



Development of the neural processing of vocal emotion during the first year of life

Chen Zhao^a, Ingo Schiessl^{b,c}, Ming Wai Wan^a, Georgia Chronaki^{b,d,*} and Kathryn M. Abel^{b,a,e,*}

^aCentre for Women's Mental Health, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK; ^bDivision of Neuroscience and Experimental Psychology, School of Biological Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK; ^cGeoffrey Jefferson Brain Research Centre, The Manchester Academic Health Science Centre, Northern Care Alliance NHS Group, University of Manchester, Manchester, UK; ^dDevelopmental Cognitive Neuroscience (DCN) Laboratory, School of Psychology, Faculty of Science and Technology, University of Central Lancashire, Preston, UK; ^eGreater Manchester Mental Health NHS Foundation Trust, Manchester, UK

ABSTRACT

Human infants are “wired” to respond to social information, an important capacity for survival. The ability to discriminate vocal emotion in others is likely to play a key role in successful social interactions with caregivers, which facilitate the rapid social-communicative development that infants typically undergo in the latter half of their first year. Infants have voice-sensitive brain regions that have been shown previously to be responsive to emotional prosody by 7 months. This study aimed to investigate the developmental trajectory of vocal emotion processing in temporal regions using functional near-infrared spectroscopy (fNIRS) to measure brain sensitivity to angry, happy, and neutral vocalizations in the same infant at 6, 9, and 12 months. We found significant and increasing temporal cortical activation in response to vocal emotional stimuli over the three time points, suggesting consistent enhanced responses for happy compared to angry vocalizations, and vocal anger sensitivity is developing incrementally. The findings suggest that the neural processing of angry and happy prosody may follow distinct developmental pathways and is gradually “tuned” to become specialized between 6 and 12 months. This first longitudinal study of vocal emotion brain processing between 6 and 12 months highlights the need for more research to understand what drives typical and atypical social cognitive development across infancy and for follow-up into the second year.

ARTICLE HISTORY

Received 4 September 2019
Accepted 16 November 2020

KEYWORDS

fNIRS; infant; social-emotional; voice; prosody

Voice recognition is fundamental to human social interaction and has long been investigated as a foundation for social cognition and language development in infants (Blasi et al., 2015; Lloyd-Fox et al., 2013) and children (Chronaki et al., 2015, 2018). Studies in

CONTACT Chen Zhao  chenzhao510@gmail.com  Centre for Women's Mental Health, Faculty of Biology, Medicine and Health, University of Manchester, UK; Georgia Chronaki  GChronaki@uclan.ac.uk  Division of Neuroscience and Experimental Psychology, School of Biological Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK; Kathryn M. Abel  Kathryn.M.Abel@manchester.ac.uk  Centre for Women's Mental Health, Faculty of Biology, Medicine and Health, University of Manchester, UK

*Joint senior authors

 Supplemental data for this article can be accessed <https://doi.org/10.1080/09297049.2020.1853090>

© 2020 Informa UK Limited, trading as Taylor & Francis Group

utero and of newborns suggest that humans are born with a readiness to discriminate their mother's voice (Kisilevsky et al., 2003; Ockleford et al., 1988). Neuroimaging suggests that the bilateral superior temporal cortices are implicated as voice-selective areas in children (Rogier et al., 2010) and adults (Belin & Zatorre, 2000; Belin et al., 2000). Infants do not show distinct neural responses to unfamiliar human voices before the age of 4 months, and their brains may develop neural sensitivity to human voices by six or seven months of age (Grossmann et al., 2010; McDonald et al., 2019).

Emotional prosody refers to changes in the speaker's vocal intonation according to their emotional state (Banse & Scherer, 1996; Hargrove, 1997). The ability of infants to distinguish emotional prosodic features in voices may play a key role in a range of developmental domains, including infant-caregiver attachment formation (Trevarthen, 2017), infant social cognition, such as social referencing (Mumme et al., 1996) and infant learning (Doan, 2010). Behavioral studies suggest that infants consistently prefer happy voices from birth (Mastropieri & Turkewitz, 1999; Singh et al., 2002) and behaviorally respond preferentially to their mother's speech at one month, but only when the mother's speech has natural prosody (Mehler et al., 1978). By five months, infants are able to detect prosodic change between happy and sad vocalizations (Blasi et al., 2015, 2011; Walker-Andrews & Grolnick, 1983) and respond behaviorally differently to positive and negative infant-directed speech (Fernald, 1993).

The ability to extract prosodic features from vocal sounds provides the pre-verbal infant with salient information about the status of their environment. Infants' well-established preference for infant-directed speech (Cooper & Aslin, 1990; Hayashi et al., 2001) seems to be based on the heightened valence of positive emotion and wider range of expressed emotion used by parents (Panneton et al., 2006; Singh et al., 2002) and not on pitch *per se* (Kitamura & Burnham, 1998). Positive vocal prosody, such as that found in infant-directed speech, may signal safety and positive social value to the infant (Lohaus et al., 2001), while vocal negativity provides information about the potential threat or danger in the environment (Striano & Rochat, 2000; Vaish & Striano, 2004).

The neural correlates of infant behavioral preferences may start to become apparent in specialized cortical and subcortical brain responses to emotional vocalizations by around seven months of age (review by Grossmann & Johnson, 2007). Studies using a range of brain imaging techniques suggest that this neural sensitivity may emerge as early as the first month of life (Blasi et al., 2015, 2011; Zhang et al., 2017). An fNIRS study of sleeping newborns reported that emotional (happy, angry, and fearful) vocalizations compared with neutral pseudo-speech elicited greater right temporal activation (Zhang et al., 2017). Two fMRI studies reported insular and bilateral frontal responses to sad vocalizations in sleeping infants aged 3–7 months (Blasi et al., 2015, 2011). Furthermore, an fNIRS study of awake 7-month-old infants found increased right inferior frontal and superior temporal cortical responses to happy and angry, but not to neutral, speech, respectively (Grossmann et al., 2010). Neural imaging studies on vocal emotional processing in infants from 8 months onwards are scarce (see Morningstar et al., 2018). One recent ERP study showed an enhanced central and temporal response to crying and laughing compared to neutral vocalizations at 8 months (Missana et al., 2017).

While several studies implicate the temporal region for infant vocal and vocal emotion processing, understanding of this neural response, especially over the first year, is still very limited. Lloyd-Fox et al. (2017) conducted the only longitudinal between-subjects

study to date on voice sensitivity development in human infants. Tracking a rural Gambian cohort of infants aged between 2 and 24 months of age, they reported stable and relatively consistent early functional specialization of selective neural response to human voice from 9 to 13 months onwards (Lloyd-Fox et al., 2017). Whether infants become able to discriminate *prosodic* vocal content along the same timeline has not yet been investigated, yet this seems a very important competence to develop as emotional vocalizations carry information of high social value.

The current study represents the first longitudinal, within-subject fNIRS study of infant vocal emotion processing at 6, 9, and 12 months of age. We hypothesized that, at all time points, infants would show increased brain responses to emotional (angry, happy) compared to neutral vocalizations in voice-sensitive temporal regions. Second, we anticipated that infant neural responses to vocal emotional stimuli would become stronger with increasing infant age. Third, we explored whether developmental trajectories of prosodic processing would differ as a function of emotion type. No specific prediction was made given that both positive and negative prosody was expected to carry high social value, although studies suggest an early happy vocal preference from birth (Mastropieri & Turkewitz, 1999; Singh et al., 2002).

Methods

Participants

Forty infants of healthy mothers were recruited to the present study from three Manchester (UK) community health centers. All infants were born at the normal birth weight (>2500 g); 39 were full term (37–42 weeks gestation); one was born at 36 weeks gestation (corrected gestational age). All infants were typically developing and none had hearing difficulties according to parental report. Mothers consented for their infants. The UK National Health Service ethics committee approved the study (Ref: 15/NW/0684).

Forty infants were tested at 6 months (20 female, 20 male, age range: 175–214 days, $M = 189.48$ days, $SD = 9.27$), 39 were re-tested at 9 months (19 female, 20 male; age range: 263–302 days, $M = 279.08$ days, $SD = 9.46$; drop out: $N = 1$), and 38 infants were re-tested at 12 months (18 female, 20 male; age range: 360–394 days, $M = 377.24$ days, $SD = 8.61$; drop out: $N = 1$). [Figure 1](#) describes the numbers participated and analyzed at each time point.

Experimental paradigm and procedure

[Figure 2](#) presents the experimental paradigm used at all three time points. Infants wore the NIRS headband, sat on their mothers' laps in front of a laptop during the task and listened to the vocal stimuli. The task started with a 20 s rest period, followed by a 5 s trial presented through loudspeakers ($SPL = 70$ dB). Consistent with previous research (Grossmann et al., 2010), a 5 s silent cartoon was played during each trial to attract infants' attention and reduce motion artifacts. After each trial, a 10 s silent baseline (consisted of a 9 s post-stimulus baseline and a 1 s pre-stimulus baseline) along with the blurred cartoon was presented (Blasi et al., 2015, 2011; Lloyd-Fox et al., 2012). The task was presented with PsychoPy software (Peirce, 2007). The same emotional expression did not occur consecutively. There were eight trials per condition (angry, happy, neutral

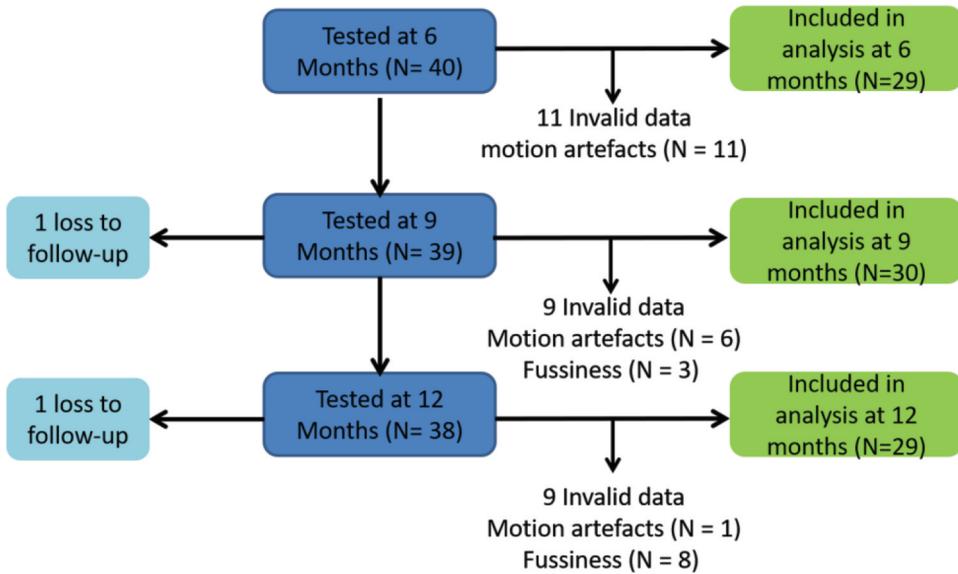


Figure 1. Total number of infants tested and included in analysis at 6, 9, and 12 months.

presented sequentially), and a total number of 24 trials. The total length of the testing session was 6 min and 20 s.

Vocal emotional stimuli

The stimulus material consisted of 15 female non-linguistic vocalizations of angry, happy prosody, and neutral vocalizations (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 et al., 2015, 2012). Five normalized 1 s stimuli from the same emotion condition were formed sequentially into a 5 s trial.

All vocal stimuli were normalized with Praat sound-analysis software (Boersma & van Heuven, 2001) to the same duration and mean intensity (see Table A1 in the Appendix for details on stimuli acoustic properties). Vocal emotional stimuli were the same at all three time points.

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 hemodynamic 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

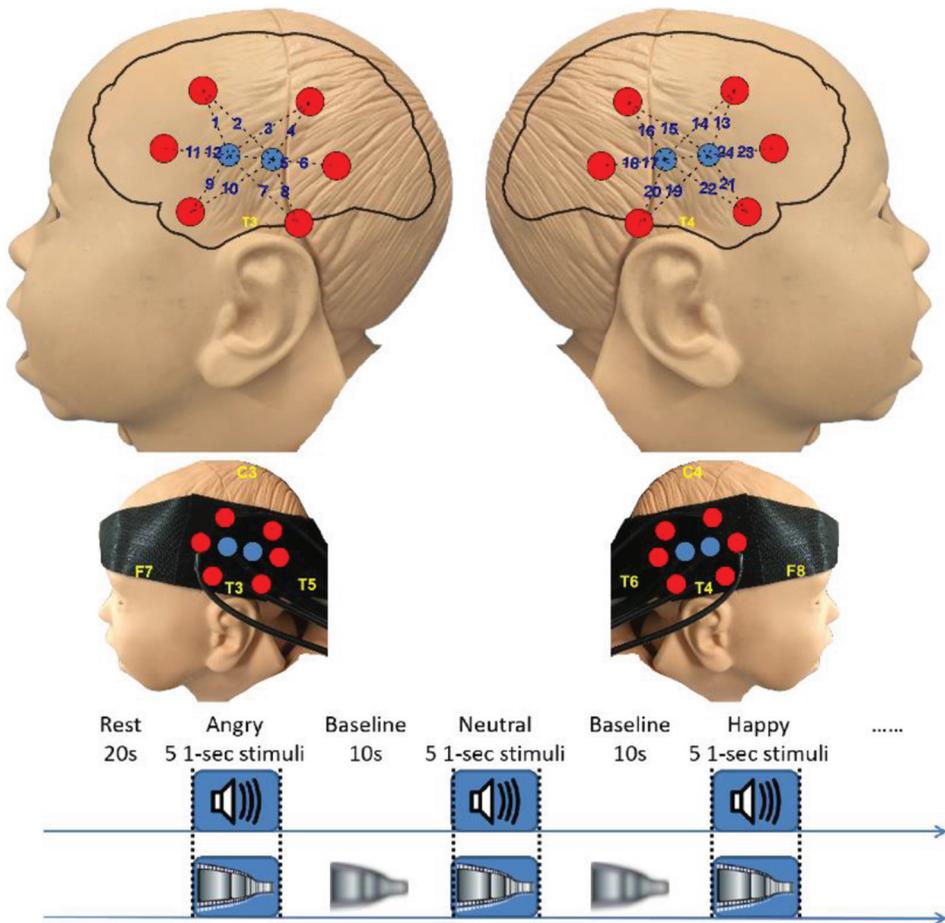


Figure 2. Study experimental task design and channel distribution. The head model illustrates the source-detector distribution where red dots represent sources (6 in each hemisphere) and blue dots represent detectors (2 in each hemisphere). Sources and detectors from 12 recording channels in each hemisphere, 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.

Biomedical Optics Research Laboratory (Dept. of Medical Physics and Bioengineering, University College London) and applied with 780 nm and 850 nm 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 et al., 2010; Lloyd-Fox et al., 2012; Pena et al., 2003; Taga & Asakawa, 2007) and adults (Belin et al., 2000; Ethofer 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 head growth standards from the World Health Organization (World Health

Organization (WHO), 2003), and from previous infants imaging studies (Li et al., 2015; Lloyd-Fox et al., 2017), the head circumference and skull thickness of 6-to 12-month-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 are presented in Figure 2.

Data analysis

To determine inclusion into the analysis, infants had to have attended to the screen without large motion artifacts for at least four out of eight trials per condition, based on videotaped observations of the experimental task. The datasets included (6 months: $N = 29$; 9 months: $N = 30$; 12 months: $N = 29$) were of a rate within the standard range for infant fNIRS 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 $N = 21$ would give 80% power to achieve a medium effect size $f = 0.29$ (Cohen, 1969, p. 348). Our sample size is consistent with previous fNIRS studies in similarly aged infants (Grossmann et al., 2010; Zhang et al., 2017).

The included datasets were filtered at 0.01 to 0.5 Hz with a 3rd order Butterworth filter, to eliminate slow drifts, instrument noise, and physiological artifacts such as heartbeats (Cooper et al., 2012; Fox et al., 2013; Grossmann et al., 2010), then converted to optical density data in HOMER2 NIRS toolbox (version 2.1, <http://homer-fnirs.org/>, Huppert et al., 2009). The remaining artifacts 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 artifact) was marked as artifacts. The time period of motion artifact 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 artifact 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 $20 \times$ 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 = 5 s; tMask = 1 s; STDEVthresh = 20; AMPthresh = 5.

After pre-processing, data were converted to Oxy- and Deoxy-Hemoglobin concentration changes in HOMER2 and averaged across trials in the same emotional condition within each dataset, with the time window of 1 s before and 15 s after the stimulation onset. The averaged time course of each channel was corrected by subtracting the mean of the 1 s before the stimulation. The analysis focused on Oxy-Hemoglobin concentration changes which seem to be a sensitive indicator of changes in cerebral blood flow

(Grossmann et al., 2010; Meek, 2002). Based on earlier work showing that the hemodynamic response reaches the peak around 2 to 4 s post stimulus (Brigadoi et al., 2014), we targeted a time window of 2 s to 9 s after stimulus onset. Mean amplitudes of cortical hemodynamic responses (Oxy- and Deoxy-Hemoglobin waveforms) were averaged over the time window of 2 s to 9 s after stimulus onset.

Repeated measures Analyses of Variance (ANOVA) were initially carried out to identify emotion sensitive channels for which there were significant differences in Oxy-Hemoglobin concentration change. Averaged hemodynamic responses were analyzed by emotion condition (angry, happy, neutral), followed by post hoc pairwise comparisons for each age (6, 9, and 12 months). Five channels showed significant Oxy-Hemoglobin concentration changes to emotional prosody at 6, 9, and 12 months (Table 1). Then, we focused on the 21 infants with valid data at all time points, and conducted a 3-way repeated measures ANOVA with age (6, 9, and 12 months), emotion (angry, happy, neutral), and location (emotion sensitive channel 2, 9, 14, 16, and 21) as within-subject factors, and Oxy-Hemoglobin concentration changes as the dependent measure. Partial eta-squared (Cohen, 1973; Kennedy, 1970) was used to estimate effect sizes of main effects and contrasts, with thresholds of 0.02, 0.13, and 0.26 indicate 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 & Hochberg, 1995) correction was applied to resolve the issue of multiple statistical comparisons. P values arranged in ascending order with an order number index allowed us to calculate adjusted α values: $\alpha_{\text{adjust}} = (\text{order index}/\text{total number of comparisons}) * 0.05$. P-value < adjusted α value remained significant (Field et al., 2012).

Table 1. Summary of emotion (Angry, Happy, Neutral) ANOVA effects on Oxy-Hemoglobin concentration changes at 6, 9, and 12 months.

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

*Comparison survived FDR correction. ^aA = Angry, H = Happy, N = Neutral

Results

The participant characteristics of those included in the final datasets analysis are described in [Figure 1](#). For the included datasets, the percentages of discarded trials at each age point are (mean \pm SD): at 6 months $3.74\% \pm 3.40\%$; at 9 months $10.00\% \pm 14.72\%$; at 12 months $4.74\% \pm 11.12\%$.

The emotion sensitive channels at each time point were identified and shown in [Table 1](#); those that have passed FDR correction are shown in [Figure 3](#). Both happy (channel 16) and angry (channel 2) vocalizations evoked significantly greater neural responses at 6 months; happy vocalizations have evoked significantly greater neural responses at 9 (channel 21) and 12 (channel 9) months.

Additionally, a 3 (emotion) \times 3 (age) \times 5 (channel) ANOVA showed a significant main effect of emotion only ($F(2, 40) = 3.86, p = .029, \eta_p^2 = .16$) and not age ($F(2, 40) = 1.22, p = .307, \eta_p^2 = .06$) or channel ($F(2, 80) = .86, p = .494, \eta_p^2 = .04$). Pairwise comparisons highlighted a significant happy $>$ angry effect ($F(1, 20) = 8.01, p = .010, \eta_p^2 = .29$, survived FDR correction), but not between happy and neutral or angry and neutral. Further, a significant age \times emotion \times location interaction effect emerged ($F(16, 320) = 2.04, p = .011, \eta_p^2 = .09$).

To further interpret the 3-way interaction effect, an age \times emotion repeated measures ANOVA was conducted for each channel location, followed by pairwise comparisons. There was no significant main effect or interaction in channel 2, 14, and 21. A significant main effect of emotion was found in channel 9 ($F(2, 40) = 4.39, p = .019, \eta_p^2 = .18$), which was attributed to the happy $>$ angry condition ($F(1, 20) = 7.83, p = .011, \eta_p^2 = .28$, survived FDR correction); age and age \times emotion interaction effects were not significant. In channel 16, a significant age

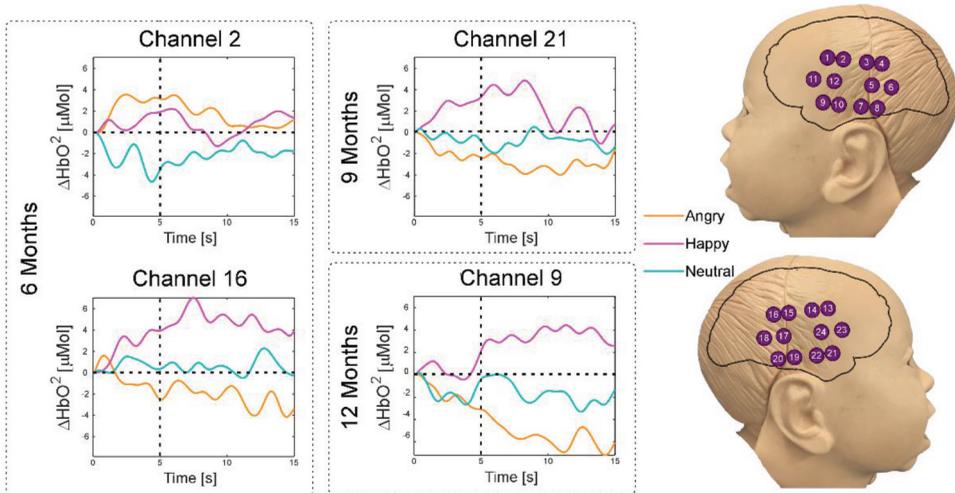


Figure 3. Averaged time courses of Oxy-Hemoglobin concentration changes in channels showing significant emotion effect at each age point (6, 9 and 12 month) per vocal emotion (Angry in orange, Happy in pink and Neutral in light blue) in the time period of 15 s (5 s stimulus and 10 s baseline). The channel location is marked in the right 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.

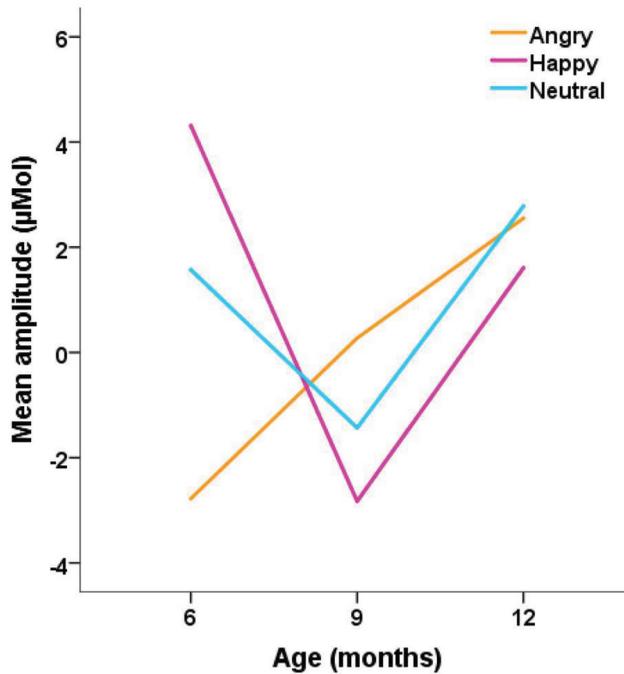


Figure 4. Mean amplitudes of Oxy-Hemoglobin concentration changes for angry, happy and neutral in channel 16 at ages of 6, 9 and 12 months.

effect ($F(2, 40) = 3.40, p = .043, \eta_p^2 = .15$) was due mainly to the increased Oxy-Hemoglobin concentration changes at $12 > 9$ months ($F(1, 20) = 8.02, p = .010, \eta_p^2 = .29$, survived FDR correction). The effect of emotion was not significant in channel 16, while an age \times emotion interaction ($F(4, 80) = 2.98, p = .024, \eta_p^2 = .13$) was localized to the happy vs. angry comparisons at 6 to 9 months and 6 to 12 months ($F(1, 20) = 14.80, p = .001, \eta_p^2 = .43$; and $F(1, 20) = 8.32, p = .009, \eta_p^2 = .29$, respectively, both survived FDR correction). Responses were stronger for happy than angry vocalizations at 6 months and reversed (angry $>$ happy) at 9 and 12 months (see Figure 4). To further test neural sensitivity to angry prosody in channel 16 over time, a Pearson correlation was conducted using data over 6–12 months, and found increased local neural responses with age ($r = .27, p = .030$).

Discussion

The present study is unique in using a within-subject design to track the developmental trajectory of temporal cortical activation to human emotional (angry, happy) non-speech vocalizations across the first year of an infant's life. There are three main findings of note. First, at all age points (6, 9, and 12 months), we found a significant main effect of activation in temporal cortices in response to vocal emotional stimuli, particularly for happy (versus angry) vocalizations. Second, despite a lack of overall main age effect, right temporal (channel 16) responses to vocal emotional stimuli increased significantly with

age, especially between 9 and 12 months. Third, infants' neural responses to vocal anger in the right hemisphere (channel 16) increased significantly with age, while sensitivity to vocal happiness did not, suggesting that infants may follow distinct developmental trajectories for processing angry and happy prosody.

Consistent with our prediction, temporal cortical activations to vocal emotions overall were consistently elicited in infants aged 6, 9, and 12 months. The temporal cortices play a key role in decoding acoustic features in human vocalizations (Belin & Zatorre, 2000; Zhang et al., 2018), even before infants fully acquire language skills (Blasi et al., 2015; Grossmann et al., 2010; Missana et al., 2017). Emotional vocalizations carry both human voice and emotion features and have been found to be meaningful in infant neural responses at 3–7 months (Blasi et al., 2015; Grossmann et al., 2010, 2005). Findings from the present study not only expand the understanding of vocal emotion neural mechanisms beyond 7 months of infancy but also confirm the ongoing development of temporal regions in decoding prosodic features in vocalizations between 9 and 12 months. There was no significant main effect for location in the longitudinal data and the age \times emotion \times location interaction was driven by the age \times emotion interaction effect in channel 16. Although neural responses to emotional stimuli in channel 16 in the right hemisphere resemble the right lateralized effect found in other infant and adult studies (Alba-Ferrara, Ellison & Mitchell, 2012; Grossmann et al., 2010; Ross & Monnot, 2011; Zhang et al., 2018), we were unable to provide a laterality effect, which may require a larger sample.

Our findings suggest that there is a neural sensitivity preference in typically developing young infants for positive prosody over negative prosody, in line with findings from other neuroimaging studies of 7- and 8-month-olds (Grossmann et al., 2010, 2005; Missana et al., 2017) and adults (Pinheiro et al., 2017). Our report of enhanced brain responses to happy prosody supports previous data that typically developing infants prefer happy voices early in life from birth (Mastropieri & Turkewitz, 1999; Singh et al., 2002); alongside the emotional tones of infant-directed speech (Mehler et al., 1978). Social interactions with caregivers in healthy mother and young infant dyads are typically characterized by smiling faces and happy voices (Eisenberg et al., 1998; Malatesta et al., 1986), which promote mother-infant attachment and infant emotion regulation (Leigh et al., 2011; Lohaus et al., 2001; Thompson, 1997). Infants' sensitivity to, and interest in, happy vocalizations from birth and their familiarity with positive vocalizations from interactions with caregivers may drive infant cortical responses to prioritize responses to happy prosody. In 6-month-old infants, we also found enhanced left temporal responses (channel 2) to angry compared to neutral voices consistent with previous research (Grossmann et al., 2010), possibly suggesting enhanced neural processing of negative stimuli early in development. Sensitivity to negative vocal emotion serves as an adaptive function, which is not detected in the temporal region in brain studies until the latter half of an infant's first year (Vaish et al., 2008). Surprisingly, we have not found significantly greater neural responses to angry compared to neutral or happy at 9 or 12 months. This is the first study, as far as we are aware, with longitudinal data in the same infants from 9 to 12 months of age on neural responses to emotional vocalizations, so we are unable to compare this set of findings with others. However, we did find consistently increasing neural responses to angry vocalizations with age in the right temporal cortex (channel 16), which is discussed below.

Our second prediction that, overall, across all cortical regions, neural response to emotional prosody would grow incrementally with age was not supported. In the sub-statistical analysis, we found a channel 16 right hemisphere effect of age with increasing vocal sensitivity, especially between 9 and 12 months and especially in response to angry prosody. Our findings are in line with EEG evidence that the alpha power frequency to visual and auditory stimuli increases with age between 5 months and 4 years (Marshall et al., 2002; Michel et al., 2015). The incremental neural responses to emotional vocalizations from 6 months to 12 months extend previous research reporting that the strength of voice-selective neural responses becomes more prominent with the infant's age between 3 and 7 months (Blasi et al., 2011; Lloyd-Fox et al., 2012). Previous research also suggests that infants are going through a transitional period between 4 and 8 months before voice-sensitive neural responses become relatively stable after 9 months (Lloyd-Fox et al., 2017); therefore, the increasing neural sensitivity to emotional vocalizations after 9 months of age may suggest the incremental development of cortical specialization. The question of when vocal emotion neural responses become specialized requires further investigation in infants beyond 1 year.

It should be noted that the longitudinal neural responses to emotional vocalizations did not follow a linear pathway. This non-linear neural development suggests that some channels may be sensitive to emotional vocalizations in general and not specialized to respond to specific emotions. Furthermore, this finding may reflect infants' rapid and variable pace of cognitive and social emotional development between 6 and 12 months, as with cognitive skills, such as language (Bates et al., 1991; Fenson et al., 1994); and/or individual differences in exposure to vocal and communicative input by caregivers and others (e.g., siblings, daycare staff). Additionally, we speculate that the drop in neural responses to neutral and happy vocalizations in channel 16 at 9 months may reflect a widely known phenomenon, stranger anxiety, that peaks around 9 months and may lead infants to perceive all stranger vocal sounds as negative; or otherwise parse the incoming auditory stimuli differently than at other times (e.g., Schore, 2001). However, this is the first study to examine vocal emotion processing longitudinally in the same infants between 6 and 12 months and we caution against further speculation in this emerging field.

Consistent with our third research question, the longitudinal analysis revealed distinct developmental trajectories for angry and happy vocalizations across the three time points. There were consistently enhanced neural responses to happy vocalizations over all time points (6, 9, and 12 months), while infant neural responses to vocal anger were stronger with age. Infants prioritize interest to happy emotional vocalizations from birth (Mastropieri & Turkewitz, 1999); they are better able to discriminate happy than negative expressions until around 6 months of age when they begin to shift attention to more negative vocalization (Grossmann et al., 2010; Vaish et al., 2008). This social-emotional developmental trend is evident in social referencing by the end of the first year. From 9 months onwards, infants tend to show less behavioral exploration on hearing negative vocalizations in ambiguous situations (see Walker-Andrews, 1997) (Mumme et al., 1996). Our earlier study of the current sample reports that, at 6 months, infant neural responses to angry, in contrast to neutral, vocalizations were positively associated with maternal directiveness, i.e., the degree to which caregivers interact or comment in a caregiver-centered way (Zhao et al., 2019). It is possible that infants perceive caregivers' re-directive or incongruent (to the infant) behavior and this may evoke negative emotions in the infant, which may be reflected in enhanced neural responses to angry vocalizations. These findings

support the non-linear development of the temporal region for processing vocal emotion in infancy (especially negative emotion such as anger) 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). Our finding of changing neural response to vocal emotion across different infant ages is consistent with the notion that the brain undergoes a process of “fine-tuning” to vocal emotional signals across the first year of life (Johnson et al., 2009; Kolb & Gibb, 2011; Leppänen & Nelson, 2009).

This is the first longitudinal, within-subject fNIRS study of the neural correlates of vocal emotion development in human infants; however, there remain important limitations. We did not find a location effect. Replication of the present paradigm, with an even larger sample, may clarify localization effects in infant neural processing of vocal emotion. Since our emotional stimuli consisted of one negative and one positive emotion, it is unclear whether the neural activations were emotion-specific (e.g., anger) or valence-specific (e.g., negative). It is also unclear whether the sensitivity for happy vocalizations reflects an innate preference or a learned one, as a result of familiarity with positive affect. This may be important in the context of family and parenting interventions. In addition, the stimuli used were unfamiliar female voices, which may be processed differently to familiar voices or caregiver voices. Infants may be more responsive to their own mother’s voice than other voices (Dehaene-Lambertz et al., 2010; Walker-Andrews et al., 2011). Furthermore, subcortical regions and frontal regions may also be implicated in vocal emotion processing (Blasi et al., 2015, 2011), but technical limitations of our particular fNIRS system and the present source-detector setting restricted our observation of these regions. Future studies of neural responses to emotional vocalizations should include frontal and subcortical regions, such as insula and hippocampus.

The present study provides novel evidence for the neural development of vocal emotion processing between 6 and 12 months of age. This is the first longitudinal study tracking infants’ neural responses to emotional vocalizations beyond 8 months of age. Our findings support a pattern of the progressive development of superior temporal cortical sensitivity to vocal emotion prosody in typically developing infants. These results have implications for understanding social-emotional development in typically developing infants. Our findings suggest that infant social emotional development is rapid; such developmental trajectories might begin from an even earlier age before 6 months and continue beyond the first year. Future research would benefit not only from examining broader age ranges of infants but might also consider how inter-individual differences, cognitive development, early experience of caregivers and their mother’s own emotion processing behavior influence this social-emotional developmental trajectory. In the future, our potential to intervene effectively in at-risk infants, or in a trajectory of atypical development, will depend on our ability to understand what influences healthy development and the parameters of typical and atypical trajectories from infancy.

Acknowledgments

The authors wish to thank the families who participated in the study.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Kathryn M. Abel  <http://orcid.org/0000-0002-5244-6528>

References

- 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. <https://doi.org/10.1016/j.brs.2011.06.004>
- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70(3), 614–636. <https://doi.org/10.1037/0022-3514.70.3.614>
- Bates, E., Bretherton, L., & Snyder, L. S. (1991). *From first words to grammar: Individual differences and dissociable mechanisms* (Vol. 20). Cambridge University Press.
- Belin, P., & Zatorre, R. J. (2000). ‘What’, ‘where’ and ‘how’ in auditory cortex. *Nature Neuroscience*, 3(10), 965–966. <https://doi.org/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. <https://doi.org/10.1038/35002078>
- 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. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.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. <https://doi.org/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. <https://doi.org/DOI10.1016/j.cub.2011.06.009>
- Boersma, P., & van Heuven, V. (2001). Speak and unSpeak with PRAAT. *Glott International*, 5(9/10), 341–347.
- 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. <https://doi.org/10.1016/j.neuroimage.2013.04.082>
- 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. <https://doi.org/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. <https://doi.org/10.1016/j.dcn.2011.11.007>
- Chronaki, G., Wigelsworth, M., Pell, M. D., & Kotz, S. A. (2018). The development of cross-cultural recognition of vocal emotion during childhood and adolescence. *Scientific Reports*, 8(1), 8659. <https://doi.org/10.1038/s41598-018-26889-1>
- Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. Academic Press.
- Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and Psychological Measurement*, 33(1), 107–112. <https://doi.org/10.1177/001316447303300111>
- 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. <https://doi.org/10.3389/fnins.2012.00147>
- Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, 61(5), 1584–1595. <https://doi.org/10.2307/1130766>

- Dehaene-Lambertz, G., Montavont, A., Jobert, A., Alliol, 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. <https://doi.org/10.1016/j.bandl.2009.09.003>
- Doan, S. N. (2010). The role of emotion in word learning. *Early Child Development and Care*, 180(8), 1065–1078. <https://doi.org/10.1080/03004430902726479>
- Eisenberg, N., Cumberland, A., & Spinrad, T. L. (1998). Parental socialization of emotion. *Psychological Inquiry*, 9(4), 241–273. https://doi.org/10.1207/s15327965pli0904_1
- 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. <https://doi.org/10.1097/01.wnr.0000199466.32036.5d>
- 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):093705. <https://doi.org/10.1063/1.2038567> doi Artn:093705
- 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. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- 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), i–185. <https://doi.org/10.2307/1166093>
- Fernald, A. (1993). Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development*, 64(3), 657–674. <https://doi.org/10.2307/1131209>
- Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Sage publications.
- 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. *Frontiers in Human Neuroscience*, 7. <https://doi.org/10.3389/Fnhum.2013.00089>
- 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. <https://doi.org/10.1364/Ao.42.002881>
- 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. <https://doi.org/10.1038/nn1392>
- Grossmann, T., & Johnson, M. H. (2007). The development of the social brain in human infancy. *European Journal of Neuroscience*, 25(4), 909–919. <https://doi.org/10.1111/j.1460-9568.2007.05379.x>
- 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. <https://doi.org/10.1016/j.neuron.2010.03.001>
- Grossmann, T., Striano, T., & Friederici, A. D. (2005). Infants' electric brain responses to emotional prosody. *Neuroreport*, 16(16), 1825–1828. <https://doi.org/10.1097/01.wnr.0000185964.34336.b1>
- Hargrove, P. M. (1997). Prosodic aspects of language impairment in children. *Topics in Language Disorders*, 17(4), 76–83. <https://doi.org/10.1097/00011363-199708000-00009>
- Hayashi, A., Tamekawa, Y., & Kiritani, S. (2001). Developmental change in auditory preferences for speech stimuli in Japanese infants. *Journal of Speech, Language, and Hearing Research*, 44(6), 1189–1200. [https://doi.org/10.1044/1092-4388\(2001\)092](https://doi.org/10.1044/1092-4388(2001)092)
- 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. <https://doi.org/10.1364/AO.48.00D280>
- 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. <https://doi.org/10.1037/a0014548>

- Kennedy, J. J. (1970). The eta coefficient in complex ANOVA designs. *Educational and Psychological Measurement*, 30(4), 885–889. <https://doi.org/10.1177/001316447003000409>
- 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. <https://doi.org/10.1111/1467-9280.02435>
- Kitamura, C., & Burnham, D. (1998). *Acoustic and affective qualities of IDS in English*. Paper presented at the fifth international conference on spoken language processing. Sydney, Australia.
- Kolb, B., & Gibb, R. (2011). Brain plasticity and behaviour in the developing brain. *J Can Acad Child Adolesc Psychiatry*, 20(4), 265–276.
- 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. <https://doi.org/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. <https://doi.org/10.1038/nrn2554>
- Li, Z., Park, B. K., Liu, W., Zhang, J., Reed, M. P., Rupp, J. D., ... Hu, J. (2015). A statistical skull geometry model for children 0-3 years old. *Plos One*, 10(5), e0127322. <https://doi.org/10.1371/journal.pone.0127322>
- 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*, 25, 92–104. <https://doi.org/10.1016/j.dcn.2016.11.005>
- Lloyd-Fox, S., Blasi, A., & Elwell, C. E. (2010). Illuminating the developing brain: The past, present and future of functional near infrared spectroscopy. *Neuroscience and Biobehavioral Reviews*, 34(3), 269–284. <https://doi.org/10.1016/j.neubiorev.2009.07.008>
- 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), 1758. <https://doi.org/10.1098/rspb.2012.3026>
- 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. <https://doi.org/10.1080/17470919.2011.614696>
- 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. https://doi.org/10.1207/S15327922par0104_1
- 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. <https://doi.org/10.2307/1130587>
- Marshall, P. J., Bar-Haim, Y., & Fox, N. A. (2002). Development of the EEG from 5 months to 4 years of age. *Clinical Neurophysiology*, 113(8), 1199–1208. [https://doi.org/10.1016/S1388-2457\(02\)00163-3](https://doi.org/10.1016/S1388-2457(02)00163-3)
- Mastropieri, D., & Turkewitz, G. (1999). Prenatal experience and neonatal responsiveness to vocal expressions of emotion. *Developmental Psychobiology*. 35(3):204–214. [https://doi.org/10.1002/\(Sici\)1098-2302\(199911\)35\(3\)204-214](https://doi.org/10.1002/(Sici)1098-2302(199911)35(3)204-214)
- Maurage, P., Joassin, F., Philippot, P., & Campanella, S. (2007). A validated battery of vocal emotional expressions. *Neuropsychological Trends*, 2(1), 63–74.
- McDonald, N. M., Perdue, K. L., Eilbott, J., Loyal, J., Shic, F., & Pelphrey, K. A. (2019). Infant brain responses to social sounds: A longitudinal functional near-infrared spectroscopy study. *Developmental Cognitive Neuroscience*, 36, 100638. <https://doi.org/10.1016/j.dcn.2019.100638>
- Meek, J. (2002). Basic principles of optical imaging and application to the study of infant development. *Developmental Science*, 5(3), 371–380. <https://doi.org/10.1111/1467-7687.00376>
- Mehler, J., Bertoncini, J., Barriere, M., & Jassikgerschenfeld, D. (1978). Infant recognition of mothers voice. *Perception*, 7(5), 491–497. <https://doi.org/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. <https://doi.org/10.1016/j.neuroimage.2015.06.042>

- Missana, M., Altwater-Mackensen, N., & Grossmann, T. (2017). Neural correlates of infants' sensitivity to vocal expressions of peers. *Developmental Cognitive Neuroscience*, 26, 39–44. <https://doi.org/10.1016/j.dcn.2017.04.003>
- Morningstar, M., Nelson, E. E., & Dirks, M. A. (2018). Maturation of vocal emotion recognition: Insights from the developmental and neuroimaging literature. *Neuroscience and Biobehavioral Reviews*, 90, 221–230. <https://doi.org/10.1016/j.neubiorev.2018.04.019>
- 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. <https://doi.org/Doi10.2307/1131775>
- Murphy, K. R., Myors, B., & Wolach, A. (2014). *Statistical power analysis: A simple and general model for traditional and modern hypothesis tests* (4th ed.). Routledge.
- Ockleford, E. M., Vince, M. A., Layton, C., & Reader, M. R. (1988). Responses of neonates to parents and other voices. *Early Human Development*, 18(1), 27–36. [https://doi.org/10.1016/0378-3782\(88\)90040-0](https://doi.org/10.1016/0378-3782(88)90040-0)
- Panneton, R., Kitamura, C., Mattock, K., & Burnham, D. (2006). Slow speech enhances younger but not older infants' perception of vocal emotion. *Research in Human Development*, 3(1), 7–19. https://doi.org/10.1207/s15427617rhd0301_2
- Peirce, J. W. (2007). PsychoPy - psychophysics software in python. *Journal of Neuroscience Methods*, 162(1–2), 8–13. <https://doi.org/10.1016/j.jneumeth.2006.11.017>
- 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. <https://doi.org/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. <https://doi.org/10.1016/j.cortex.2017.03.018>
- 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. <https://doi.org/10.1016/j.ijpsycho.2009.10.013>
- 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. <https://doi.org/10.1016/j.neuropsychologia.2010.12.024>
- Scholkmann, F., Spichtig, S., Muehleemann, 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. <https://doi.org/10.1088/0967-3334/31/5/004>
- Schore, A. N. (2001). Effects of a secure attachment relationship on right brain development, affect regulation, and infant mental health. *Infant Mental Health Journal: Official Publication of the World Association for Infant Mental Health*, 22(1-2), 7–66. [https://doi.org/10.1002/1097-0355\(200101/04\)22:1<7::AID-IMHJ2>3.0.CO;2-N](https://doi.org/10.1002/1097-0355(200101/04)22:1<7::AID-IMHJ2>3.0.CO;2-N)
- Singh, L., Morgan, J. L., & Best, C. T. (2002). Infants' listening preferences: baby talk or happy talk? *Infancy*, 3(3), 365–394. https://doi.org/10.1207/S15327078in0303_5
- Striano, T., & Rochat, P. (2000). Emergence of selective social referencing in infancy. *Infancy*, 1(2), 253–264. https://doi.org/10.1207/S15327078in0102_7
- 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. <https://doi.org/10.1016/j.neuroimage.2007.04.037>
- Thompson, R. A. (1997). Sensitivity and security: New questions to ponder. *Child Development*, 68(4), 595–597. <https://doi.org/10.2307/1132109>
- Trevarthen, C. (2017). The function of emotions in early infant communication and development *New perspectives in early communicative development* (pp. 48–81): Routledge.
- 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. <https://doi.org/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. <https://doi.org/10.1111/j.1467-7687.2004.00344.x>
- Villringer, A., & Chance, B. (1997). Non-invasive optical spectroscopy and imaging of human brain function. *Trends in Neurosciences*, 20(10), 435–442. [https://doi.org/10.1016/S0166-2236\(97\)01132-6](https://doi.org/10.1016/S0166-2236(97)01132-6)
- Walker-Andrews, A. S. (1997). Infants' perception of expressive behaviors: Differentiation of multimodal information. *Psychological Bulletin*, 121(3), 437–456. <https://doi.org/10.1037/0033-2909.121.3.437>
- Walker-Andrews, A. S., & Grolnick, W. (1983). Discrimination of vocal expressions by young infants. *Infant Behavior & Development*, 6(4), 491–498. [https://doi.org/10.1016/S0163-6383\(83\)90331-4](https://doi.org/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. <https://doi.org/10.1037/a0024435>
- World Health Organization (WHO). (2003). *Child growth standards*.
<http://www.who.int/childgrowth/standards/>
- 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. <https://doi.org/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. <https://doi.org/10.1038/s41598-017-18683-2>
- Zhao, C., Chronaki, G., Schiessl, I., Wan, M. W., & Abel, K. M. (2019). Is infant neural sensitivity to vocal emotion associated with mother-infant relational experience? *Plos One*, 14(2), e0212205. <https://doi.org/10.1371/journal.pone.0212205>

Appendices

The vocal stimuli were evaluated as for the following acoustic parameters: i) mean duration which was identical for all stimuli (5 s); 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 ($t = 4.47$, $p < .05$).

Additional results

The following is a summary of Deoxy-Hemoglobin concentration changes at 6, 9, and 12 months.

6-month-olds

The analysis of Deoxy-Hemoglobin concentration changes complemented the Oxy-Hemoglobin changes. Specifically, there was a significant effect of emotion on Deoxy-Hemoglobin concentration changes in channel 2 in the left hemisphere ($F(2, 56) = 4.04$, $p = .020$, $\eta_p^2 = .13$), with decreased Deoxy-Hemoglobin concentration changes in response to angry voice compared to neutral voice ($F(1, 28) = 10.26$, $p = .003$, $\eta_p^2 = .27$, survived FDR correction). Furthermore, there was a significant effect of emotion in channel 16 in the right hemisphere ($F(2, 56) = 3.62$, $p = .030$, $\eta_p^2 = .11$), with decreased Deoxy-Hemoglobin concentration changes in response to happy voice compared to angry voice ($F(1, 28) = 7.45$, $p = .010$, $\eta_p^2 = .21$, survived FDR correction).

9-month-olds

Table A1. 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

Two channels showed a significant effect of emotion in Deoxy-Hemoglobin 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.

12-month-olds

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

The 3-way ANOVA of emotion (angry, happy, and neutral) x age (6, 9, and 12 months) x channel (channel 2, 9, 10 and 16) showed a significant main age x emotion x channel interaction effect ($F(12, 240) = 2.32, p = .008, \eta_p^2 = .10$). Main effects of emotion ($F(2, 40) = 1.41, p = .257, \eta_p^2 = .07$), age ($F(2, 40) = 0.01, p = .986, \eta_p^2 = .001$) and channel ($F(2, 60) = .32, p = .813, \eta_p^2 = .02$) were not significant. The 3-way interaction was further explored, and the age x emotion interaction was found in channel 16 only ($F(4, 80) = 3.39, p = .013, \eta_p^2 = .15$). That is, at 6 months, neural responses were decreased for happy compared to angry voices, while at 12 months, responses were decreased for angry compared to happy voices ($F(1, 20) = 11.95, p = .002, \eta_p^2 = .37$).