses to human vocalisations versus other matched sounds. By 4-8 months, they reported distinct temporal responses both to human vocalisations and to non-human sounds and suggested this may represent a transition period (Lloyd-Fox et al., 2017; Lloyd-Fox et al., 2012). Lloyd-Fox and colleagues' findings further suggest that infants from the age of 9 months have robust and stable neural responses to human vocalisations compared to non-human sounds located in the right temporal cortex (Lloyd-Fox et al., 2017). Lloyd-Fox's longitudinal findings were in line with those of Dehaene-Lambertz and colleagues' fMRI study of 2-month-old infants in which no significant cortical response differences were found to human voices and environmental sounds (Dehaene-Lambertz et al., 2010). fMRI studies also support Lloyds-Fox et al.'s developmental time course for the processing of auditory social stimuli, in which 3- to 7- month old infants' temporal cortex responded to human vocalisations and non-human sounds during passive listening (Blasi et al., 2015; Blasi et al., 2011). Grossmann et al.'s fNIRS study, which compared infant brain activation to human vocalisations (neutral speech and non-speech) versus non-vocal sounds, reported overlapping findings with those from Lloyd-Fox et al.: in 7-, but not 4-month-old infants, bilateral superior temporal cortical activation was significantly greater to human vocalisations than to non-vocal sounds, suggesting that sensitivity to the human voice emerges sometime between 4 and 7 months of age (Grossmann, Oberecker, et al., 2010).

In summary, previous research findings suggested that infant cortical activation to human speech or emotional stimuli may develop between 4-8 months of age and become stable from 9 months. However, it remains unclear whether speech and emotional components within vocalisations influence infant temporal cortical activation. In addition, very little is known about the developmental trajectory of neural processing of emotionally neutral, non-speech voice stimuli. The present study aimed to address this gap in the literature and to examine infant neural responses to emotionally neutral non-speech voice and non-vocal environmental sounds in infants longitudinally at 6, 9 and 12 months using fNIRS. In line with previous research, we predicted that at 6 months infants would show increased neural responses to non-vocal sounds compared to the human voice, and that such enhanced neural responses to the human voice would be robust at 9 and 12 months.

7.3 Methods

Participants

Forty infants of healthy mothers were recruited to the present study from three Manchester community health centres. All infants were born at normal birth weight (>2500 g); 39 were full term (37–42 weeks gestation); one was born at 36 weeks gestation (corrected gestational age). Infants were typically developing and had no hearing difficulties according to parental report. Mothers consented on behalf of their infants. The UK National Health Service ethics committee approved the study (REF: 15/NW/0684).

A total of 40 infants were tested at 6 months (20 girls and 20 boys, aged between 175 and 214 days, M = 189.48 days, SD= 9.27). From those 40 infants, a total of 39 were tested at 9 months after 1 infant dropped out (19 girls and 20 boys, between 263 and 302 days, M = 279.08 days, SD = 9.46). An additional infant dropped out at 12 months, therefore a total of 38 infants were tested at 12 months (18 girls and 20 boys, between 360 and 394 days, M = 377.24 days, SD= 8.61). The flow chart in Figure 7.1 describes the longitudinal participation.

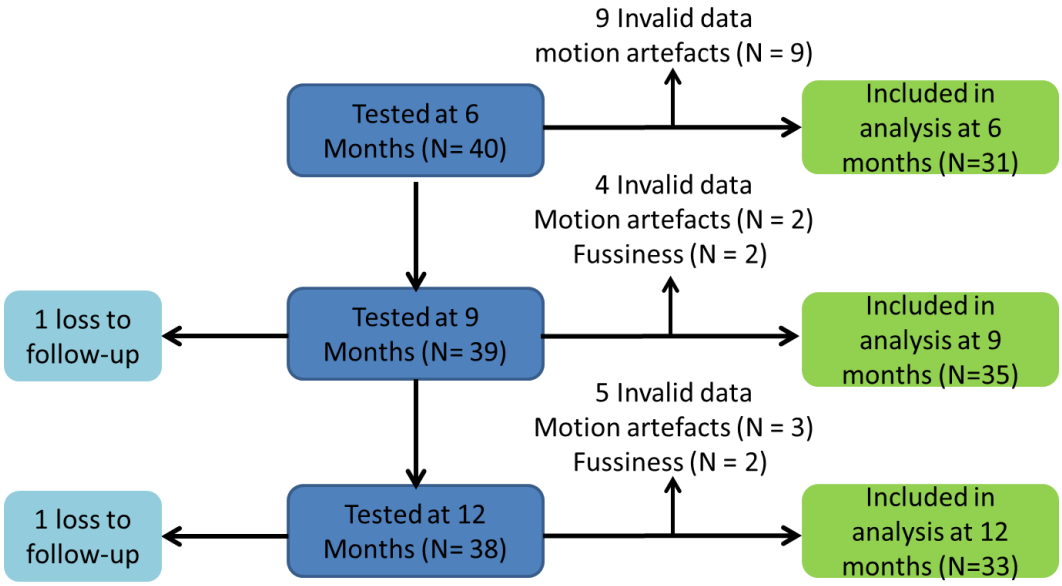


Figure 7.1. Flow charts of participant retention and infants included and excluded in the analysis at 6, 9, and 12 months.

Experimental paradigm and procedure

Figure 7.2 presents the experimental paradigm used at all three time points. Infants wore the NIRS headband, sitting in their mothers' laps in front of a laptop during the task and listened to the acoustic stimuli. The task started with a 20 sec rest period, followed by a 5 sec trial presented through loudspeakers (SPL = 70dB). Consistent with previous research (Grossmann, Oberecker, et al., 2010), a 5 sec silent cartoon was played along with each trial to attract infants' attention and reduce motion artefacts. After each trial, a 10 sec silent baseline along with the blurred cartoon was presented. The same video was played for both voice and non-vocal conditions. The task was presented with PsychoPy software (Peirce, 2007).

Sound stimuli

Two types of stimuli from a validated battery (Capilla et al., 2013) were presented to infants: (a) in the voice condition there were 5 female, non-speech, neutral

vocalisations (e.g. interjection ‘ah’, yawn and etc.); (b) in the non-vocal condition there were 5 environmental sounds (e.g. sounds of running water, a bell, a horn and etc.). All stimuli were normalised to the same duration and mean intensity with Praat sound-analysis software (Boersma & van Heuven, 2001). Each condition was repeated 8 times, the length of the testing session was 4 minutes 20 seconds. The stimuli, paradigm and measurement procedures were the same at all three measurement time points.

fNIRS Data acquisition

The infants’ cerebral responses were recorded with a multichannel NIRS data collection system. The system was built by Biomedical Optics Research Laboratory, (Dept. of Medical Physics and Bioengineering, University College London) and applied with 780nm and 850nm continuous wavelengths and 10 Hz sampling rate (Everdell et al., 2005). Two detectors and six sources formed 12 source-detector pairs in each hemisphere, and were distributed at temporal regions, which have been shown to be voice sensitive in previous research in infants (Grossmann, Oberecker, et al., 2010; Lloyd-Fox et al., 2012; Pena et al., 2003; Taga & Asakawa, 2007) and adults (Belin et al., 2000; Ethofer, Anders, Erb, et al., 2006; Grandjean et al., 2005). To achieve the best spatial sensitivity profile for infants (Fukui et al., 2003), the distances between source and detectors were fixed between 1.5 and 2.5 cm. Channels were distributed according to the 10-20 system and attached to a custom-made Velcro headband. According to the infants’ head growth standards from the World Health Organisation (World Health Organization (WHO), 2003) and from one previous longitudinal pilot study (Lloyd-Fox et al., 2017), there are no significant changes in the head circumference of 6- to 12-months-old infants. Therefore, the application of a fixed source-detector array across three age time points was

reasonable and practical (Lloyd-Fox et al., 2017). In each occasion, the headband was adjusted by calculating the distance between the glabella and the ear, ensuring that T3 and T4 were between the two bottom sources in each hemisphere. This procedure was carried out for all the infants at each time point. The locations of the channels and the channel positions with respect to the 10-20 system are presented in Figure 7.2.

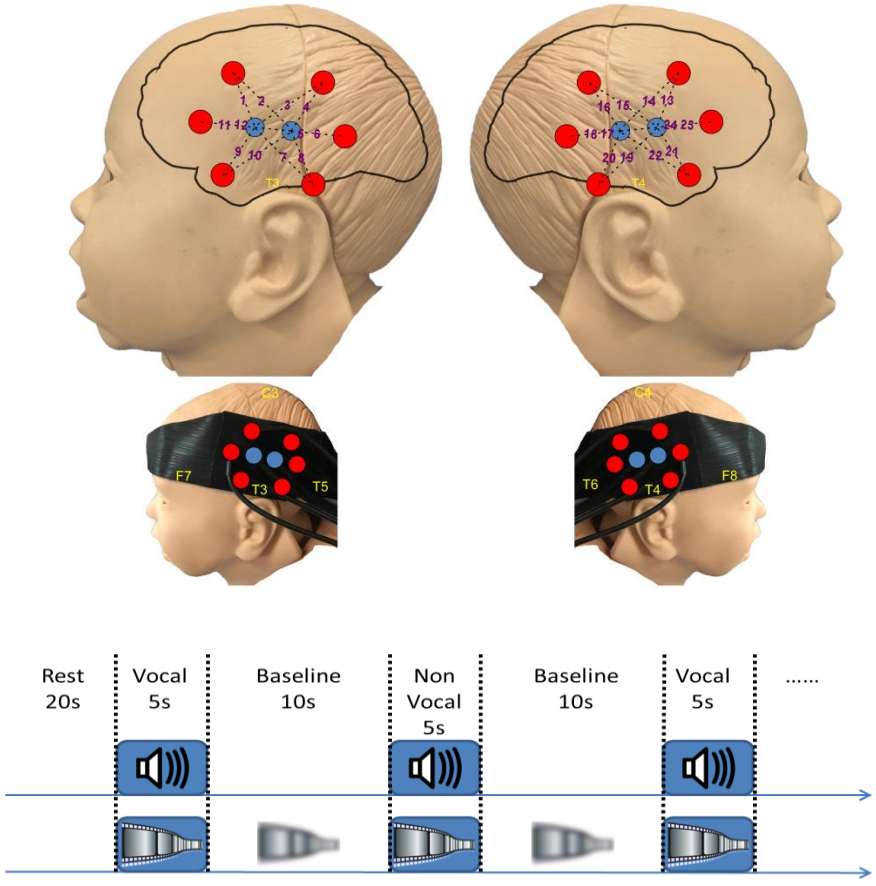


Figure 7.2. Study experimental task design and channel distribution. The head model illustrates the source-detector distribution where red dots represent sources (6 in each hemisphere) and blue dots represent detectors (2 in each hemisphere). Sources and detectors form 12 recording channels in each hemisphere, which are marked in purple numbers (upper head models), and are held by Velcro headband. Scalp

landmarks with respect to 10-20 system are marked in yellow (middle head models). The bottom streamline demonstrates the timeline of the experimental task.

Data analysis

Infant behaviour was video-recorded during the experimental task. A member of the research team viewed the videos and coded the number of trials in which an infant attended to the screen without large motion artefacts. Four out of eight trials per condition were set as a criterion for inclusion of each infant dataset. As a result, 31 datasets were included in the 6 months' time point, 35 datasets were included in the 9 months' time point and 33 datasets were included in the 12 months' time point. This inclusion rate is consistent with criteria for inclusions in previous infant NIRS studies (Lloyd-Fox et al., 2010).

All the datasets included in the analysis were filtered at 0.01 to 0.5Hz with a 3rd order Butterworth filter to eliminate slow drifts, instrument noise and physiological artefacts, such as heartbeats (Cooper et al., 2012; Fox et al., 2013; Grossmann, Oberecker, et al., 2010). The remaining artefacts were identified on a channel by channel basis with the algorithm 'hmrMotionArtifactByChannel' implemented in the HOMER2 NIRS toolbox (version 2.1, http://homer-fnirs.org/, Huppert et al., 2009 (Huppert et al., 2009)). Within the time interval (tMotion), if the change of the signal amplitude exceeded the threshold (AMPthresh) or the standard deviation changes were greater than a factor (STDEVthresh) multiplied by original channel standard deviation, the time period (tMask time before and after the motion artefact) was marked as artefact. The time period of motion artefact within the channel was corrected with a cubic spline interpolation algorithm with p set to 0.99 as recommended (Cooper et al., 2012; Scholkmann et al., 2010). Since the algorithm

works on a channel by channel basis, the actual standard deviation threshold for the motion artefact varies according to the standard deviation of the original channel; the setting of the STDEVthresh is the multiplication factor rather than a fixed threshold (i.e. in the current study the standard deviation threshold is 20*standard deviation of the channel). This means that the standard deviation threshold varies from channel to channel and subject to subject. All the values were set as follows: tMotion=5s; tMask=1s; STDEVthresh=20; AMPthresh=5.

After pre-processing, data were converted to Oxy- and Deoxy-Haemoglobin concentration changes in HOMER2 and averaged across trials in the same condition within each dataset, with the time window of 1 sec before and 15 sec after the stimulation onset. The averaged time course of each channel was corrected by subtracting the mean of the 1 sec before the stimulation. The analysis focused on Oxy-Haemoglobin concentration changes as the most sensitive indicator of changes in cerebral blood flow. Kolmogorov-Smirnov tests were conducted to test the normality of the haemodynamic time courses, results showed that p values were larger than 0.05, which indicate that haemodynamic data were normally distributed. Based on earlier work showing that the haemodynamic response reaches the peak around 2 to 4 sec post stimulus (Brigadoi et al., 2014), we targeted a time window of 2 sec to 9 sec after stimulus onset. Mean amplitudes of cortical haemodynamic responses (Oxy- and Deoxy-Haemoglobin waveforms) were averaged over the time window of 2 sec to 9 sec after stimulus onset. The averaged haemodynamic responses to the expression conditions (voice and non-vocal) were evaluated with paired sample t-tests.

7.4 Results

Of the tested 40 six-month-old infants, 31 datasets were included in the final analysis (13 girls and 18 boys, between 180 and 214 days, M = 190.58 days, SD=9.59). From the tested 39 infants at 9 months, 35 datasets were included in the analysis (16 girls and 19 boys, between 264 and 296 days, M = 278.91 days, SD = 8.70). Out of the tested 38 infants at 12 months, 33 datasets were included in the analysis (16 girls and 17 boys, between 360 and 394 days, M = 376.33 days, SD = 8.64). Derivation of the final full analysis sample is described in Figure 7.1.

At 6 months, statistical analysis of Oxy-haemoglobin concentration changes to voice and non-vocal conditions revealed a significant difference in channel 1, in the left hemisphere, in which non-vocal sounds evoked stronger responses than voice ($t(30) = 2.45$, $p = 0.021$, uncorrected, Figure 7.3, Table 7.1). Deoxy-Haemoglobin concentration changes complemented Oxy-haemoglobin concentration changes, showing significantly greater decrease in changes to non-vocal sounds compared to voice sounds ($t(30) = 2.43$, $p = 0.021$, uncorrected, Table 7.2).

At 9 months, significant Oxy-haemoglobin concentration changes were found in channel 17 and 13 in the right hemisphere, in which Oxy-haemoglobin concentration changes were significantly increased in the non-vocal condition compared to the voice condition (channel 17, $t(34) = 2.33$, $p = 0.026$; channel 13, $t(34) = 2.56$, $p = 0.015$, uncorrected, Figure 7.4, Table 7.1). Deoxy-haemoglobin concentration changes showed a similar response trend. There were significantly greater decreased changes to the non-vocal sounds than to the voice in channels 17 and 13 (channel 17, $t(34) = 2.95$, $p = 0.006$; channel 13, $t(34) = 3.05$, $p = 0.004$, uncorrected, Table 7.2).

At 12 months, there was no significant difference in either Oxy-or Deoxy-haemoglobin concentration changes between non-vocal and voice conditions.

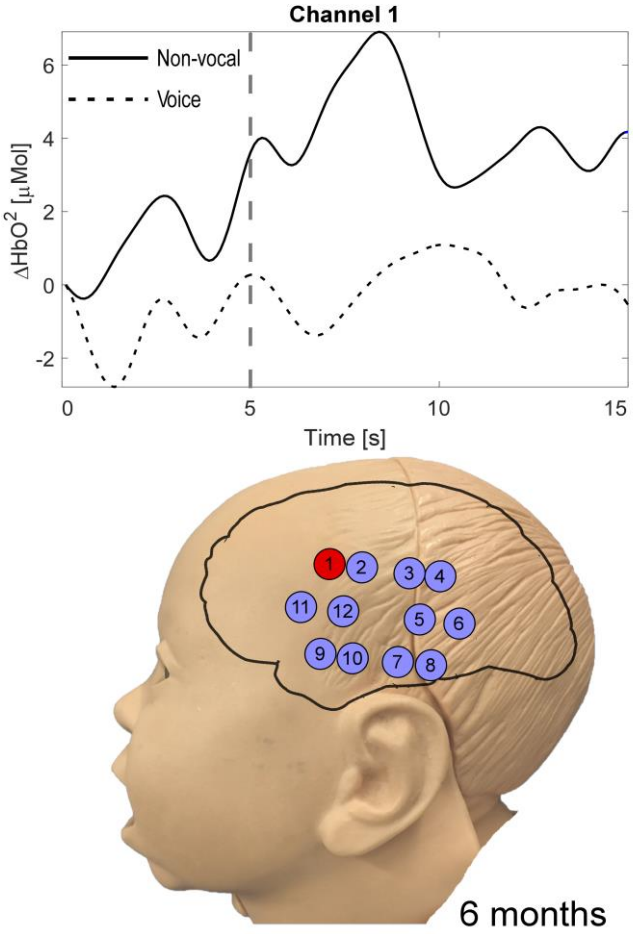


Figure 7.3 Averaged time courses of Oxy-Haemoglobin concentration changes across all datasets in 6-month-olds in channel 1 (in the left hemisphere) per condition (voice in dashed line, non-vocal in solid line) in the time period of 15 sec (5 sec stimulus and 10 sec baseline). The channel location is marked in the lower panel. The stimulus end time is marked by the vertical grey dashed line. The time (in sec) and change in amplitude (μMol) are in the x and y axis respectively.

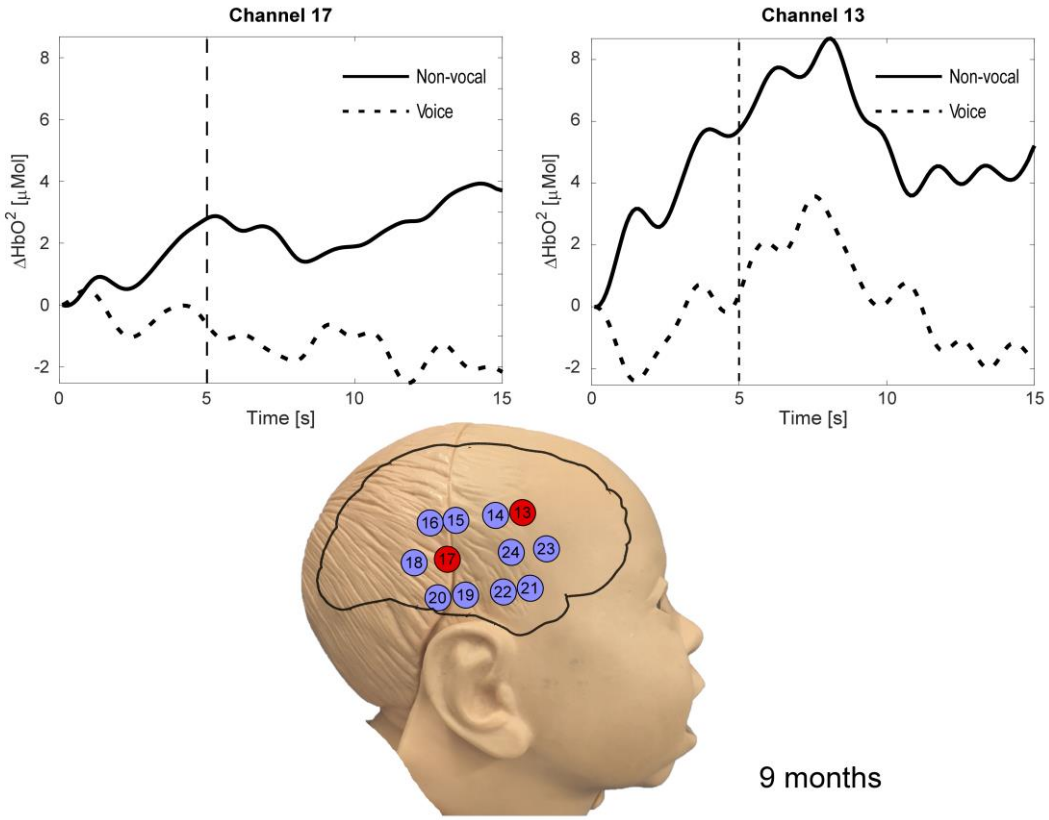


Figure 7.4 Averaged time courses of Oxy-Haemoglobin concentration changes across all datasets in 9-month-olds in channel 13 and channel 17 (in the right hemisphere) per condition (voice in dashed line, non-vocal in solid line) in the time period of 15 sec (5 sec stimulus and 10 sec baseline). The channel location is marked in the lower panel. The stimulus end time is marked by the vertical grey dashed line. The time (in sec) and change in amplitude (μMol) are in the x and y axis respectively.

Table 7.1 Oxy-Haemoglobin concentration changes comparisons between conditions

Channel	Condition	Mean ± SEM	t	p	Comparison
6 months					
Channel 1	Non-vocal	3.69 ± 1.38	2.45	**0.021***	Non-vocal > voice
	Voice	-0.54 ± 1.05			
9 months					
Channel 17	Non-vocal	1.93 ± 1.00	2.33	**0.026***	Non-vocal > voice
	Voice	-0.92 ± 0.90			
Channel 13	Non-vocal	6.25 ± 1.45	2.56	**0.015***	Non-vocal > voice
	Voice	1.19 ± 1.21			

*p < 0.05

Table 7.2 Deoxy-Haemoglobin concentration changes comparisons between conditions

Channel	Condition	Mean ± SEM	t	p	Comparison
6 months					
Channel 1	Non-vocal	-1.57 ± 1.39	2.43	**0.021***	Non-vocal < voice
	Voice	2.25 ± 1.07			
9 months					
Channel 17	Non-vocal	-1.46 ± 0.96	2.95	**0.006****	Non-vocal < voice
	Voice	2.08 ± 1.04			
Channel 13	Non-vocal	-3.77 ± 1.54	3.05	**0.004****	Non-vocal < voice
	Voice	1.75 ± 1.41			

*p < 0.05; ** p < 0.01

7.5 Discussion

The present study provides the first longitudinal observation of infant neural responses to emotionally neutral non-speech voice compared to those to non-vocal sounds across three timepoints in the first year of life. We intended broadly to replicate previous findings, which suggested an increase in neural responses to human voice and non-vocal stimuli before the age of 9 months as well as the presence of robust neural responses to human voice stimuli after the age of 9 months. Further, we intended to extend these findings and anticipated there would be similar infant neural response patterns to emotionally neutral non-speech voice and environmental sounds as in previous studies. Findings from the present study partially supported the hypotheses. We found stronger temporal cortical responses to non-vocal sounds than to emotionally neutral non-speech voice stimuli at both 6 and 9 months similar with findings from previous studies in 3- to 8-month-old infants (Lloyd-Fox et al., 2017; Lloyd-Fox et al., 2012). At 6 months, non-vocal stimuli evoked greater neural activation in the left temporal cortex compared to neutral non-speech voice stimuli; at 9 months, non-vocal stimuli evoked greater neural activation in the right temporal cortex compared to neutral non-speech voice stimuli. Surprisingly, we did not find greater neural responses to voice than to non-vocal sounds at any time point.

In the current study, we did not find increased neural activation to voice than to environmental sounds, unlike previous infant studies (Blasi et al., 2011; Grossmann, Oberecker, et al., 2010; Lloyd-Fox et al., 2017; Lloyd-Fox et al., 2012). This distinction in findings may be derived from the difference in the stimuli. The present study compared non-speech emotionally neutral vocalisations to non-vocal sounds, while previous studies either used emotional vocalisations or speech to compare with

non-vocal sounds. This finding may be interpreted as: the preference for unfamiliar, non-speech, emotionally neutral voice sounds has not yet developed in infants, before the age of 9 months compared to adults (Belin et al., 2000; Grandjean et al., 2005). The missing effect at the age of 12 months indicates the less preference to non-vocal sounds, or the undifferentiated neural processing of the stimuli from two conditions. This finding may suggest that after the age of 9 months, infants start to differentiate non-speech, emotionally neutral vocalisations compared non-vocal sounds. However, this assumption requires further evidence from infants after the first year of life.

The current study applied emotionally neutral, non-speech vocalisations in the voice condition, while previous studies have either explicitly or implicitly used stimuli with emotional content (Blasi et al., 2011; Lloyd-Fox et al., 2017; Lloyd-Fox et al., 2012) or speech-embedded voice stimuli (Grossmann, Oberecker, et al., 2010). Although from birth, infants (and even as fetuses) are able to discriminate voices from parents vs unfamiliar adults (Kisilevsky et al., 2003; Ockleford et al., 1988), their interest is lost if the voice lacks prosody and intonation (Mehler et al., 1978; Singh et al., 2002). These findings suggest that the infant neural responses to voice may depend on familiarity and prosody. Neuroimaging studies show that emotional vocalisations evoke greater awareness compared to neutral vocalisations in infants (Grandjean et al., 2005; Grossmann, Oberecker, et al., 2010; Grossmann et al., 2005), children (Chronaki, Benikos, et al., 2015; Chronaki, Garner, et al., 2015) and adults (Ethofer, Anders, Wiethoff, et al., 2006; Fecteau et al., 2007). It is worth considering that infant-direct speech, to which infants are exposed, is characterised by overt emotional vocalisations. Consequently, infants are more familiar with, and prefer, emotion-embedded vocalisations than adult-direct speech (Saint-Georges et

al., 2013; Singh et al., 2002; Werker & McLeod, 1989). The stimuli in the voice condition were emotionally neutral, non-speech vocalisations, which did not share similarities of emotion or speech with infant-directed speech, and therefore, did not attract the infants' interest or attention. Even though pre-linguistic infants have limited understanding of speech, their ability to understand their native language strengthens and becomes specialised from the age of 6 months (Kuhl, 2004; Kuhl et al., 2014). Specialised speech perception is not only evident in behavioural responses, but also in localised neural responses. Supporting evidence from Minagawa-Kawai et al. (2011) showed 4-month-old infants' neural responses to speech (native and non-native) were lateralised to the left temporal cortex; and non-speech emotional vocalisations were localised to the right temporal cortex (Minagawa-Kawai, van der Lely, et al., 2011). Thus, infants may preferentially process maternal and/or emotional vocalisations at an early age, but this only generalises to vocal expressions by other people later on.

The increased temporal activation to non-vocal compared to voice stimuli between 6 and 9 months of age are comparable to the findings of previous studies in infants at the ages of 3 to 8 months (Blasi et al., 2011; Grossmann, Oberecker, et al., 2010; Lloyd-Fox et al., 2012). Infants' enhanced neural responses to non-vocal stimuli may represent their heightened attention to novel sounds.

It is possible that non-vocal environmental sounds (e.g. bell, horn and running water) were more novel and attractive to infants compared to emotionally neutral non-speech human voice (yawn and neutral 'ah'). Infants may perceive the environmental sounds to be similar to toy sounds, for example, the sound of a horn resembles the sound of a toy car or a toy train. Infants may also find the sound of a bell novel and alerting. For infants as young as 6 and 9 months old, without sufficient knowledge of

language, environmental sounds in contrast to neutral voice may be more meaningful and have evoked stronger neural responses. It is worth considering that infant-direct speech, which infants are exposed to, is characterised by overt emotional vocalisations. Consequently, infants are more familiar with, and prefer, emotion embedded vocalisations than adult-direct speech (Saint-Georges et al., 2013; Singh et al., 2002; Werker & McLeod, 1989). The stimuli in the voice condition were emotionally neutral, non-speech vocalisations, which did not share similarity of emotion or speech with infant-directed speech, and therefore, did not attract the infants' interest or attention.

At birth, the infant brain is less connected and specialised, and gradually but relatively rapidly undergoes a "fine-tuning" process to become a differentiated and specialised dependent on an interaction within the brain regions and the infant's environmental/ maternal experience postnatally (See Grossmann et al., 2007; Kolb & Gibb, 2011). From the age of 4 months, studies have reported stronger neural responses to voice stimuli in contrast to non-vocal sounds, and stronger neural responses to non-vocal versus voice stimuli (Blasi et al., 2011; Lloyd-Fox et al., 2017; Lloyd-Fox et al., 2012), suggesting that infants have developed distinct neural responses to human and non-human sounds. However, these responses have not been reported to be specialised. The literature suggests that, by the end of the first year of life, infant neural responses to non-vocal sounds become weaker whereas responses to voice become more robust (Lloyd-Fox et al., 2017). Our findings of neural responses to non-vocal sounds followed a similar developmental trend: there was stronger brain activation to non-vocal sounds than to voice stimuli at 6 and 9 months of age, which diminished by the age of 12 months.

There are several limitations in the present study: first, the current experimental paradigm did not allow us to know whether the heightened activations in temporal regions to non-vocal sounds were because of the less developed infancy voice perception, or were they representing attentional process. Second, there is a lack of a speech or emotional vocalisation condition, in order to directly compare with emotionally neutral non-speech voice and with non-vocal sounds. Third, neural responses were only measured within infant temporal regions. Other brain regions that may be activated by voice and non-vocal stimuli were not investigated, such as frontal and parietal regions.

The present study measured infant neural responses to emotionally neutral non-speech voice and non-vocal sounds with fNIRS at the age of 6, 9, and 12 months. Infants showed significantly increased temporal cortical activation to non-vocal sounds compared to voice stimuli at the age of 6 and 9 months. Our findings may suggest that specialised neural responses to neutral non-speech voice develop throughout the first year of life. Infant neural responses to voice may be dependent on the medium via which the voice is carried (i.e. speech or emotion). The findings also suggest that the novelty of the stimuli and the infant's familiarity with the stimuli may play an important role in infant auditory perception. Further studies are needed to replicate the present paradigm beyond the first year of life, and to provide comprehensive interpretation to infant neural responses. Future studies observing infant responses to voice should take stimuli properties into consideration. Elements such as speech, emotion, novelty, and the infant's familiarity with the sounds may lead to the diversity in results and hinder comparisons between studies.

Chapter 8 General Discussion

8.1 Overview

The general aims of this PhD were (a) to develop and pilot an fNIRS protocol that can be acceptably and feasibly applied in a routine clinical setting in order potentially to monitor and evaluate the neurological underpinnings of early language and social-emotional development; (b) to track longitudinal changes in neural correlates of human voice and emotional vocalisation processing in typically developing infants at the age of 6, 9 and 12 months; (c) to explore the role of early maternal caregiving behaviour as a possible mediator in the healthy development of voice and vocal emotion processing in infants.

This thesis features 5 papers: paper 1 (in Chapter 3) reported infant fNIRS data analysis approaches; paper 2 (Chapter 4) reported the feasibility and acceptability of using infant fNIRS longitudinally in a clinical setting; paper 3 (Chapter 5), reported vocal emotion processing in 6-month-old infants and its correlation with maternal caregiving behaviour; paper 4 (Chapter 6) reported longitudinal infant neural processing of emotion vocalisations; paper 5 (Chapter 7) reported longitudinal infant neural processing of human voice and non-vocal sounds.

This chapter will summarise and discuss findings from this PhD, review the strengths and limitations of the work presented in previous chapters, along with implications and suggestions for future studies.

8.2 Key findings

8.2.1 The acceptability and feasibility of using infant fNIRS longitudinally in a clinical setting

The present study recruited 40 healthy mothers and their infants who successfully completed two fNIRS experimental tasks (voice versus non-vocal and vocal emotion tasks) in the same infants at the age of 6, 9 and 12 months. This portable imaging procedure was successfully applied in a routine community clinical setting, i.e. at local children's centres. Acceptability was tested in a number of ways: indirectly by the retention rate of mother-infant pairs at each time point, and directly by asking mothers for their feedback on the procedure and details of the paradigm after each session. The retention rate of the infants and their mothers across the three-time periods was 95%; this suggests that recruited mothers were successfully engaged in the study and that the procedures were well tolerated. Mothers also reported a high satisfaction rate with the fNIRS measurement for their infants and were happy to engage and return for the follow-up visits. Furthermore, our carefully developed experimental paradigm and measurement procedures allowed for very low data loss and secured 84.7% and 75.2% of fNIRS data for the voice versus non-vocal and vocal emotion task respectively. This provides evidence for the feasibility of our fNIRS paradigm in non-research settings and of our data collection and analysis methods. Also, the Bayley infant language scores (Bayley, 2006) suggest that participating infants were developing typically with respect to their language acquisition and did not have any language delays in their first year.

8.2.2 Infant vocal emotion neural processing

Our first longitudinal investigation was on the infant haemodynamic responses to emotional, non-speech vocalisations at three time points (6, 9 and 12 months) in the first year (Chapter 5, paper 4). A key finding was the consistency of activations in the temporal cortex throughout the first year, in particular to happy voices. Increased left superior temporal activation to angry compared to neutral was only found at 6 months. Emotional vocalisations evoked temporal cortical activation regions differed at the ages of 6, 9 and 12 months. Specifically, activation was bilateral and superior at 6 months, whereas it was unilaterally localised inferiorly to the right at 9 months and to the left at 12 months. In addition, there was an age effect in the longitudinal infants' neural responses to vocal emotional stimuli, specifically right temporal activation to angry prosody increased with age in the second half of the first year.

8.2.3 Infant voice versus non-vocal neural processing

Our second longitudinal investigation was on the infant haemodynamic responses to non-speech voice (that was emotionally neutral) in contrast to non-vocal stimuli at three time points in the first year (Chapter 7, paper 5). Our results indicate greater temporal cortical activation in response to non-vocal sounds than to non-speech voice stimuli at both 6 and 9 months, but not at 12 months. At 6 months, non-vocal stimuli evoked greater neural activation in the left superior temporal cortex; at 9 months, non-vocal stimuli evoked greater neural activation in the right superior and inferior temporal cortex. Unlike other similar studies, we did not find greater neural responses to voice than to non-vocal sounds at any time point.

8.2.4 Maternal caregiving behaviour and infant vocal emotion sensitive neural responses

At 6 months, we measured maternal caregiving behaviours, from recorded mother-infant play interactions of our sample, using the Manchester Assessment of Caregiver-Infant Interaction (MACI (Wan, 2015; Wan et al., 2017)). We then tested whether maternal caregiving behaviour ratings were associated with infant cortical activation patterns to vocal emotions. There was a significant correlation between 6-month-old infants' neural responses to vocal anger (neural responses to angry minus neural responses to neutral) and maternal directiveness. Therefore, infants with stronger neural discrimination to angry vocal stimuli (in the left superior temporal cortex) had experienced more directive interactions. However, we did not find a link between maternal sensitive caregiving behaviour and infant neural responses to emotional vocalisations.

8.2.5 fNIRS data analysis

Two approaches were piloted to analyse infant fNIRS data: the standardised analysis approach and ICA implanted analysis method. The standardised analysis approach is the commonly accepted way of analysing fNIRS data. The ICA approach is sensitive to high noise levels and significantly improved the data quality, both in the task-related haemodynamic time courses and in the signal-to-noise ratio (SNR). However, the ICA approach has to be balanced against the loss of amplitude information.

8.3 General discussion

8.3.1 The acceptability and feasibility of using fNIRS longitudinally in infants

The repeated application of fNIRS in a routine clinical setting may offer the potential for the use of clinical neuroimaging in infants and children. This project has shown not only that is it feasible to administer fNIRS in a routine community setting, but also that it is well tolerated even in infants as young as 6 months; and that parents are willing and able to attend repeated appointments. Furthermore, and importantly, by 12 months of age, the fNIRS remains well tolerated even as the infant becomes more physically and mentally active. This is an important advance compared to other neuroimaging approaches because it allows for the testing of functional networks in awake infants as they develop over time.

8.3.2 Infant neural processing of human voice and non-vocal sounds

Increased neural responses to non-vocal and vocal emotion stimuli were consistently detected in the temporal cortices; these activation patterns were comparable to those in previous infant human voice processing studies (Blasi et al., 2015; Blasi et al., 2011; Grossmann, Oberecker, et al., 2010; Lloyd-Fox et al., 2012). The present activation patterns were also within the temporo-frontal network for adult voice and vocal emotion processing (Belin et al., 2000; Ethofer et al., 2012; Fecteau et al., 2005; Frühholz et al., 2012; Frühholz & Grandjean, 2013; Witteman et al., 2012) and also as part of the 'social brain' (Adolphs, 2009; Grossmann & Johnson, 2007). This finding suggests that the infant brain starts to emerge adult-like neural responses to human voice stimuli in the first year of life.

However, the activation locations within temporal cortices in response to voice, non-vocal and vocal emotion stimuli were lack of consistency between 6 and 12 months of age. The different activation regions to the same auditory stimuli over time may suggest the unstable and yet specialised neural development in the first year in infants.

Moreover, the right temporal activation to angry prosody showed an age effect of neural responses to angry voices, which increased between 6 and 12 months. Although it did not survive multiple comparison corrections, the right temporal cortex also showed an age effect inasmuch as there was an increase in cortical activations to emotional vocalisations, especially greater neural responses in 12-month-old infants compared to 9-month-old infants. This age effect extends previous findings which suggest neural responses to voice and non-vocal sounds became more distinct with the increasing age in 3-to 7-month-old infants (Blasi et al., 2011; Lloyd-Fox et al., 2012). The age effect in infants' haemodynamic responses is also in line with EEG evidence that infants' alpha power frequency increases with age (Marshall, Bar-Haim, & Fox, 2002; Michel et al., 2015).

The present experimental paradigm allowed for the comparison between positive and negative vocal emotions processing in infants. It is well-known that infants prefer happy vocalisations from birth (Mastropieri & Turkewitz, 1999), and respond differently to positive and negative infant-directed speech from the age of 5 months (Fernald, 1993). At the age of 6 months, the infant still prioritises attention and response to happy speech over neutral and negative speech (Singh et al., 2002). It is possible that the infant preference for positive vocal emotion develops earlier, from birth, and that responses to negative emotion develop later, from around 5 months of age. In line with infant behavioural responses to emotional vocalisations, key

findings from this PhD consistently demonstrate increased neural responses to happy compared to angry prosody over time; enhanced left superior temporal responses to angry compared to neutral voices at 6 months; and we found that right temporal activations to angry vocalisations increased between 6 and 12 months. Our findings suggest that there may be distinct neural developmental trajectories for positive and negative vocal emotion processing in infants. Particularly, the perception of positive emotion emerges earlier and develops robustly over time, while the perception of negative emotion emerges later and develops rapidly over the first year of life.

The other longitudinal investigation in the present study was to present infants with human neutral vocalisations i.e. those specified as without any emotional features. Infants showed greater temporal cortical responses to non-vocal sounds compared with human neutral, non-speech vocal sounds at 6 and 9 months of age; these increased neural responses diminished by the age of 12 months. The increased neural responses to non-vocal sounds relative to voice stimuli is broadly consistent with the previous literature in infants younger than 7 month old (Blasi et al., 2011; Grossmann, Oberecker, et al., 2010; Lloyd-Fox et al., 2012). Although we tried to match voice and non-vocal stimuli, non-vocal environmental sounds (e.g. bells, horns and running water) may have been more novel and attractive to infants compared to emotionally neutral non-speech human voice (yawn and neutral 'ah'). This may imply that infants' enhanced neural responses to our non-vocal stimuli may represent a heightened attention to novel or more interesting sounds. However, we did not find greater neural responses to the voice than to non-vocal sounds at any age point, which were reported in previous studies (Blasi et al., 2011; Grossmann, Oberecker, et al., 2010; Lloyd-Fox et al., 2012). This finding may suggest that the preference for unfamiliar, non-speech, emotionally neutral voice sounds has not yet

developed in infants, before the age of 9 months compared to adults (Belin et al., 2000; Grandjean et al., 2005). Previous studies which used speech or emotional vocalisations compared with non-vocal sounds, found greater neural responses to voice than to non-vocal sounds (Blasi et al., 2011; Grossmann, Oberecker, et al., 2010; Lloyd-Fox et al., 2012); our findings may suggest that the infant perception of human vocalisations depend on the medium through which voice is carried (i.e. human neutral speech or emotionally valent vocalisations). Infants may preferentially process maternal and/or emotionally valent vocalisations at an early age, but this preference may only generalises to other vocal expressions by other people later on.

Infants' neural responses to voice stimuli at each age point resemble those of adults'. However, infants' temporal activation patterns to voice stimuli were not stable over the first year of life and there was a significant age effect suggesting the continuously developing neural responses to human vocalisations. All these findings supported the emergence of neural activation to voice stimuli at 6 months of age and continue to be shaped and to develop between 6 to 12 months. This developmental time course is also in line with the view of 'social brain' development, which suggested that the less developed and specialised infant 'social brain' shows similar neural activations to social stimuli as in the adult brain (see Grossmann et al., 2007).

8.3.3 Maternal caregiving behaviour and infant vocal emotion sensitive neural responses

Temporal cortices, as an integral part of the social brain, undergo substantial "fine-tuning" in the early years both internally and in interaction with the environment to

develop specialised functions (Johnson et al., 2009). Neural correlates changes in the temporal processing of social stimuli over the latter half of the first year may reflect such a fine-tuning process. In line with this notion, we found a correlation between 6-month-old infants' neural responses to vocal anger and maternal directiveness. Maternal directiveness typically involves vocal and behavioural demands, intrusions and/or critical utterances. Although not all such directiveness carries vocal negativity, being the recipient of high caregiver directiveness may involve being exposed to negative emotion more often as a guide to acceptable behaviour, giving rise to a neural bias towards the processing of negative prosody. This association may suggest the potential influence of the infant's early experience on the neural processing of voice stimuli.

No evidence was found suggesting that infants who receive more maternal sensitive caregiving (versus lower sensitivity) showed a differential neural response to auditory voices, as we did not report significant associations; and there is very limited literature observing this link. Two EEG studies reported associations between maternal positive affect and infants' neural responses (Bernier et al., 2016; Swingler et al., 2017). Bernier et al. (2016) found an association between maternal positive affect during the interaction with the 5-month-old infant and later frontal resting EEG power in the infant at 10 and 24 months of age (Bernier et al., 2016). Swingler et al. (2017) reported an non-significant correlation between maternal positive affect and EEG power-related attentional cortical activity, although maternal positive affect was positively associated with infants' attention behaviour at 10 months (Swingler et al., 2017). As we did not assess the longitudinal change of maternal behaviour, or its correlation with changing infant neural responses to voice

stimuli, we were unable to investigate the longitudinal influence of maternal caregiving behaviour on the infant's neural responses to social stimuli.

8.3.4 fNIRS data analysis

This PhD study also tried to optimise infant fNIRS data analysis. Neuroimaging data from infants tend to contain more artefacts compared to data from healthy adult populations, the situation could be worse in awake infant studies. Unlike EEG and fMRI, there is a lack of optimum processing procedure in fNIRS data analysis (Hocke et al., 2018; Pinti et al., 2018). Infant studies have applied various analysis approaches to secure clean data (Gervain et al., 2008; Grossmann, Oberecker, et al., 2010; Lloyd-Fox et al., 2012; Minagawa-Kawai, van der Lely, et al., 2011; Zhang et al., 2017), whereas the different selection of algorithms or parameter settings may significantly impact on the results, which may handicap the direct comparisons between studies (Hocke et al., 2018). The suggested ICA implanted analysis approach intended to post-process the data that has been processed with normally used processing procedures. ICA has significantly improved SNR after converting neural responses to univariate, zero-mean time courses. In order to apply ICA in infant fNIRS data analysis, more work is needed to resume the amplitude information in the independent components. However, current tasks are not ideal as infant haemodynamic response time courses for the stimuli are still under investigation. ICA approach will benefit from piloting on well-established fNIRS data, such as from adult data. Therefore the infant fNIRS data presented in papers 3 to 5 were analysed using standardised analysis methods.

8.4 Strengths and limitations

8.4.1 Strengths:

As far as we are aware, this is the first longitudinal infant fNIRS study that has been piloted in a routine clinical setting in 40 typically developing infants over three age points of 6, 9 and 12 months. The longitudinal assessment was clearly acceptable to mother-infant pairs with extremely good retention and high satisfaction rates were adequate. The two experimental tasks were able to be undertaken reliably and feasibly within a short length of time (11 minutes) with relatively small loss of infant data. This is particularly important for a technique which may be used in the future for high risk, hard to reach groups, perhaps especially for longitudinal studies of typical and atypical neurodevelopment.

Previous studies either focused on one age group (i.e. 7 months or 8 months; (Grossmann, Oberecker, et al., 2010; Grossmann et al., 2005; Missana et al., 2017), or have averaged infants' data across ages (i.e. 3 to 7 months; (Blasi et al., 2015; Blasi et al., 2011; Lloyd-Fox et al., 2012)). There is a lack of information about the developmental trajectory of infant neural processing of human vocalisations. This study provided infant haemodynamic responses to human non-speech vocalisations (neutral and emotional vocalisations) and to non-vocal sounds at a fixed age; and longitudinally over 6 months within the same infants. Infants' neural responses to voice stimuli at different ages allowed for the observation of key developmental stages; and added to our understanding of the timeline of infant brain development.

The study paradigm provided individual comparison between neural responses to emotional and neutral vocalisations (i.e. angry vs neutral, happy vs neutral and angry vs happy), which facilitated the observation of infant neural responses to different

types of emotions over time; and provides a neural basis to the behavioural findings from previous studies (Fernald, 1993; Mastropieri & Turkewitz, 1999; Singh et al., 2002).

Previous studies provided limited evidence on infant neural processing of speech and/or emotional vocalisations. The stimuli in the present study were carefully selected human non-speech vocalisations from validated batteries (Capilla et al., 2013; Maurage et al., 2007); and were based on the infant's language and social-emotional developmental level, which were suggested by previous studies (Grossmann, Oberecker, et al., 2010; Lloyd-Fox et al., 2012; Mastropieri & Turkewitz, 1999; Singh et al., 2002; Walker-Andrews & Grolnick, 1983). Findings from the present study have made a significant contribution to the literature on infant neural processing of non-speech human vocalisations distinguishing, as they have, between responses to vocal and non-vocal, neutral and emotional stimuli.

The present study tried to explore the association between maternal caregiving behaviours and infants' neural responses to emotional vocalisations. We report the first preliminary evidence of a link between an individual infant brain's neural responsivity to vocal anger and maternal directiveness in a healthy sample. These findings assist our understanding of the influence of early social experience on the infant's functional cortical responses.

The present study used rigorous analysis strategies to exclude infant movement artefacts from the data, and also secure as many infant data as possible. On average we have secured 84.7% and 75.2% of collected infant fNIRS data for the voice and vocal emotion task respectively. These rates are higher than the average data inclusion rate of 60% (reviewed by (Lloyd-Fox et al., 2010)). In addition, we

highlight the potential use of ICA to improve signal detection and reduce data loss in infant fNIRS data analysis. The involvement of ICA improved the SNR in the data, as well as the quality of task-related neural response waveforms.

8.4.2 Limitations:

The study strengths notwithstanding, there remain a number of important limitations which need to be considered for discussion:

(i) First of all, we observed brain responses in the temporal cortical regions and did not investigate the involvement of other (e.g. frontal) regions implicated in vocal processing (Schirmer & Kotz, 2006).

(ii) We did not examine comparisons of cross-task conditions, i.e. brain responses to non-vocal and to vocal emotional stimuli. This is because of two methodological factors during infant data collection: first, infants took off their headsets after each task. Six-month-old infants took off the headsets when they needed, to have a break between tasks. In contrast, most 9- and 12-month-old infants completed both tasks without break. The decision of whether to have a break was dependent on the infant state and was made by the mother. Second, stimuli in the two tasks were from two independent sound batteries (voice and non-vocal stimuli were from Capilla et al.'s sound battery and vocal emotion stimuli were from Maurage et al.'s sound battery), between which there are differences in low-level acoustic features, such as duration and pitch.

(iii) All vocal stimuli in this study were from women who were unfamiliar to the infants. It is well-described that the infant's neural responses and behaviours are particularly sensitive to the own mother's voices (Dehaene-Lambertz et al., 2010; Walker-Andrews et al., 2011).

(iv) There was a main effect of age but did not survive correction for multiple comparison, this may due to the relatively small sample size in the longitudinal age effect analysis (N = 21).

(v) Vocal emotional stimuli in the current study consisted of one type of negative (i.e. angry) and one type of positive (i.e. happy) valence, happy and angry emotion may not be able to represent the whole positive and negative category respectively.

(vi) Responses to non-vocal stimuli were presented uncorrected in the current analysis, and there is a requirement of replication from in a larger sample, with rigorous statistical analysis method.

(vii) The current experimental paradigm did not allow us to know whether the heightened activations in temporal regions to non-vocal sounds were because of the less developed infancy voice perception, or were they representing the attentional process.

(viii) Since maternal and infant anger was not measured directly in this study, whether more directive caregivers actually used more angry vocal expressions and/or whether their infants experienced more anger (or irritation) as a result of their social interactions is unknown.

(ix) We did not measure the longitudinal change of maternal caregiving behaviour, nor did we link it with longitudinal changes of infant neural correlates to voice stimuli.

(x) ICA implanted data analysis requires further work to reconstruct rescaled task-related components into waveforms.

8.5 Implications and suggestions for future studies

This PhD project successfully piloted vocal emotion and voice versus non-vocal fNIRS paradigm in a non-research, routine community clinical setting in typically developing infants at their 6, 9 and 12 months of age. The fNIRS paradigm, longitudinal measurement procedures as well as data analysis approach are transferable to future infant or child neural developmental observations. Future studies should explore further the use of ICA and related techniques to optimise data acquisition in infant fNIRS. Future studies will benefit from improving the customised headgear, especially the development of infant-/ child-friendly headgear; as well as a comfortable testing environment, for example, testing in home settings. Future research should also consider expanding the stimulation paradigm to include other types of positive and negative emotional voices; and involve voices from own mothers that are meaningful to infants, which may evoke stronger responses than unfamiliar women's voices.

Longitudinal infant neural responses to human non-speech voice stimuli (emotional and neutral vocalisations) added novel evidence to the literature and may assist early identification of neural basis for early language and social-emotional development. Findings from the present study highlight the need for the exploration of neural processing of human vocal expressions beyond the first year of life and its correlations with language and social-emotional behavioural development outcomes.

Findings from this study and future replications may have important implications for infants at risk for atypical neurodevelopmental outcomes. Individual differences in infants' responsiveness to vocal emotion may be valuable in the assessment of at-risk groups, such as those with parental mental illness, or living in poverty and

deprivation. These children represent a significant and growing group. However, most will remain resilient and so the ability to detect those most likely to express risk represents a potentially important biomarker. Future research detailing neurodevelopment in at-risk infants compared to healthy controls is needed.

In summary, future studies should focus on improving the comfort of fNIRS system to make studies more child-friendly (e.g. wireless/lightweight and soft headband), along with reliable data analysis methods that would improve SNR and reduce data loss. Future studies would also benefit from expanding stimulation paradigm to compare other types of stimuli, as well as in participants beyond the first year of life. More importantly, the exploration of associations between infant neural responses/ correlate changes with behavioural achievements would be advantageous for the potential future clinical application. Last but not least, the combination of fNIRS and other psychophysiological tools (such as eye-tracking) would provide new evidence to support the interpretation of haemodynamic changes in the brain.

Although fNIRS has shown the great potential in assisting clinic diagnosis (Lloyd-Fox et al., 2013), its application in cognitive neuroscience is still in a relatively early stage. A lot of improvement on the method, and the exploration of various cognitive questions are expected in future studies. In order to make fNIRS applicable in a clinic a few considerations should be taken into account: (1) the association between neural responses and cognitive functioning/ development should be clear and robust, without confounding factors. (2) Considering the huge variation in infant and child development, careful attention should be given to individual differences. (2) The assessment procedures should be simple, short which should be acceptable by infants and children, as well as feasible for clinicians to carry out. (3) Data analysis methods should be reliable, and achievable by clinicians without the requirement of special

data analysis skills. (4) The assessment outcomes should be accurate, and the presentation of results should be easy to understand.

8.6 Conclusions

This PhD study has made significant contributions to the literature by providing a novel, scalable and acceptable longitudinal neuroimaging paradigm for future use in at-risk populations. Findings provide valuable evidence for infant neural responses to human non-speech vocalisations (emotional and neutral) and non-vocal sounds. Findings suggest there is a continuous process of "fine-tuning" in temporal cortices responding to human vocal expressions across infancy. Infants' neural responses to emotional vocalisations at the age of 6, 9, and 12 months support consistent neural activation to happy prosody and increasing right temporal responses to angry over time, which suggests a preference to and distinct developmental trajectories of emotionally positive and negative vocalisations in the first year of life. The finding of the association between individual infant neural sensitivity to vocal anger and maternal directive behaviour provides the first preliminary evidence of how experiences in early infancy may influence key aspects of cognitive development. Infants in the first year of life differentiate positive and negative emotions irrespective of speech; they also distinguish human non-speech, neutral vocalisations from non-vocal sounds, but show increased neural responses to the latter. This may suggest that the infant neural responses to human vocalisations are dependent on contextual information (e.g. emotion, speech, familiarity, novelty and etc.).

The present longitudinal study provided an acceptable and feasible measurement analysis protocol for infant neural responses to human vocalisations, which may

serve as a part of communication and social-emotional development. Future research will benefit from exploring the link between neural responses to human vocalisations and communicational social-emotion development in typically and atypically developing infants and children.

References

Adolphs, R. (2009). The social brain: neural basis of social knowledge. *Annual Review of Psychology, Vol 66, 60*, 693-716. doi:10.1146/annurev.psych.60.110707.163514

Ainsworth, M. D. S. (1978). *Patterns of attachment : a psychological study of the strange situation*. Hillsdale, N.J ; New York; London.

Akhtar, N., Dunham, F., & Dunham, P. J. (1991). Directive Interactions and Early Vocabulary Development - the Role of Joint Attentional Focus. *Journal of Child Language, 18*(1), 41-49. doi:Doi 10.1017/S0305000900013283

Alba-Ferrara, L., Ellison, A., & Mitchell, R. L. C. (2012). Decoding emotional prosody: Resolving differences in functional neuroanatomy from fMRI and lesion studies using TMS. *Brain Stimulation, 5*(3), 347-353. doi:10.1016/j.brs.2011.06.004

Alba-Ferrara, L., Hausmann, M., Mitchell, R. L., & Weis, S. (2011). The Neural Correlates of Emotional Prosody Comprehension: Disentangling Simple from Complex Emotion. *Plos One, 6*(12). doi:10.1371/journal.pone.0028701

Aslin, R. N. (2007). What's in a look? *Developmental Science, 10*(1), 48-53.

Aslin, R. N. (2012). Questioning the questions that have been asked about the infant brain using near-infrared spectroscopy. *Cognitive Neuropsychology, 29*(1-2), 7-33.

Aslin, R. N., & Mehler, J. (2005). Near-infrared spectroscopy for functional studies of brain activity in human infants: promise, prospects, and challenges. *Journal of Biomedical Optics, 10*(1). doi:Artn 011009. Doi 10.1117/1.1854672

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70*(3), 614-636. doi:Doi 10.1037/0022-3514.70.3.614

Barrett, K. C., Campos, J. J., & Emde, R. N. (1996). Infants' use of conflicting emotion signals. *Cognition & Emotion, 10*(2), 113-135. doi:Doi 10.1080/026999396380295

Barrett, K. C., Zahn-Waxler, C., & Cole, P. M. (1993). Avoiders vs. amenders: Implications for the investigation of guilt and shame during toddlerhood? *Cognition & Emotion, 7*(6), 481-505.

Bates, E., Bretherton, I., & Snyder, L. S. (1991). *From first words to grammar: Individual differences and dissociable mechanisms* (Vol. 20): Cambridge University Press.

Bates, E., Thal, D., & Janowsky, J. S. (1992). Early language development and its neural correlates. *Handbook of neuropsychology, 7*, 69-69.

Baum, K. M., & Nowicki, S. (1998). Perception of emotion: Measuring decoding accuracy of adult prosodic cues varying in intensity. *Journal of Nonverbal Behavior, 22*(2), 89-107. doi:Doi 10.1023/A:1022954014365

Baumwell, L., TamisLeMonda, C. S., & Bornstein, M. H. (1997). Maternal verbal sensitivity and child language comprehension. *Infant Behavior & Development, 20*(2), 247-258. doi:Doi 10.1016/S0163-6383(97)90026-6

Bayley, N. (2006). *Bayley Scales of Infant and Toddler Development–Third Edition.* San Antonio, TX: Harcourt Assessment.

Beitchman, J. H., Brownlie, E. B., Inglis, A., Wild, J., Ferguson, B., Schachter, D., . . . Matthews, R. (1996). Seven-year follow-up of speech/language impaired and control children: Psychiatric outcome. *Journal of Child Psychology and*

Psychiatry and Allied Disciplines, 37(8), 961-970. doi:DOI 10.1111/j.1469-7610.1996.tb01493.x

Beitchman, J. H., Jiang, H., Koyama, E., Johnson, C. J., Escobar, M., Atkinson, L., . . . Vida, R. (2008). Models and determinants of vocabulary growth from kindergarten to adulthood. *Journal of Child Psychology and Psychiatry, 49*(6), 626-634. doi:10.1111/j.1469-7610.2008.01878.x

Beitchman, J. H., Nair, R., Clegg, M., Ferguson, B., & Patel, P. G. (1986). Prevalence of Psychiatric-Disorders in Children with Speech and Language Disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 25*(4), 528-535. doi:Doi 10.1016/S0002-7138(10)60013-1

Beitchman, J. H., Wilson, B., Johnson, C. J., Atkinson, L., Young, A., Adlaf, E., . . . Douglas, L. (2001). Fourteen-year follow-up of speech/language-impaired and control children: Psychiatric outcome. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(1), 75-82. doi:Doi 10.1097/00004583-200101000-00019

Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends in Cognitive Sciences, 8*(3), 129-135. doi:DOI 10.1016/j.tics.2004.01.008

Belin, P., & Zatorre, R. J. (2000). 'What', 'where' and 'how' in auditory cortex. *Nature Neuroscience, 3*(10), 965-966. doi:Doi 10.1038/79890

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature, 403*(6767), 309-312. doi:Doi 10.1038/35002078

Benedict, H. (1979). Early Lexical Development - Comprehension and Production. *Journal of Child Language, 6*(2), 183-200. doi:Doi 10.1017/S0305000900002245

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Methodological, 57*(1), 289-300.

Benton, A. L. (1968). Differential Behavioral Effects in Frontal Lobe Disease. *Neuropsychologia, 6*(1), 53-&. doi:Doi 10.1016/0028-3932(68)90038-9

Bergelson, E., & Swingley, D. (2012). At 6-9 months, human infants know the meanings of many common nouns. *Proc Natl Acad Sci U S A, 109*(9), 3253-3258. doi:10.1073/pnas.1113380109

Bergelson, E., & Swingley, D. (2015). Early Word Comprehension in Infants: Replication and Extension. *Lang Learn Dev, 11*(4), 369-380. doi:10.1080/15475441.2014.979387

Bernier, A., Calkins, S. D., & Bell, M. A. (2016). Longitudinal Associations Between the Quality of Mother-Infant Interactions and Brain Development Across Infancy. *Child Development, 87*(4), 1159-1174. doi:10.1111/cdev.12518

Blacher, J., Baker, B. L., & Kaladjian, A. (2013). Syndrome specificity and mother-child interactions: examining positive and negative parenting across contexts and time. *Journal of Autism and Developmental Disorders, 43*(4), 761-774. doi:10.1007/s10803-012-1605-x

Blasi, A., Lloyd-Fox, S., Sethna, V., Brammer, M. J., Mercure, E., Murray, L., . . . Johnson, M. H. (2015). Atypical processing of voice sounds in infants at risk for autism spectrum disorder. *Cortex, 71*, 122-133. doi:10.1016/j.cortex.2015.06.015

Blasi, A., Mercure, E., Lloyd-Fox, S., Thomson, A., Brammer, M., Sauter, D., . . . Murphy, D. G. M. (2011). Early Specialization for Voice and Emotion

Processing in the Infant Brain. *Current Biology, 21*(14), 1220-1224. doi:DOI 10.1016/j.cub.2011.06.009

Blonder, L. X., Bowers, D., & Heilman, K. M. (1991). The Role of the Right-Hemisphere in Emotional Communication. *Brain, 114*, 1115-1127. doi:DOI 10.1093/brain/114.3.1115

Boersma, P., & van Heuven, V. (2001). Speak and unSpeak with PRAAT. *Glot International, 5:341-347*.

Bornstein, M. H., Tamis-LeMonda, C. S., & Haynes, O. M. (1999). First words in the second year: Continuity, stability, and models of concurrent and predictive correspondence in vocabulary and verbal responsiveness across age and context. *Infant Behavior and Development, 22*(1), 65-85.

Botting, N., Durkin, K., Toseeb, U., Pickles, A., & Conti-Ramsden, G. (2016). Emotional health, support, and self-efficacy in young adults with a history of language impairment. *British Journal of Developmental Psychology, 34*(4), 538-554. doi:10.1111/bjdp.12148

Bowlby, J. (1969). Attachment, Vol. 1 of Attachment and loss: New York: Basic Books.

Brigadoi, S., Ceccherini, L., Cutini, S., Scarpa, F., Scatturin, P., Selb, J., . . . Cooper, R. J. (2014). Motion artifacts in functional near-infrared spectroscopy: A comparison of motion correction techniques applied to real cognitive data. *Neuroimage, 85*, 181-191. doi:10.1016/j.neuroimage.2013.04.082

Brown, M. A., McIntyre, L. L., Crnic, K. A., Baker, B. L., & Blacher, J. (2011). Preschool children with and without developmental delay: Risk, parenting, and child demandingness. *Journal of Mental Health Research in Intellectual Disabilities, 4*(3), 206-226.

Burgoon, J. K., Guerrero, L. K., & Floyd, K. (2016). *Nonverbal communication*: Routledge.

Cantiani, C., Riva, V., Piazza, C., Bettoni, R., Molteni, M., Choudhury, N., . . . Benasich, A. A. (2016). Auditory discrimination predicts linguistic outcome in Italian infants with and without familial risk for language learning impairment. *Developmental Cognitive Neuroscience, 20*, 23-34.

Capilla, A., Belin, P., & Gross, J. (2013). The Early Spatio-Temporal Correlates and Task Independence of Cerebral Voice Processing Studied with MEG. *Cereb Cortex, 23*(6), 1388-1395. doi:10.1093/cercor/bhs119

Caron, A. J., Caron, R. F., & MacLean, D. J. (1988). Infant discrimination of naturalistic emotional expressions: the role of face and voice. *Child Development, 59*(3), 604-616.

Carton, J. S., Kessler, E. A., & Pape, C. L. (1999). Nonverbal decoding skills and relationship well-being in adults. *Journal of Nonverbal Behavior, 23*(1), 91-100. doi:Doi 10.1023/A:1021339410262

Choudhury, N., & Benasich, A. A. (2011). Maturation of auditory evoked potentials from 6 to 48 months: prediction to 3 and 4 year language and cognitive abilities. *Clin Neurophysiol, 122*(2), 320-338. doi:10.1016/j.clinph.2010.05.035

Chronaki, G., Benikos, N., Fairchild, G., & Sonuga-Barke, E. J. S. (2015). Atypical neural responses to vocal anger in attention-deficit/hyperactivity disorder. *Journal of Child Psychology and Psychiatry, 56*(4), 477-487. doi:10.1111/jcpp.12312

Chronaki, G., Broyd, S., Garner, M., Hadwin, J. A., Thompson, M. J. J., & Sonuga-Barke, E. J. S. (2012). Isolating N400 as neural marker of vocal anger

processing in 6-11-year old children. *Developmental Cognitive Neuroscience, 2*(2), 268-276. doi:10.1016/j.dcn.2011.11.007

Chronaki, G., Garner, M., Hadwin, J. A., Thompson, M. J. J., Chin, C. Y., & Sonuga-Barke, E. J. S. (2015). Emotion-recognition abilities and behavior problem dimensions in preschoolers: evidence for a specific role for childhood hyperactivity. *Child Neuropsychology, 21*(1), 25-40.

Chronaki, G., Hadwin, J. A., Garner, M., Maurage, P., & Sonuga-Barke, E. J. S. (2015). The development of emotion recognition from facial expressions and non-linguistic vocalizations during childhood. *British Journal of Developmental Psychology, 33*(2), 218-236. doi:10.1111/bjdp.12075

Chronaki, G., Wigelsworth, M., Pell, M. D., & Kotz, S. A. (2018). The development of cross-cultural recognition of vocal emotion during childhood and adolescence. *Sci Rep, 8*(1), 8659. doi:10.1038/s41598-018-26889-1

Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and psychological measurement, 33*(1), 107-112.

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.

Cole, P. M., Tamang, B. L., & Shrestha, S. (2006). Cultural variations in the socialization of young children's anger and shame. *Child Development, 77*(5), 1237-1251. doi:DOI 10.1111/j.1467-8624.2006.00931.x

Conti-Ramsden, G., & Botting, N. (2008). Emotional health in adolescents with and without a history of specific language impairment (SLI). *Journal of Child Psychology and Psychiatry, 49*(5), 516-525. doi:10.1111/j.1469-7610.2007.01858.x

Cooper, R. J., Seib, J., Gagnon, L., Phillip, D., Schytz, H. W., Iversen, H. K., . . . Boas, D. A. (2012). A systematic comparison of motion artifact correction techniques for functional near-infrared spectroscopy. *Frontiers in Neuroscience, 6*. doi:10.3389/fnins.2012.00147

Cui, X., Bray, S., & Reiss, A. L. (2010). Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. *Neuroimage, 49*(4), 3039-3046. doi:10.1016/j.neuroimage.2009.11.050

Decasper, A. J., & Fifer, W. P. (1980). Of Human Bonding - Newborns Prefer Their Mothers Voices. *Science, 208*(4448), 1174-1176. doi:DOI 10.1126/science.7375928

Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science, 298*(5600), 2013-2015. doi:DOI 10.1126/science.1077066

Dehaene-Lambertz, G., Montavont, A., Jobert, A., Allirol, L., Dubois, J., Hertz-Pannier, L., & Dehaene, S. (2010). Language or music, mother or Mozart? Structural and environmental influences on infants' language networks. *Brain and Language, 114*(2), 53-65. doi:10.1016/j.bandl.2009.09.003

Delpy, D. T., & Cope, M. (1997). Quantification in tissue near–infrared spectroscopy. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 352*(1354), 649-659.

Demonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., . . . Frackowiak, R. (1992). The Anatomy of Phonological and Semantic Processing in Normal Subjects. *Brain, 115*, 1753-1768. doi:DOI 10.1093/brain/115.6.1753

Dykas, M. J., & Cassidy, J. (2011). Attachment and the Processing of Social Information Across the Life Span: Theory and Evidence. *Psychological Bulletin, 137*(1), 19-46. doi:10.1037/a0021367

Eisenberg, N., Cumberland, A., & Spinrad, T. L. (1998). Parental socialization of emotion. *Psychological Inquiry, 9*(4), 241-273. doi:DOI 10.1207/s15327965pli0904_1

Eisenberg, N., Vidmar, M., Spinrad, T. L., Eggum, N. D., Edwards, A., Gaertner, B., & Kupfer, A. (2010). Mothers' teaching strategies and children's effortful control: A longitudinal study. *Developmental Psychology, 46*(5), 1294.

Elsabbagh, M., & Johnson, M. H. (2007). Infancy and autism: progress, prospects, and challenges. *From Action to Cognition, 164*, 355-383. doi:10.1016/S0079-6123(07)64020-5

Elsabbagh, M., Mercure, E., Hudry, K., Chandler, S., Pasco, G., Charman, T., . . . Team, B. (2012). Infant Neural Sensitivity to Dynamic Eye Gaze Is Associated with Later Emerging Autism. *Current Biology, 22*(4), 338-342. doi:DOI 10.1016/j.cub.2011.12.056

Ethofer, T., Anders, S., Erb, M., Herbert, C., Wiethoff, S., Kissler, J., . . . Wildgruber, D. (2006). Cerebral pathways in processing of affective prosody: A dynamic causal modeling study. *Neuroimage, 30*(2), 580-587. doi:10.1016/j.neuroimage.2005.09.059

Ethofer, T., Anders, S., Wiethoff, S., Erb, M., Herbert, C., Saur, R., . . . Wildgruber, D. (2006). Effects of prosodic emotional intensity on activation of associative auditory cortex. *Neuroreport, 17*(3), 249-253. doi:10.1097/01.wnr.0000199466.32036.5d

Ethofer, T., Bretscher, J., Gschwind, M., Kreifelts, B., Wildgruber, D., & Vuilleumier, P. (2012). Emotional Voice Areas: Anatomic Location, Functional Properties, and Structural Connections Revealed by Combined fMRI/DTI. *Cereb Cortex, 22*(1), 191-200. doi:10.1093/cercor/bhr113

Everdell, N. L., Gibson, A. P., Tullis, I. D. C., Vaithianathan, T., Hebden, J. C., & Delpy, D. T. (2005). A frequency multiplexed near-infrared topography system for imaging functional activation in the brain. *Review of Scientific Instruments, 76*(9). doi:Artn 093705 10.1063/1.2038567

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods, 39*(2), 175-191.

Fecteau, S., Armony, J. L., Joanette, Y., & Belin, P. (2005). Sensitivity to voice in human prefrontal cortex. *Journal of Neurophysiology, 94*(3), 2251-2254. doi:10.1152/jn.00329.2005

Fecteau, S., Belin, P., Joanette, Y., & Armony, J. L. (2007). Amygdala responses to nonlinguistic emotional vocalizations. *Neuroimage, 36*(2), 480-487. doi:10.1016/j.neuroimage.2007.02.043

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in Early Communicative Development. *Monographs of the Society for Research in Child Development, 59*(5), R5-+.

Fernald, A. (1993). Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development, 64*(3), 657-674.

Field, T., Guy, L., & Umbel, V. (1985). Infants' responses to mothers' imitative behaviors. *Infant Mental Health Journal, 6*(1), 40-44.

Flom, R., & Bahrick, L. E. (2007). The development of infant discrimination of affect in multimodal and unimodal stimulation: The role of intersensory redundancy. *Developmental Psychology, 43*(1), 238-252. doi:10.1037/0012-1649.43.1.238

Fox, S. E., Wagner, J. B., Shrock, C. L., Tager-Flusberg, H., & Nelson, C. A. (2013). Neural processing of facial identity and emotion in infants at high-risk for autism spectrum disorders. *Front Hum Neurosci, 7*. doi:10.3389/Fnhum.2013.00089

Friederici, A. D. (2005). Neurophysiological markers of early language acquisition: from syllables to sentences. *Trends in Cognitive Sciences, 9*(10), 481-488. doi:DOI 10.1016/j.tics.2005.08.008

Frühholz, S., Ceravolo, L., & Grandjean, D. (2012). Specific Brain Networks during Explicit and Implicit Decoding of Emotional Prosody. *Cereb Cortex, 22*(5), 1107-1117. doi:10.1093/cercor/bhr184

Frühholz, S., & Grandjean, D. (2013). Processing of emotional vocalizations in bilateral inferior frontal cortex. *Neurosci Biobehav Rev, 37*(10), 2847-2855. doi:10.1016/j.neubiorev.2013.10.007

Fukui, Y., Ajichi, Y., & Okada, E. (2003). Monte Carlo prediction of near-infrared light propagation in realistic adult and neonatal head models. *Applied Optics, 42*(16), 2881-2887. doi:10.1364/Ao.42.002881

Gervain, J., Macagno, F., Cogoi, S., Pena, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proc Natl Acad Sci U S A, 105*(37), 14222-14227. doi:10.1073/pnas.0806530105

Gervain, J., Mehler, J., Werker, J. F., Nelson, C. A., Csibra, G., Lloyd-Fox, S., . . . Aslin, R. N. (2011). Near-infrared spectroscopy: A report from the McDonnell infant methodology consortium. *Developmental Cognitive Neuroscience, 1*(1), 22-46. doi:10.1016/j.dcn.2010.07.004

Goldstein, M. H., Schwade, J. A., & Bornstein, M. H. (2009). The Value of Vocalizing: Five-Month-Old Infants Associate Their Own Noncry Vocalizations With Responses From Caregivers. *Child Development, 80*(3), 636-644. doi:DOI 10.1111/j.1467-8624.2009.01287.x

Gómez, D. M., Berent, I., Benavides-Varela, S., Bion, R. A. H., Cattarossi, L., Nespor, M., & Mehler, J. (2014). Language universals at birth. *Proceedings of the National Academy of Sciences of the United States of America, 111*(16), 5837-5841. doi:10.1073/pnas.1318261111

Graham, A. M., Fisher, P. A., & Pfeifer, J. H. (2013). What Sleeping Babies Hear: A Functional MRI Study of Interparental Conflict and Infants' Emotion Processing. *Psychological Science, 24*(5), 782-789. doi:10.1177/0956797612458803

Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nature Neuroscience, 8*(2), 145-146. doi:10.1038/nn1392

Gredeback, G., Johnson, S., & von Hofsten, C. (2010). Eye tracking in infancy research. *Developmental Neuropsychology, 35*(1), 1-19. doi:10.1080/87565640903325758

Green, S., Caplan, B., & Baker, B. (2014). Maternal supportive and interfering control as predictors of adaptive and social development in children with and without developmental delays. *Journal of Intellectual Disability Research, 58*(8), 691-703.

Groome, L. J., Mooney, D. M., Holland, S. B., Smith, L. A., Atterbury, J. L., & Dykman, R. A. (1999). Behavioral state affects heart rate response to low-intensity sound in human fetuses. *Early Human Development, 54*(1), 39-54. doi:Doi 10.1016/S0378-3782(98)00083-8

Grossmann, T., & Johnson, M. H. (2007). The development of the social brain in human infancy. *European Journal of Neuroscience, 25*(4), 909-919. doi:10.1111/j.1460-9568.2007.05379.x

Grossmann, T., Lloyd-Fox, S., & Johnson, M. H. (2013). Brain responses reveal young infants' sensitivity to when a social partner follows their gaze. *Developmental Cognitive Neuroscience, 6*, 155-161. doi:DOI 10.1016/j.dcn.2013.09.004

Grossmann, T., Oberecker, R., Koch, S. P., & Friederici, A. D. (2010). The Developmental Origins of Voice Processing in the Human Brain. *Neuron, 65*(6), 852-858. doi:10.1016/j.neuron.2010.03.001

Grossmann, T., Parise, E., & Friederici, A. D. (2010). The detection of communicative signals directed at the self in infant prefrontal cortex. *Front Hum Neurosci, 4*, 201. doi:10.3389/fnhum.2010.00201

Grossmann, T., Striano, T., & Friederici, A. D. (2005). Infants' electric brain responses to emotional prosody. *Neuroreport, 16*(16), 1825-1828. doi:10.1097/01.wnr.0000185964.34336.b1

Grossmann, T., Vaish, A., Franz, J., Schroeder, R., Stoneking, M., & Friederici, A. D. (2013). Emotional Voice Processing: Investigating the Role of Genetic Variation in the Serotonin Transporter across Development. *Plos One, 8*(7). doi:10.1371/journal.pone.0068377

Guzell, J. R., & Vernon-Feagans, L. (2004). Parental perceived control over caregiving and its relationship to parent–infant interaction. *Child Development, 75*(1), 134-146.

Hargrove, P. M. (1997). Prosodic aspects of language impairment in children. *Topics in Language Disorders, 17*(4), 76-83.

Hart, S. L., Carrington, H. A., Tronick, E. Z., & Carroll, S. R. (2004). When infants lose exclusive maternal attention: Is it jealousy? *Infancy, 6*(1), 57-78.

Hebert-Myers, H., Guttentag, C. L., Swank, P. R., Smith, K. E., & Landry, S. H. (2006). The importance of language, social, and behavioral skills across early and later childhood as predictors of social competence with peers. *Applied Developmental Science, 10*(4), 174-187.

Heilman, K. M., Scholes, R., & Watson, R. T. (1975). Auditory Affective Agnosia - Disturbed Comprehension of Affective Speech. *Journal of Neurology Neurosurgery and Psychiatry, 38*(1), 69-72. doi:Doi 10.1136/Jnnp.38.1.69

Hepper, P. G., & Shahidullah, B. S. (1994a). Development of Fetal Hearing. *Archives of Disease in Childhood, 71*(2), F81-F87. doi:Doi 10.1136/Fn.71.2.F81

Hepper, P. G., & Shahidullah, B. S. (1994b). The development of fetal hearing. *Fetal and Maternal Medicine Review, 6*(3), 167-179.

Herrmann, M. J., Walter, A., Ehlis, A. C., & Fallgatter, A. J. (2006). Cerebral oxygenation changes in the prefrontal cortex: Effects of age and gender. *Neurobiology of Aging, 27*(6), 888-894. doi:10.1016/j.neurobiolaging.2005.04.013

Hocke, L. M., Oni, I. K., Duszynski, C. C., & Corrigan, A. V. (2018). Automated Processing of fNIRS Data—A Visual Guide to the Pitfalls and Consequences. *Algorithms, 11*(5), 67.

Hornik, R., Risenhoover, N., & Gunnar, M. (1987). The Effects of Maternal Positive, Neutral, and Negative Affective Communications on Infant Responses to New

Toys. *Child Development, 58*(4), 937-944. doi:DOI 10.1111/j.1467-8624.1987.tb01431.x

Huppert, T. J., Diamond, S. G., Franceschini, M. A., & Boas, D. A. (2009). HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain. *Applied Optics, 48*(10), D280-D298.

Hyvärinen, A., & Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural Networks, 13*(4-5), 411-430. doi:10.1016/S0893-6080(00)00026-5

Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology, 75*(4), 887-900. doi:Doi 10.1037//0022-3514.75.4.887

Izzetoglu, M., Devaraj, A., Bunce, S., & Onaral, B. (2005). Motion artifact cancellation in NIR spectroscopy using Wiener filtering. *Ieee Transactions on Biomedical Engineering, 52*(5), 934-938. doi:10.1109/Tbme.2005.845243

Jobsis, F. F. (1977). Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science, 198*(4323), 1264-1267.

Jobsis, F. F. (1999). Discovery of the near-infrared window into the body and the early development of near-infrared spectroscopy. *Journal of Biomedical Optics, 4*(4), 392-397.

Johnson, M. H., Grossmann, T., & Kadosh, K. C. (2009). Mapping functional brain development: Building a social brain through interactive specialization. *Developmental Psychology, 45*(1), 151.

Johnstone, T., van Reekum, C. M., Oakes, T. R., & Davidson, R. J. (2006). The voice of emotion: an FMRI study of neural responses to angry and happy vocal expressions. *Social Cognitive and Affective Neuroscience, 1*(3), 242-249. doi:10.1093/scan/nsl027

Jonsson, C., Clinton, D., Fahrman, M., Mazzaglia, G., Novak, S., & Sörhus, K. (2001). How do mothers signal shared feeling-states to their infants? An investigation of affect attunement and imitation during the first year of life. *Scandinavian Journal of Psychology, 42*(4), 377-381.

Kabdebon, C., Leroy, F., Simmonet, H., Perrot, M., Dubois, J., & Dehaene-Lambertz, G. (2014). Anatomical correlations of the international 10-20 sensor placement system in infants. *Neuroimage, 99*, 342-356. doi:10.1016/j.neuroimage.2014.05.046

Katura, T., Sato, H., Fuchino, Y., Yoshida, T., Atsumori, H., Kiguchi, M., . . . Tanaka, N. (2008). Extracting task-related activation components from optical topography measurement using independent components analysis. *Journal of Biomedical Optics, 13*(5), 054008.

Kennedy, J. J. (1970). The eta coefficient in complex ANOVA designs. *Educational and psychological measurement, 30*(4), 885-889.

Kisilevsky, B. S., Hains, S. M. J., Lee, K., Xie, X., Huang, H. F., Ye, H. H., . . . Wang, Z. P. (2003). Effects of experience on fetal voice recognition. *Psychological Science, 14*(3), 220-224. doi:10.1111/1467-9280.02435

Kitayama, S., Markus, H. R., & Matsumoto, H. (1995). Culture, self, and emotion: A cultural perspective on" self-conscious" emotions.

Kivijarvi, M., Voeten, M. J. M., Niemela, P., Raiha, H., Lertola, K., & Piha, J. (2001). Maternal sensitivity behavior and infant behavior in early interaction. *Infant Mental Health Journal, 22*(6), 627-640. doi:Doi 10.1002/Imhj.1023

Koessler, L., Maillard, L., Benhadid, A., Vignal, J. P., Felblinger, J., Vespignani, H., & Braun, M. (2009). Automated cortical projection of EEG sensors: Anatomical correlation via the international 10-10 system. *Neuroimage, 46*(1), 64-72. doi:10.1016/j.neuroimage.2009.02.006

Kolb, B., & Gibb, R. (2011). Brain plasticity and behaviour in the developing brain. *J Can Acad Child Adolesc Psychiatry, 20*(4), 265-276.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5*(11), 831-843. doi:10.1038/Nrn1533

Kuhl, P. K. (2007). Is speech learning 'gated' by the social brain? *Developmental Science, 10*(1), 110-120. doi:10.1111/j.1467-7687.2007.00572.x

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*(1493), 979-1000.

Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: implications for the" Critical Period". *Language Learning and Development, 1*(3-4), 237-264.

Kuhl, P. K., Ramirez, R. R., Bosseler, A., Lin, J. F. L., & Imada, T. (2014). Infants' brain responses to speech suggest Analysis by Synthesis. *Proceedings of the National Academy of Sciences of the United States of America, 111*(31), 11238-11245. doi:DOI 10.1073/pnas.1410963111

Kuhl, P. K., & Rivera-Gaxiola, M. (2008). Neural substrates of language acquisition. *Annual Review of Neuroscience, 31*, 511-534. doi:DOI 10.1146/annurev.neuro.30.051606.094321

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic Experience Alters Phonetic Perception in Infants by 6 Months of Age. *Science, 255*(5044), 606-608. doi:DOI 10.1126/science.1736364

Landry, S. H., Smith, K. E., MillerLoncar, C. L., & Swank, P. R. (1997). Predicting cognitive-language and social growth curves from early maternal behaviors in children at varying degrees of biological risk. *Developmental Psychology, 33*(6), 1040-1053. doi:Doi 10.1037/0012-1649.33.6.1040

Leigh, P., Nievar, M. A., & Nathans, L. (2011). Maternal Sensitivity and Language in Early Childhood: A Test of the Transactional Model. *Perceptual and Motor Skills, 113*(1), 281-299. doi:10.2466/10.17.21.28.PMS.113.4.281-299

Leppänen, J. M., & Nelson, C. A. (2009). Tuning the developing brain to social signals of emotions. *Nature Reviews Neuroscience, 10*(1), 37-47. doi:10.1038/nrn2554

Lloyd-Fox, S., Begus, K., Halliday, D., Pirazzoli, L., Blasi, A., Papademetriou, M., . . . Moore, S. (2017). Cortical specialisation to social stimuli from the first days to the second year of life: A rural Gambian cohort. *Developmental Cognitive Neuroscience*.

Lloyd-Fox, S., Blasi, A., & Elwell, C. E. (2010). Illuminating the developing brain: the past, present and future of functional near infrared spectroscopy. *Neurosci Biobehav Rev, 34*(3), 269-284.

Lloyd-Fox, S., Blasi, A., Elwell, C. E., Charman, T., Murphy, D., & Johnson, M. H. (2013). Reduced neural sensitivity to social stimuli in infants at risk for autism. *Proceedings of the Royal Society B-Biological Sciences, 280*(1758).

Lloyd-Fox, S., Blasi, A., Mercure, E., Elwell, C. E., & Johnson, M. H. (2012). The emergence of cerebral specialization for the human voice over the first months

of life. *Social Neuroscience, 7*(3), 317-330. doi:10.1080/17470919.2011.614696

Lloyd-Fox, S., Blasi, A., Volein, A., Everdell, N., Elwell, C. E., & Johnson, M. H. (2009). Social Perception in Infancy: A Near Infrared Spectroscopy Study. *Child Development, 80*(4), 986-999.

Lloyd-Fox, S., Papademetriou, M., Darboe, M. K., Everdell, N. L., Wegmuller, R., Prentice, A. M., . . . Elwell, C. E. (2014). Functional near infrared spectroscopy (fNIRS) to assess cognitive function in infants in rural Africa. *Scientific Reports, 4*.

Lloyd-Fox, S., Wu, R., Richards, J. E., Elwell, C. E., & Johnson, M. H. (2013). Cortical activation to action perception is associated with action production abilities in young infants. *Cereb Cortex, 25*(2), 289-297.

Lohaus, A., Keller, H., Ball, J., Elben, C., & Voelker, S. (2001). Maternal Sensitivity: Components and Relations to Warmth and Contingency. *Parenting-Science and Practice, 1*(4), 267-284. doi:10.1207/S15327922par0104_1

Lohaus, A., Keller, H., Ball, J., Voelker, S., & Elben, C. (2004). Maternal sensitivity in interactions with three-and 12-month-old infants: stability, structural composition, and developmental consequences. *Infant and Child Development, 13*(3), 235-252.

Määttä, S., Laakso, M. L., Tolvanen, T. A., Westerholm, J., & Aro, T. (2016). Continuity From Prelinguistic Communication to Later Language Ability: A Follow-Up Study From Infancy to Early School Age. *J Speech Lang Hear Res, 59*(6), 1357-1372. doi:10.1044/2016_JSLHR-L-15-0209

Maggio, V., Granana, N. E., Richaudeau, A., Torres, S., Giannotti, A., & Suburo, A. M. (2014). Behavior Problems in Children With Specific Language Impairment. *Journal of Child Neurology, 29*(2), 194-202. doi:Doi 10.1177/0883073813509886

Makeig, S., Bell, A. J., Jung, T. P., & Sejnowski, T. J. (1996). Independent component analysis of electroencephalographic data. *Advances in Neural Information Processing Systems 8, 8*, 145-151.

Malatesta, C. Z., Grigoryev, P., Lamb, C., Albin, M., & Culver, C. (1986). Emotion Socialization and Expressive Development in Preterm and Full-Term Infants. *Child Development, 57*(2), 316-330. doi:Doi 10.2307/1130587

Markham, J., White, B. R., Zeff, B. W., & Culver, J. P. (2009). Blind Identification of Evoked Human Brain Activity With Independent Component Analysis of Optical Data. *Human Brain Mapping, 30*(8), 2382-2392. doi:Doi 10.1002/Hbm.20678

Marshall, P. J., Bar-Haim, Y., & Fox, N. A. (2002). Development of the EEG from 5 months to 4 years of age. *Clin Neurophysiol, 113*(8), 1199-1208.

Mastropieri, D., & Turkewitz, G. (1999). Prenatal experience and neonatal responsiveness to vocal expressions of emotion. *Developmental Psychobiology, 35*(3), 204-214. doi:10.1002/(Sici)1098-2302(199911)35:3<204::Aid-Dev5>3.0.Co;2-V

Mathis, E. T. B., & Bierman, K. L. (2015). Dimensions of Parenting Associated with Child Prekindergarten Emotion Regulation and Attention Control in Low-income Families. *Social Development, 24*(3), 601-620. doi:10.1111/sode.12112

Maurage, P., Joassin, F., Philippot, P., & Campanella, S. (2007). A validated battery of vocal emotional expressions. *Neuropsychological Trends, 2*(1), 63-74.

Medvedev, A. V., Kainerstorfer, J., Borisov, S. V., Barbour, R. L., & VanMeter, J. (2008). Event-related fast optical signal in a rapid object recognition task: Improving detection by the independent component analysis. *Brain Research, 1236*, 145-158. doi:10.1016/j.brainres.2008.07.122

Mehler, J., Bertoncini, J., Barriere, M., & Jassikgerschenfeld, D. (1978). Infant Recognition of Mothers Voice. *Perception, 7*(5), 491-497. doi:Doi 10.1068/P070491

Michel, C., Stets, M., Parise, E., Reid, V. M., Striano, T., & Hoehl, S. (2015). Theta- and alpha-band EEG activity in response to eye gaze cues in early infancy. *Neuroimage, 118*, 576-583. doi:10.1016/j.neuroimage.2015.06.042

Minagawa-Kawai, Y., Cristia, A., Vendelin, I., Cabrol, D., & Dupoux, E. (2011). Assessing signal-driven mechanisms in neonates: brain responses to temporally and spectrally different sounds. *Frontiers in Psychology, 2*. doi:10.3389/Fpsyg.2011.00135

Minagawa-Kawai, Y., van der Lely, H., Ramus, F., Sato, Y., Mazuka, R., & Dupoux, E. (2011). Optical brain imaging reveals general auditory and language-specific processing in early infant development. *Cereb Cortex, 21*(2), 254-261. doi:10.1093/cercor/bhq082

Missana, M., Altvater-Mackensen, N., & Grossmann, T. (2017). Neural correlates of infants' sensitivity to vocal expressions of peers. *Dev Cogn Neurosci, 26*, 39-44. doi:10.1016/j.dcn.2017.04.003

Mitchell, R. L. C., Elliott, R., Barry, M., Cruttenden, A., & Woodruff, P. W. R. (2003). The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia, 41*(10), 1410-1421. doi:10.1016/S0028-3932(03)00017-4

Miyai, I., Tanabe, H. C., Sase, I., Eda, H., Oda, I., Konishi, I., . . . Kubota, K. (2001). Cortical mapping of gait in humans: A near-infrared spectroscopic topography study. *Neuroimage, 14*(5), 1186-1192. doi:DOI 10.1006/nimg.2001.0905

Molfese, D. L. (2000). Predicting dyslexia at 8 years of age using neonatal brain responses. *Brain and Language, 72*(3), 238-245. doi:DOI 10.1006/brln.2000.2287

Molfese, D. L., & Molfese, V. J. (1985). Electrophysiological Indexes of Auditory-Discrimination in Newborn-Infants - the Bases for Predicting Later Language-Development. *Infant Behavior & Development, 8*(2), 197-211. doi:Doi 10.1016/S0163-6383(85)80006-0

Molfese, D. L., & Molfese, V. J. (1997). Discrimination of language skills at five years of age using event-related potentials recorded at birth. *Developmental Neuropsychology, 13*(2), 135-156.

Moore, D. R. (2002). Auditory development and the role of experience. *Br Med Bull, 63*, 171-181.

Moriguchi, Y., & Hiraki, K. (2009). Neural origin of cognitive shifting in young children. *Proceedings of the National Academy of Sciences of the United States of America, 106*(14), 6017-6021. doi:10.1073/pnas.0809747106

Moser, S. J., Cutini, S., Weber, P., & Schroeter, M. L. (2009). Right prefrontal brain activation due to Stroop interference is altered in attention-deficit hyperactivity disorder - A functional near-infrared spectroscopy study. *Psychiatry Research-Neuroimaging, 173*(3), 190-195. doi:10.1016/j.pscychresns.2008.10.003

Mumme, D. L., Fernald, A., & Herrera, C. (1996). Infants' responses to facial and vocal emotional signals in a social referencing paradigm. *Child Development, 67*(6), 3219-3237. doi:Doi 10.2307/1131775

Murphy, K. R., Myors, B., & Wolach, A. (2014). *Statistical power analysis: A simple and general model for traditional and modern hypothesis tests*: Routledge.

Nakano, T., Watanabe, H., Homae, F., & Taga, G. (2009). Prefrontal cortical involvement in young infants' analysis of novelty. *Cereb Cortex, 19*(2), 455-463. doi:10.1093/cercor/bhn096

Nowicki, S., & Duke, M. P. (1992). The association of children's nonverbal decoding abilities with their popularity, locus of control, and academic achievement. *The Journal of genetic psychology, 153*(4), 385-393.

Ockleford, E. M., Vince, M. A., Layton, C., & Reader, M. R. (1988). Responses of Neonates to Parents and Others Voices. *Early Human Development, 18*(1), 27-36. doi:10.1016/0378-3782(88)90040-0

Olson, S. L., Bayles, K., & Bates, J. E. (1986). Mother-child interaction and children's speech progress: A longitudinal study of the first two years. *Merrill-Palmer Quarterly (1982-)*, 1-20.

Ortiz-Mantilla, S., Hamalainen, J. A., & Benasich, A. A. (2012). Time course of ERP generators to syllables in infants: A source localization study using age-appropriate brain templates. *Neuroimage, 59*(4), 3275-3287. doi:10.1016/j.neuroimage.2011.11.048

Palfrey, J. S., Singer, J. D., Walker, D. K., & Butler, J. A. (1987). Early identification of children's special needs: a study in five metropolitan communities. *The Journal of pediatrics, 111*(5), 651-659.

Peeters, G., & Czapinski, J. (1990). Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. *European review of social psychology, 1*(1), 33-60.

Peirce, J. W. (2007). PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1-2), 8-13. doi:10.1016/j.jneumeth.2006.11.017

Pell, M. D., Rothermich, K., Liu, P., Paulmann, S., Sethi, S., & Rigoulot, S. (2015). Preferential decoding of emotion from human non-linguistic vocalizations versus speech prosody. *Biological Psychology, 111*, 14-25. doi:10.1016/j.biopsycho.2015.08.008

Pena, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., & Mehler, J. (2003). Sounds and silence: An optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences of the United States of America, 100*(20), 11702-11705. doi:10.1073/pnas.1934290100

Pinheiro, A. P., Barros, C., Vasconcelos, M., Obermeier, C., & Kotz, S. A. (2017). Is laughter a better vocal change detector than a growl? *Cortex, 92*, 233-248. doi:10.1016/j.cortex.2017.03.018

Pinti, P., Tachtsidis, I., Hamilton, A., Hirsch, J., Aichelburg, C., Gilbert, S., & Burgess, P. W. (2018). The present and future use of functional near-infrared spectroscopy (fNIRS) for cognitive neuroscience. *Annals of the New York Academy of Sciences*.

Pinto-Martin, J. A., Dunkle, M., Earls, M., Fliedner, D., & Landes, C. (2005). Developmental stages of developmental screening: Steps to implementation of a successful program. *American Journal of Public Health, 95*(11), 1928-1932. doi:10.2105/Ajph.2004.052167

Piper, S. K., Krueger, A., Koch, S. P., Mehnert, J., Habermehl, C., Steinbrink, J., . . . Schmitz, C. H. (2014). A wearable multi-channel fNIRS system for brain imaging in freely moving subjects. *Neuroimage, 85*, 64-71. doi:10.1016/j.neuroimage.2013.06.062

Posner, M. I., Petersen, S. E., Fox, P. T., & Raichle, M. E. (1988). Localization of Cognitive Operations in the Human-Brain. *Science, 240*(4859), 1627-1631. doi:DOI 10.1126/science.3289116

Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2005). Brain potentials to native and non-native speech contrasts in 7-and 11-month-old American infants. *Developmental Science, 8*(2), 162-172. doi:DOI 10.1111/j.1467-7687.2005.00403.x

Robertson, F. C., Douglas, T. S., & Meintjes, E. M. (2010). Motion Artifact Removal for Functional Near Infrared Spectroscopy: A Comparison of Methods. *Ieee Transactions on Biomedical Engineering, 57*(6), 1377-1387. doi:10.1109/Tbme.2009.2038667

Rogier, O., Roux, S., Belin, P., Bonnet-Brilhault, F., & Bruneau, N. (2010). An electrophysiological correlate of voice processing in 4-to 5-year-old children. *International Journal of Psychophysiology, 75*(1), 44-47. doi:10.1016/j.ijpsycho.2009.10.013

Rosenblum, K. L., Dayton, C. J., & Muzik, M. (2009). Infant social and emotional development: The emergence of self in a relational context. *In C. H. Zeanah, Ed. Handbook of infant mental health, NY: Guilford Press, 3rd ed.*, 80-103.

Ross, E. D., & Monnot, M. (2011). Affective prosody: What do comprehension errors tell us about hemispheric lateralization of emotions, sex and aging effects, and the role of cognitive appraisal. *Neuropsychologia, 49*(5), 866-877. doi:10.1016/j.neuropsychologia.2010.12.024

Ruddy, M. G., & Bornstein, M. H. (1982). Cognitive correlates of infant attention and maternal stimulation over the first year of life. *Child Development*, 183-188.

Saint-Georges, C., Chetouani, M., Cassel, R., Apicella, F., Mahdhaoui, A., Muratori, F., . . . Cohen, D. (2013). Motherese in Interaction: At the Cross-Road of Emotion and Cognition? (A Systematic Review). *Plos One, 8*(10). doi:10.1371/journal.pone.0078103

Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology, 32*(1), 76-92. doi:10.1177/0022022101032001009

Schiessl, I., Wang, W., & McLoughlin, N. (2008). Independent components of the haemodynamic response in intrinsic optical imaging. *Neuroimage, 39*(2), 634-646. doi:10.1016/j.neuroimage.2007.09.022

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences, 10*(1), 24-30. doi:10.1016/j.tics.2005.11.009

Scholkmann, F., Spichtig, S., Muehlemann, T., & Wolf, M. (2010). How to detect and reduce movement artifacts in near-infrared imaging using moving standard deviation and spline interpolation. *Physiological Measurement, 31*(5), 649-662. doi:10.1088/0967-3334/31/5/004

Schupp, H. T., Ohman, A., Junghofer, M., Weike, A. I., Stockburger, J., & Hamm, A. O. (2004). The facilitated processing of threatening faces: An ERP analysis. *Emotion, 4*(2), 189-200. doi:10.1037/1528-3542.4.2.189

Shultz, S., & Vouloumanos, A. (2010). Three-month-olds prefer speech to other naturally occurring signals. *Language Learning and Development, 6*(4), 241-257.

Silva, P. A., Mcgee, R., & Williams, S. M. (1983). Developmental Language Delay from 3 to 7 Years and Its Significance for Low Intelligence and Reading

Difficulties at Age 7. *Developmental Medicine and Child Neurology, 25*(6), 783-793.
Singh, L., Morgan, J. L., & Best, C. T. (2002). Infants' Listening Preferences: Baby Talk or Happy Talk? *Infancy, 3*(3), 365-394. doi:10.1207/S15327078in0303_5
Sorce, J. F., Emde, R. N., Campos, J., & Klinnert, M. D. (1985). Maternal Emotional Signaling - Its Effect on the Visual-Cliff Behavior of 1-Year-Olds. *Developmental Psychology, 21*(1), 195-200. doi:Doi 10.1037//0012-1649.21.1.195
Stifter, C. A., & Fox, N. A. (1987). Preschool children's ability to identify and label emotions. *Journal of Nonverbal Behavior, 11*(1), 43-54.
Striano, T., & Rochat, P. (2000). Emergence of Selective Social Referencing in Infancy. *Infancy, 1*(2), 253-264. doi:10.1207/S15327078in0102_7
Suto, T., Fukuda, M., Ito, M., Uehara, T., & Mikuni, M. (2004). Multichannel near-infrared spectroscopy in depression and schizophrenia: Cognitive brain activation study. *Biological Psychiatry, 55*(5), 501-511. doi:10.1016/j.biopsych.2003.09.008
Swingler, M. M., Perry, N. B., Calkins, S. D., & Bell, M. A. (2017). Maternal Behavior Predicts Infant Neurophysiological and Behavioral Attention Processes in the First Year. *Developmental Psychology, 53*(1), 13-27. doi:10.1037/dev0000187
Sylvestre, A., & Merette, C. (2010). Language delay in severely neglected children: A cumulative or specific effect of risk factors? *Child Abuse & Neglect, 34*(6), 414-428. doi:DOI 10.1016/j.chiabu.2009.10.003
Taga, G., & Asakawa, K. (2007). Selectivity and localization of cortical response to auditory and visual stimulation in awake infants aged 2 to 4 months. *Neuroimage, 36*(4), 1246-1252. doi:10.1016/j.neuroimage.2007.04.037
Tak, S., & Ye, J. C. (2014). Statistical analysis of fNIRS data: a comprehensive review. *Neuroimage, 85 Pt 1*, 72-91. doi:10.1016/j.neuroimage.2013.06.016
Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children's achievement of language milestones. *Child Development, 72*(3), 748-767.
Teinonen, T., Fellman, V., Naatanen, R., Alku, P., & Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *Bmc Neuroscience, 10*. doi:Artn 21 10.1186/1471-2202-10-21
Thompson, R. A. (1991). Emotional Regulation and Emotional Development. *Educational Psychology Review, 3*(4), 269-307. doi:Doi 10.1007/Bf01319934
Toppelberg, C. O., & Shapiro, T. (2000). Language disorders: A 10-year research update review. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*(2), 143-152. doi:Doi 10.1097/00004583-200002000-00011
Trevarthen, C. (2017). The function of emotions in early infant communication and development *New perspectives in early communicative development* (pp. 48-81): Routledge.
Trevithick, L., McAllister-Williams, R. H., Blamire, A., Branton, T., Clark, R., Downey, D., . . . Anderson, I. M. (2015). Study protocol for the randomised controlled trial: Ketamine augmentation of ECT to improve outcomes in depression (Ketamine-ECT study). *BMC Psychiatry, 15*, 257. doi:10.1186/s12888-015-0641-4
Tronick, E. (1989). Emotions and Emotional Communication in Infants. *American Psychologist, 44*(2), 112-119. doi:Doi 10.1037//0003-066x.44.2.112
Tronick, E., Als, H., Adamson, L. B., Wise, S., & Brazelton, T. B. (1979). The infant's response to entrapment between contradictory messages in face-to-face

interaction. *Journal of the American Academy of Child psychiatry, 17*(1), 1-13.

Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development, 75*(4), 1067-1084. doi:DOI 10.1111/j.1467-8624.2004.00726.x

Tsuzuki, D., & Dan, I. (2014). Spatial registration for functional near-infrared spectroscopy: From channel position on the scalp to cortical location in individual and group analyses. *Neuroimage, 85*, 92-103. doi:10.1016/j.neuroimage.2013.07.025

Urakawa, S., Takamoto, K., Ishikawa, A., Ono, T., & Nishijo, H. (2015). Selective Medial Prefrontal Cortex Responses During Live Mutual Gaze Interactions in Human Infants: An fNIRS Study. *Brain Topogr, 28*(5), 691-701. doi:10.1007/s10548-014-0414-2

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin, 134*(3), 383-403. doi:10.1037/0033-2909.134.3.383

Vaish, A., & Striano, T. (2004). Is visual reference necessary? Contributions of facial versus vocal cues in 12-months-olds' social referencing behavior. *Developmental Science, 7*(3), 261-269. doi:10.1111/j.1467-7687.2004.00344.x

Vanlancker, D., Kreiman, J., & Cummings, J. (1989). Voice Perception Deficits - Neuroanatomical Correlates of Phonagnosia. *Journal of Clinical and Experimental Neuropsychology, 11*(5), 665-674. doi:Doi 10.1080/01688638908400923

Vigario, R., Jousmaki, V., Hamalainen, M., Hari, R., & Oja, E. (1998). Independent component analysis for identification of artifacts in magnetoencephalographic recordings. *Advances in Neural Information Processing Systems 10, 10*, 229-235.

Villringer, A., & Chance, B. (1997). Non-invasive optical spectroscopy and imaging of human brain function. *Trends in Neurosciences, 20*(10), 435-442. doi:10.1016/S0166-2236(97)01132-6

Virtanen, J., Noponen, T., Kotilahti, K., Virtanen, J., & Ilmoniemi, R. J. (2011). Accelerometer-based method for correcting signal baseline changes caused by motion artifacts in medical near-infrared spectroscopy. *Journal of Biomedical Optics, 16*(8). doi:10.1117/1.3606576

Vouloumanos, A., Hauser, M. D., Werker, J. F., & Martin, A. (2010). The Tuning of Human Neonates' Preference for Speech. *Child Development, 81*(2), 517-527.

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. doi:DOI 10.1111/j.1467-7687.2007.00549.x

Walker-Andrews, A. S. (1997). Infants' perception of expressive behaviors: differentiation of multimodal information. *Psychological Bulletin, 121*(3), 437-456.

Walker-Andrews, A. S., & Grolnick, W. (1983). Discrimination of Vocal Expressions by Young Infants. *Infant Behavior & Development, 6*(4), 491-498. doi:Doi 10.1016/S0163-6383(83)90331-4

Walker-Andrews, A. S., Krogh-Jespersen, S., Mayhew, E. M. Y., & Coffield, C. N. (2011). Young Infants' Generalization of Emotional Expressions: Effects of Familiarity. *Emotion, 11*(4), 842-851. doi:10.1037/a0024435

Wan, M. W. (2015). Manchester Assessment of Caregiver-Infant Interaction. Coding Manual, Version 2. Manchester: Unpublished manual.

Wan, M. W., Brooks, A., Green, J., Abel, K. M., & Elmadih, A. (2017). Psychometrics and validation of a brief rating measure of parent-infant interaction: Manchester assessment of caregiver–infant interaction. *International Journal of Behavioral Development, 41*(4), 542-549.

Wan, M. W., Downey, D., Strachan, H., Elliott, R., Williams, S. R., & Abel, K. M. (2014). The Neural Basis of Maternal Bonding. *Plos One, 9*(3).

Wan, M. W., Green, J., Elsabbagh, M., Johnson, M., Charman, T., Plummer, F., & Team, B. (2012). Parent-infant interaction in infant siblings at risk of autism. *Research in Developmental Disabilities, 33*(3), 924-932. doi:DOI 10.1016/j.ridd.2011.12.011

Wan, M. W., Green, J., Elsabbagh, M., Johnson, M., Charman, T., Plummer, F., & Team, B. (2013). Quality of interaction between at-risk infants and caregiver at 12-15months is associated with 3-year autism outcome. *Journal of Child Psychology and Psychiatry, 54*(7), 763-771. doi:Doi 10.1111/Jcpp.12032

Wan, M. W., Green, J., & Scott, J. (2018). A systematic review of parent–infant interaction in infants at risk of autism. *Autism*, 1362361318777484.

Watanabe, H., Homae, F., Nakano, T., Tsuzuki, D., Enkhtur, L., Nemoto, K., . . . Taga, G. (2013). Effect of auditory input on activations in infant diverse cortical regions during audiovisual processing. *Human Brain Mapping, 34*(3), 543-565. doi:10.1002/hbm.21453

Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology/Revue canadienne de psychologie, 43*(2), 230.

Werker, J. F., & Tees, R. C. (2002). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development, 25*(1), 121-133. doi:10.1016/S0163-6383(02)00093-0

Wilcox, T., Bortfeld, H., Woods, R., Wruck, E., & Boas, D. A. (2005). Using near-infrared spectroscopy to assess neural activation during object processing in infants. *Journal of Biomedical Optics, 10*(1). doi:Doi 10.1117/1.1852551

Witteman, J., Van Heuven, V. J. P., & Schiller, N. O. (2012). Hearing feelings: A quantitative meta-analysis on the neuroimaging literature of emotional prosody perception. *Neuropsychologia, 50*(12), 2752-2763. doi:10.1016/j.neuropsychologia.2012.07.026

World Health Organization (WHO). (2003). Child growth standards.

Wykes, T., Haro, J. M., Belli, S. R., Obradors-Tarrago, C., Arango, C., Ayuso-Mateos, J. L., . . . consortium, R. (2015). Mental health research priorities for Europe. *Lancet Psychiatry, 2*(11), 1036-1042. doi:10.1016/S2215-0366(15)00332-6

Yuan, Z. (2013). Spatiotemporal and time-frequency analysis of functional near infrared spectroscopy brain signals using independent component analysis. *Journal of Biomedical Optics, 18*(10). doi:Doi 10.1117/1.Jbo.18.10.106011

Zhang, D. D., Liu, Y. Z., Hou, X. L., Sun, G. Y., Cheng, Y. W., & Luo, Y. J. (2014). Discrimination of fearful and angry emotional voices in sleeping human neonates: a study of the mismatch brain responses. *Frontiers in Behavioral Neuroscience, 8*. doi:10.3389/Fnbeh.2014.00422

Zhang, D. D., Zhou, Y., Hou, X. L., Cui, Y., & Zhou, C. L. (2017). Discrimination of emotional prosodies in human neonates: A pilot fNIRS study. *Neuroscience Letters, 658*, 62-66. doi:10.1016/j.neulet.2017.08.047

Zhang, D. D., Zhou, Y., & Yuan, J. J. (2018). Speech Prosodies of Different Emotional Categories Activate Different Brain Regions in Adult Cortex: an fNIRS Study. *Scientific Reports, 8*. doi:10.1038/s41598-017-18683-2

Zhang, H., Zhang, Y. J., Lu, C. M., Ma, S. Y., Zang, Y. F., & Zhu, C. Z. (2010). Functional connectivity as revealed by independent component analysis of resting-state fNIRS measurements. *Neuroimage, 51*(3), 1150-1161. doi:10.1016/j.neuroimage.2010.02.080

Zhao, C., Valentini, E., & Hu, L. (2015). Functional features of crossmodal mismatch responses. *Exp Brain Res, 233*(2), 617-629. doi:10.1007/s00221-014-4141-4

Zweifel, C., Castellani, G., Czosnyka, M., Helmy, A., Manktelow, A., Carrera, E., . . . Smielewski, P. (2010). Noninvasive Monitoring of Cerebrovascular Reactivity with Near Infrared Spectroscopy in Head-Injured Patients. *Journal of Neurotrauma, 27*(11), 1951-1958. doi:10.1089/neu.2010.1388

Appendix 1: Participant information sheet

Study Title: Novel biomarkers in infants: developing optical imaging solutions for the measurement of early vocal brain development

PARTICIPANT INFORMATION SHEET

A Study of the mothering role in baby's brain development

If you wish, one of our researchers will go through this information sheet with you and answer any questions you have. You are also free to talk to others about the study if you wish. Please ask us if there is anything that is not clear or if you would like more information. It is important that you take time to decide whether or not you wish to take part.

What is the purpose of the study?

A baby's vocal development (using the voice and reacting to voices) is important for speech and language development. An abnormally developed vocal ability (problems with speaking and/ or understanding) is related to behavioural problems and an increased risk of mental disorders. Mothers make up the majority of primary caregivers across cultures and undertake most caring responsibilities. Mothers' different caregiving styles will have different effects on children's language development outcomes.

This study will focus on developing a method to study and monitor brain patterns related to the care a mother gives. We want to examine the brain processing of vocal sounds and speech in babies from the age of 6 - 18 months old within the same mother-baby pairs, and to understand the link between the mother's care and a baby's vocal development.

Fig 1

Fig 2

Why have I been chosen?

We are looking for new healthy mothers and their babies to take part in this study.

You are being invited because:

(i) You are aged 18 – 40, healthy and do not have any mental or physical illness;
(ii) You are a white fluent English speaker;
(iii) You are pregnant or the biological mother of a healthy full-term baby and he/ she is under/ about 6 months old;
(iv) You are taking care of your baby.

Do I have to take part?

It is up to you whether to take part or not. If you do decide to take part, you are free to withdraw at any time during the study without giving a reason. If you do not wish to take part or decide to withdraw, this will not affect the standard of care you receive.

What will happen to me if I take part?

There are four parts in the present study: a mother's questionnaire session, a mother-baby play session, three Baby Vocal Development Testing Sessions, and mother's feedback sessions.

Part I - Questionnaire

Should you wish to take part, we shall be asking you to complete a series of short questionnaires about your age, highest qualification, occupation, number of children, as well as your own childhood experiences, and how you are feeling in your mood.

Scores on these questionnaires will help us to select mothers and babies to take part in the second part of the study for the following reason: we would like our sample of mother-babies to have as broad a range of scores as possible on these measures. If we do this, we can be more confident that the group, who undertake the imaging tests (Baby Vocal Development Testing), on 3 occasions, with their babies, will represent the wide variation of language development seen in the population. If you agree, we would like to see your medical records so that we can gather information about your emotional wellbeing from questionnaires/discussions you may have completed with your Health Visitor. Only relevant information will be gathered and as with all study data will be kept securely in line with the Data Protection Act (1998).

We shall only know about which mothers will be selected when we begin to sample mothers. We shall relate an individual's scores on the initial quick and easy questionnaires to scores in the larger sample. In this way, we shall select mothers to go forward to the second stage.

Part II - Mother-Baby Play

The second part is a videotaped play session between you and your baby. All you need to do in this play session is play with your baby for 6 minutes like you always do

at home (Fig 1). You can choose to use toys or not. This play session will last 15 to 30 minutes. We will also record your voice for your baby to listen to in Part III. You can choose to do this phase at home or come to Wellcome Trust Clinical Research Facility (CRF). If you choose to come to Wellcome Trust CRF, we will pay travel costs.

Part III - Baby Vocal Brain Development Testing

The third part involves vocal brain development testing sessions of your baby, and identical procedures will take place on three different days when your baby is aged from 6 to 18 months. In these sessions we will assess your baby's speech perception by using non-invasive optical imaging method (functional Near Infrared Spectroscopy, fNIRS; details of this technique will be in another information sheet given on request, Fig 2).

In a measurement session, your baby will sit on your lap, listen to some sounds (e.g. human voices, non-human sounds, your speech and an unknown woman's speech), see some images (images of different types of transport), and watch some video clips (adult social interactions), while his/ her brain responses will be quietly collected. In addition, your baby's performance during the measurement will be videotaped for the sake of analysis. A measurement session will take 20 to 40 minutes. You can choose to do these sessions at your home or come to Wellcome Trust Clinical Research Facility (WTCRF). If you choose to come to Wellcome Trust CRF, we will pay travel costs.

Part IV – Baby Language Behaviour Measurement

After vocal brain development testing, we will assess your baby's language behaviour performance with the language subset of Bayley Scales of Infant and Toddler Development, Third Edition (Bayley-III). This assessment involves non-invasive observations and tasks with baby, and will take 15 to 30 minutes. If you agree, we will also access your baby's medical records in order to gather any developmental

information that your health visitor may have gathered. Again, we will only look at and gather information relevant to the study and all study data will be stored securely in-line with the Data Protection Act (1998).

Part V - Mother's feedback

After each measurement session, you will answer a questionnaire on how you feel about the measurement session and how you think your baby feels about the measurement session. This questionnaire will take you about 5-10 minutes.

After the last session, we would like you to take part in a short, voice recorded interview (10-15 minutes), about how you feel about the whole study, including the equipment settings, length of measurement and recruitment, and what you think your baby felt about the measurement sessions.

What are the possible risks?

fNIRS is a non-invasive tool to record brain responses, and it has been safely applied in many baby studies. We will follow the steps to carry out the study and minimise

risks to you and your baby. Your baby will sit on your lap during the vocal development testing session, so that you will be with your baby all the time. We will stop the session (both in Part II-Mother-Baby Play and Part III-Baby Vocal Development Testing) at any time your baby feels uncomfortable, or you think your baby wants to stop, so you can comfort your baby.

What are the possible benefits?

By taking part in this unique study, you will help us understand how caregiving activities affect a baby's brain development. We value your time and participation and would like to offer you £5 voucher for the questionnaire session and £45 worth of voucher for completing the remaining sessions plus up to £15 travel costs for each visit.

Unfortunately, we are not able to provide individual information on brain activation because scans will be averaged across all participants for research purposes.

Will my taking part be kept confidential?

We value your participation in the study and all information will be treated as strictly confidential. The only exception to this is if the researcher becomes concerned about risk of serious danger to yourself or others. If this happens, we will inform the clinical staff involved in your care. If you feel distressed, you can speak to researcher and the person in charge of your care (GP/midwife/health visitor). Further support is available from the card the researcher give you at the visit, for example, PANDAS (http://www.pandasfoundation.org.uk/, 08432 898 401), Mind Helpline (http://www.mind.org.uk/, 0300 123 3393), Samaritans (http://www.samaritans.org/, 08457 909090), etc.

Any information and data we collect from you will be kept in accordance with the Data Protection Act 1998. Individuals from the University of Manchester, NHS Trust or regulatory authorities may need to look at the study information to make sure the project is being carried out properly. With your permission, this will include identifiable data (they will be able to tell who it belongs to). Your interview recording will be transcribed by the researcher, the recordings will be destroyed after they have been used, and your personal details will not be disclosed when disseminating the study results. Anyone that looks at the data will have a duty of confidentiality to you and your child as research participants.

What will happen to my data?

Some of the information collected for this study could be used to support future research. With your permission, we would like to use the data or share the data with other researchers so it can be used in other studies. Your name and any other identifiable information would be removed before the data was

used or shared so no one would be able to tell who it belonged to. Your anonymised data will be held securely by the research team at The University of Manchester for 5 years after the last publication of the study or for 10 years, whichever is longer, after which point it will be destroyed. The video and audio data will be transcribed and rated

by the researchers within the research team, these data will be destroyed after they have been used, personal details will not be disclosed. Personal information (e.g. name, date of birth) will be destroyed as soon as it is no longer needed.

What if something goes wrong?

If you have a concern about any aspect of this study, please speak to the lead researcher, Chen Zhao (0161 306 7916), who will do her best to answer your questions. If you wish to make a complaint regarding the study, please contact a University Research Practice and Governance Coordinator on (0161 275 7583) or (0161 275 8093) or by email to research.complaints@manchester.ac.uk. In the event that something does go wrong and you or your child is harmed during the research you may have grounds for a legal action for compensation against the University of Manchester or NHS Trusts but you may have to pay your legal costs. The normal National Health Service complaints mechanisms will still be available to you.

Who is funding and conducting the research?

The study is funded by China Scholarship Council (CSC) in collaboration with the Centre for Women's Mental Health, Institute of Brain Behaviour & Mental Health, The University of Manchester.

The research is organised by The University of Manchester. It will be conducted by a researcher (Chen Zhao) as part of her PhD in Medicine, and supervised by Prof Kathryn Abel (Professor of Psychological Medicine).

Who has reviewed the study?

This research project has been reviewed by the National Research Ethics Service [North West - Greater Manchester Central Research Ethics Committee, ref. 15/NW/0684], and your NHS Research & Development Department [Central Manchester University Hospitals NHS Foundation Trust, ref. R04137].

What do I do now?

If you would like to take part, call/email Chen Zhao (0161 306 7916, chen.zhao-8@postgrad.manchester.ac.uk)

Thank you very much for considering taking part in our research.

Please contact us if you require further information or advice on how to take part:

Miss Chen Zhao and Professor Kathryn Abel Institute of Brain, Behaviour and Mental Health

Jean McFarlane Building, University of Manchester, Oxford Road Manchester, M13 9PL

Phone: 0161 306 7916/ 0771 765 0306 Email chen.zhao-8@postgrad.manchester.ac.uk

Appendix 2: Participant information sheet (fNIRS)

Study Title: Novel Biomarkers in Infants: Developing Optical Imaging Solutions for Measurement of Early Vocal Brain Development.

A Study of the Mothering Role in Infant Brain Development

PARTICIPANT INFORMATION SHEET

Functional Near Infrared Spectroscopy (fNIRS)


Fig 1

What is fNIRS?

Functional Near Infrared Spectroscopy (fNIRS) is a non-invasive tool used to indirectly observe brain responses.

How does it work?

When we sense stimulation (in this case, a voice or speech sound, but it could be a smell or seeing something), our brain gets the information, and neurons (nerve cells) are activated. In certain brain regions, the blood flow increases, along with an increased oxygenation requirement to fuel the brain activity. This changed blood flow (oxygenated and deoxygenated haemoglobin) in the red blood cells represents the activation of the brain region, and can be detected and measured by fNIRS, which calculates the oxy-haemoglobin and deoxy-haemoglobin absorbed infrared light. Using the fNIRS, we can give a stimulus (a voice, for example) and then watch as the baby reacts, and see the pathways that the brain responses take. Using this method can help us to understand more of the processing in the brain responding to language.

There are ‘detectors’ and ‘sources’ (See Fig 1) on the fNIRS headgear. The near infrared light, which travels through skin, skull and brain tissue, migrates between ‘source’ and ‘detector’. By calculating the absorbed light by oxy-haemoglobin and deoxy-haemoglobin, we can indirectly measure the brain responses of your baby during testing.

What is the purpose of using fNIRS?

fNIRS can provide objective and reliable results on a baby's vocal brain development. Your baby's vocal brain development will be assessed three times using fNIRS between 6-18 months old.

Why fNIRS?

fNIRS is a widely used tool in baby studies because of its advantages, listed below:

(a) it is silent, which provides a non-intrusive environment to test a baby's reaction to voice;

(b) it can give us a better understanding of the brain area that is related to the vocal development;

(c) fNIRS is not as sensitive to movement as other equipment, so that we can observe a baby's reactions while he/ she is awake;

(d) fNIRS doesn't involve magnetic fields, and doesn't use X-Rays, both of which are not friendly to babies;

(e) the baby will sit in the mother's lap during assessment, which makes it a friendly equipment in studying baby's brain development.

Will it do any harm to my baby?

There is no risk in taking part in the present study. fNIRS is non-invasive tool to measure baby's brain responses, and it will not do any harm to your baby. It has been tested on many adult and baby groups in previous studies, and has earned a very good reputation.

What will happen during fNIRS testing?

During the measurement, your baby will wear the headgear and sit in your lap in front of a screen. Sounds will be played (human voice sounds like coughing yawning, laughing and crying; other sounds, like running water, toy rattles, bells, etc; the mother's speech and an unknown female's speech), images (cars and helicopters), and videos (adults in normal social interactions). The length of the test will be 20 minutes.

If your baby becomes fussy or tired, we will stop the measurement at any time, and you can comfort your baby. The measurement in all three phases will be the same (same procedure and stimulation display).

Will my baby's fNIRS data be kept confidential?

Your baby's fNIRS data will be treated as strictly confidential. When you participate in the study, you will be given a Study ID, your baby's fNIRS data will only be identified

by the Study ID when we do the analysis work, and when we publish the data, so that your baby's information (name, address, date of birth etc.) will not be identified.

Thank you very much for considering taking part in our research.

Please contact us if you require further information or advice on how to take part:

Miss Chen Zhao and Professor Kathryn Abel

Institute of Brain, Behaviour and Mental Health

Jean McFarlane Building, University of Manchester, Oxford Road Manchester M13 9PL

Tel 0161 306 67916/ 0771 765 0306 Email chen.zhao-8@postgrad.manchester.ac.uk

Appendix 3: Consent form (General)

CONSENT FORM (General)

Study Title: Novel biomarkers in infants: developing optical imaging solutions for the measurement of early vocal brain development

Name of Researcher: Miss Chen Zhao

REC Ref: 15/NW/0684

Please initial each box and sign your name to show you agree to the items below:

1) I confirm that I have read and understand the Participant Information Sheet [Version 3, dated 26.02.2016] for the above study and have had the opportunity to ask questions and had time to think and talk about it.

2) I understand that my and my baby's participation is voluntary and that I am free to withdraw at any time without giving any reason and without it affecting our future care.

3) I understand that video and audio recordings may be taken as part of the study in line with the participant information sheet [Version 3, dated 26.02.2016], and I give consent for this.

4) I understand that data collected during the study, may be looked at by individuals from the University of Manchester, from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my data.

5) I agree that anonymised quotes from interviews can be used in the reporting of this research and future research. I understand that my personal details will not be used and it will not be possible to identify me from any published information.

6) I agree for my baby and myself to take part in the above study.

7) I agree the researchers' accessing to mine and my baby's medical records that are related to the purpose of the study.

8) I agree to my GP being informed about my involvement in the study.

9) I would like to receive a summary of the study results.

10) I understand that the information collected about me will be used to support other research in the future, and may be shared with other researchers.

Please write your name in capital letters below, sign and date.

Name of participant	**Date**	**Participant's signature**
_____	_____/____/____	_____

To be completed by the researcher

I confirm that I have explained the study and supplied the participant with an information sheet which in my judgement is suited to the understanding of the participant.

_____	_____/____/____	_____
Name of researcher	Date	Researcher's signature

Appendix 4: Study advert

Take part in a research study to explore your baby's vocal brain development

REC ref: 15/NW/0684

We are looking for healthy mothers and babies

*

This involves a mother-baby play session and baby vocal development measurement sessions which will help us to understand how mother's caregiving affects baby's vocal brain development

All participants will be compensated for their time

To find out more, please contact: Miss Chen Zhao:

chen.zhao-8@postgrad.manchester.ac.uk

0161 306 7916/ 0771 765 0306

*The photo is from the Rochester Baby Lab and the photographer Adam Fenster. The photo is from the Intelligent Nest, LLC (http://intelligentnest.com/).

Appendix 5: Demographic information

Table of Maternal demographic information at 6 months

		Mean ± SD	Range
Age (years)		34.85 ± 3.32	23 - 40
Employment	Full-time work		9 (22.50%)
	Part-time work		4 (10.00%)
	Looking after family or home		2 (5.00%)
	Maternal leave		24 (60.00%)
	Missing data		1 (2.5%)
Qualification	University degree or above		34 (85.00%)
	A-levels or equivalent		2 (5.00%)
	GCSE or equivalent		3 (7.50%)
	Missing data		1 (2.5%)
Household Income	20,000 - 55,000		9 (22.50%)
	55,001 - 80,000		13 (32.50%)
	80,001 upwards		18 (42.50%)
	Missing data		1 (2.5%)
Living with partner	Yes		39 (97.5%)
	Missing data		1 (2.5%)

Appendix 6 Feedback questionnaire

Participant ID……

Date of completion……

Feedback for Each Session

Study Title: Novel biomarkers in infants: developing optical imaging solutions for the measurement of early vocal brain development

IRAS Ref: **REC Ref:**

Feedback Interview Record Card				
Mother's feedback	Strongly Disagree	Disagree	Agree	Strongly Agree
1. The length of time of the measurement procedure was reasonable				
2. The fNIRS equipment looks good to me (e.g. the equipment didn't look to ugly or daunting)				
3. The headgear fitted for my baby's head				
4. The measurement session was interesting				
5. My baby concentrated on the sounds, images and videos at most of the time				
If your baby couldn't concentrate, what do you think might have been the reason?				
6. My baby felt comfortable during the measurement				
If you think your baby felt uncomfortable, what do you think that made him/her uncomfortable?				

7. My baby liked the measurement session				
If your baby was not able to finish the session, what do you think might be the reason:				
8. If applicable, I felt comfortable being visited at home				
9. If applicable, I felt comfortable having the session at the Wellcome Trust Clinical Research Facility (CRF)				
10. I feel happy to take part in the next measurement session				
Would you like to give any suggestions for the study? Please state				

Thanks very much for taking part in our study, your effort is appreciated.

Hope to see you next time ☺

Appendix 7: Stimuli acoustic properties

Type	Duration (s)	Mean Intensity (dB)	Max Intensity (dB)	Min Intensity (dB)	Mean Pitch (Hz)	Max Pitch (Hz)	Min Pitch (Hz)
Angry 1	1.00	69.99	75.83	57.59	345.88	492.56	271.00
Angry 2	1.00	73.18	78.39	62.88	142.17	301.61	76.26
Angry 3	1.00	70.67	80.67	52.67	151.38	248.96	97.27
Angry 4	1.00	75.83	80.87	67.67	142.89	346.95	78.54
Angry 5	1.00	73.87	80.24	52.69	210.92	317.19	116.00
Angry trial	5.00	71.59	80.87	47.28	200.15	492.56	76.26
Happy 1	1.00	69.99	80.48	61.85	282.41	393.98	180.13
Happy 2	1.00	78.55	81.29	72.01	289.16	312.25	259.06
Happy 3	1.00	73.73	83.60	67.24	347.18	524.56	219.98
Happy 4	1.00	75.31	78.08	70.42	352.31	485.99	224.65
Happy 5	1.00	74.63	79.57	65.49	233.33	327.25	139.94
Happy trial	5.00	74.13	83.62	61.84	299.88	524.56	139.94
Neutral 1	1.00	74.04	79.14	58.41	207.20	218.65	199.01
Neutral 2	1.00	71.65	79.37	53.72	190.01	197.30	178.68
Neutral 3	1.00	76.18	78.34	68.43	191.14	194.65	181.69
Neutral 4	1.00	75.17	77.91	59.45	192.91	206.03	176.71
Neutral 5	1.00	75.01	80.27	57.27	215.23	239.25	191.02
Neutral trial	5.00	73.57	80.27	52.79	197.89	239.25	176.71
Voice 1	0.50	68.77	74.01	52.69	256.54	348.86	126.45
Voice 2	0.50	67.60	72.53	56.75	349.13	486.02	288.02
Voice 3	0.50	69.82	70.93	65.82	307.41	453.77	194.64
Voice 4	0.50	68.97	74.00	62.08	334.30	487.90	215.10
Voice 5	0.50	67.74	74.53	62.15	173.18	296.30	89.21
Voice trial	5.00	68.51	74.54	51.61	295.69	489.63	126.71
Non-vocal 1	0.50	69.45	72.13	65.35	223.76	503.98	89.01
Non-vocal 2	0.50	68.29	75.32	60.88	121.27	212.04	81.04
Non-vocal 3	0.50	69.78	70.89	64.14	170.90	171.66	170.18
Non-vocal 4	0.50	68.97	73.82	62.00	90.20	180.53	68.91
Non-vocal 5	0.50	69.74	71.77	66.11	353.93	359.94	351.82
Non-vocal trial	5.00	69.21	75.32	60.81	199.16	503.98	68.90