


## RESEARCH ARTICLE

# Near-infrared spectroscopy reveals neural perception of vocal emotions in human neonates

Dandan Zhang<sup>1,2</sup>  | Yu Chen<sup>1</sup> | Xinlin Hou<sup>3</sup> | Yan Jing Wu<sup>4</sup>

<sup>1</sup>College of Psychology and Sociology, Shenzhen University, Shenzhen, China

<sup>2</sup>Shenzhen Key Laboratory of Affective and Social Cognitive Science, Shenzhen University, Shenzhen, China

<sup>3</sup>Department of Pediatrics, Peking University First Hospital, Beijing, China

<sup>4</sup>Faculty of Foreign Languages, Ningbo University, Ningbo, China

## Correspondence

Yan Jing Wu, PhD, Faculty of Foreign Languages, Ningbo University, Ningbo 315211, China.

Email: wuyanjing@nbu.edu.cn and

Xinlin Hou, PhD, Department of Pediatrics, Peking University First Hospital, #8 Xishiku Street, Xicheng District, Beijing 100034, China. Email: houxinlin66@sina.com

## Funding information

National Natural Science Foundation of China, Grant/Award Number: 31571120; Shenzhen Basic Research Project, Grant/Award Number: JCYJ20170302143246158; Project of Humanities and Social Sciences of the Ministry of Education in China, Grant/Award Number: 15YJC190002; Beijing Municipal Science & Technology Commission, Grant/Award Number: Z161100002616011; Wong Magna KC Fund in Ningbo University

## Abstract

Processing affective prosody, that is the emotional tone of a speaker, is fundamental to human communication and adaptive behaviors. Previous studies have mainly focused on adults and infants; thus the neural mechanisms underlying the processing of affective prosody in newborns remain unclear. Here, we used near-infrared spectroscopy to examine the ability of 0-to-4-day-old neonates to discriminate emotions conveyed by speech prosody in their maternal language and a foreign language. Happy, fearful, and angry prosodies enhanced neural activation in the right superior temporal gyrus relative to neutral prosody in the maternal but not the foreign language. Happy prosody elicited greater activation than negative prosody in the left superior frontal gyrus and the left angular gyrus, regions that have not been associated with affective prosody processing in infants or adults. These findings suggest that sensitivity to affective prosody is formed through prenatal exposure to vocal stimuli of the maternal language. Furthermore, the sensitive neural correlates appeared more distributed in neonates than infants, indicating a high-level of neural specialization between the neonatal stage and early infancy. Finally, neonates showed preferential neural responses to positive over negative prosody, which is contrary to the “negativity bias” phenomenon established in adult and infant studies.

## KEYWORDS

affective prosody, angular gyrus, inferior frontal gyrus, neonate, positivity preference, superior temporal gyrus

## 1 | INTRODUCTION

Speech processing and adaptive behaviors in social contexts depend on effective decoding of vocal emotions (Frühholz & Grandjean, 2013a), a skill that emerges early in human life (Grossmann, Oberecker, Koch, & Friederici, 2010; Vaish, Grossmann, & Woodward, 2008; Vaish & Striano, 2004). For instance, infants at the age of 3–7 months have been shown to respond to emotional changes in speech prosody (e.g., Blasi et al., 2011; Flom & Bahrick, 2007; Grossmann et al., 2010). However, to date, only two studies have investigated the processing of affective prosody in neonates. Mastropieri and Turkewitz (1999) showed that happy as compared to angry, sad, and neutral prosody increased eye-opening responses in neonates. Moreover, the perception of vocal emotions was observed when neonates listened to their native (i.e., maternal) language, but

not the foreign language. More recently, an event-related potential (ERP) study showed that fearful prosody, compared to happy prosody, elicited mismatch negativities to a greater amplitude in neonates, suggesting that such discrimination is neurobiologically encoded early in life (Cheng, Lee, Chen, Wang, & Decety, 2012). While this evidence suggests emotion-specific responses to speech prosody in neonates, the pattern of emotional effects was contradictory between studies. Findings of Cheng et al. (2012) are mostly consistent with the negativity bias phenomenon wherein processing priority is given to negative information (Ito, Larsen, Smith, & Cacioppo, 1998). The negativity bias phenomenon typically develops between 6 and 12 months of age in the auditory domain (e.g., Grossmann, Striano, & Friederici, 2005; Grossmann et al., 2010; Hoehl & Striano, 2010; Peltola, Leppänen, Mäki, & Hietanen, 2009) and is observed mainly in the visual domain in adults (e.g., sensitivity to negative scenes and intensity in distracting

tasks; Yuan et al., 2009; Yuan, Meng, Yang, Hu, & Yuan, 2012; Meng, Yuan, & Li, 2009). In contrast, Mastropieri and Turkewitz (1999) showed preferential responses to positive emotions in vocalizations, a finding that is consistent with infants less than 6 months old (e.g., Farroni, Menon, Rigato, & Johnson, 2007; Fernald, 1993; Rigato, Farroni, & Johnson, 2010; Vaish et al., 2008). Therefore, the emotional discrimination of speech prosody in neonates remains poorly understood and the underlying neural mechanism has not been specified.

The current study examined neural responses to speech prosody conveying different emotions in 0-to-4-day-old neonates. We used near-infrared spectroscopy (NIRS) to characterize the underlying functional neuroanatomical correlation given NIRS's high spatial resolution in comparison to ERPs (Grossmann et al., 2010) and its low degree of invasiveness as compared to other neuroimaging techniques (e.g., fMRI). We presented neonates with speech samples produced with negative (i.e., angry and fearful), positive (i.e., happy), and neutral prosodies either in their maternal language (i.e., Mandarin Chinese) or a foreign language (i.e., Portuguese). This design allows three critical questions to be addressed: (a) whether there is preferential processing of positive versus negative prosodic emotions in neonates; (b) whether the perception of prosodic emotion is acquired prenatally (e.g., Mastropieri & Turkewitz, 1999) so that neonates respond more strongly to vocal emotions presented in their maternal language compared to a foreign language; and (c) whether the neural substrates in older infants or adults exist in neonates, or there is a development stage with neural specialization.

Research on language perception in neonates and fetuses has shown preferential responses to stimuli presented in the native language over unknown languages or the second language in babies born in a bilingual environment (Kisilevsky et al., 2009; Mehler et al., 1988; Moon, Cooper, & Fifer, 1993; Sato et al., 2012; Vannasing et al., 2016). One possibility is that the prosodic cues of a familiar language may have caught more attention of neonatal and prenatal participants with stronger emotional resonance compared to those of unfamiliar languages (Mehler et al., 1988; Mastropieri & Turkewitz, 1999). However, studies on adults have repeatedly shown a cross-cultural agreement in the processing of facial and vocal expressions (Ekman et al., 1987; Pell, Monetta, Paulmann, & Kotz, 2009; Thompson & Balkwill, 2006), suggesting that the representations of basic emotions are independent of culture and language.

Our hypothesis is that neonates would respond to affective prosody in both the native and foreign languages, and that their emotional responses would be stronger in the native language due to prenatal learning (e.g., Mastropieri & Turkewitz, 1999). Previous studies have typically associated functional specialization for speech processing with activities in the left temporal regions of the brain (e.g., Peña et al., 2003). However, nonemotional, prosodic manipulations modulate activity in the right temporal regions (Arimitsu et al., 2011; Sambeth, Ruohio, Alku, Fellman, & Huotilainen, 2008; Telkemeyer et al., 2009) and in the frontal and parietal areas of the brain in neonates (Arimitsu et al., 2011; Saito et al., 2007). Moreover, studies on adults have shown systematic involvement of the right superior temporal gyrus (STG; Brück, Kreifelts, & Wildgruber, 2011; Ethofer et al., 2012; Wildgruber, Ethofer, Grandjean, & Kreifelts, 2009) and the right superior temporal sulcus (STS; Belin, Zatorre, Lafaille, Ahad, & Pike, 2000; Grandjean et al., 2005) in the perception of vocal emotions. In Grossmann et al. (2010), happy prosody, as compared to angry prosody, enhanced activation levels of the right

inferior frontal gyrus (IFG) in infants, suggesting category-specific involvement of the right IFG in the early development of emotional perception. These findings are consistent with the functional lateralization hypothesis (Van Lancker Sidtis, Pachana, Cummings, & Sidtis, 2006), which proposes that hemispheric lateralization of prosodic processing is affected by (a) the prosodic information (emotional = right hemisphere; linguistic = left hemisphere) and (b) the size of the speech units (long-range intonation vs. short-range stress) on which prosodic variations operate (small unit = left; large unit = right; Häuser & Domahs, 2014). Emotional prosody in long-range intonation changes should be associated with activations in the right hemisphere (particularly the auditory areas), given its emotional content and large unit size (for a review, see Belyk & Brown, 2014). Therefore, we expected speech prosody in the three emotional conditions to elicit enhanced neural responses in the right temporal (including the STG) and frontal-parietal areas (including the IFG) as compared to the neutral condition. Moreover, since the right STG and the right STS have been implicated in adult studies, different patterns of neural activations observed in neonates would bring insights into early neural development of the perception of vocal emotions.

## 2 | MATERIALS AND METHODS

### 2.1 | Participants

Sixty healthy full-term neonates (30 girls; gestational age: 38–41 weeks, mean =  $39.2 \pm 1.0$  weeks) with postnatal ages ranging from 0 to 4 days (mean =  $1.9 \pm 0.7$  days) participated in this study. They met the following criteria: (a) birth weight in the normal range for gestational age; (b) clinically asymptomatic at the time of NIRS recording; (c) no sedation or medication for at least 48 hr before the recording; (d) normal results of hearing screening test using evoked otoacoustic emissions (OAE, ILO88 Dpi, Otodynamics Ltd, Hatfield, UK); (e) Apgar scores at 1 min and 5 min after birth were not lower than 9; and (f) no neurologic abnormalities up to 6 months of age. In particular, the participants did not have any of the following neurological or metabolic disorders: (a) hypoxic-ischemic encephalopathy; (b) intraventricular hemorrhage or white matter damage on cranial ultrasound; (c) major congenital malformation; (d) central nervous system infection; (e) metabolic disorder; (f) clinical evidence of seizures; and (g) evidence of asphyxia. Informed consent was signed by parents or the legal guardian of the neonates to approve the use of clinical information and NIRS data for scientific purpose. The research was approved by the Ethical Committee of Peking University and the Peking University First Hospital, which specifies a maximal duration of 20 min for neuroimaging experiments performed on neonates. Therefore, each neonate was either assigned to the Chinese or the Portuguese group (i.e., between-subject design), as exposing the neonates to both the Chinese and Portuguese experiments (i.e., within-subject design) would make data collection longer than the maximal duration allowed by Ethics.

### 2.2 | Stimuli

Speech samples in Chinese and Portuguese were selected from the Database of Chinese Vocal Emotions (Liu & Pell, 2012) and the

Database of Portuguese Vocal Emotions (Castro & Lima, 2010). These databases provide “language-like” pseudo-sentences that are constructed by replacing content words with semantically meaningless words while maintaining function words to ensure that the pseudo-sentences are grammatical. Pseudo-sentences had  $9.1 \pm 1.0$  syllables on average in Chinese and  $8.0 \pm 0.9$  syllables on average in Portuguese; they individually lasted between 1 and 2 s each for a total duration of exactly 15 s per affective category (fearful, angry, happy, and neutral) per language. We concatenated 11, 11, 8, and 9 pseudo-sentences of fearful, angry, happy, and neutral prosodies in Chinese, and 13, 11, 10, and 10 pseudo-sentences for the corresponding conditions in Portuguese (see Minagawa-Kawai et al., 2011 and Peña et al., 2003). Speech rates varied between 5.1 and 6.5 syllables per second, but the differences in speech rate between affective conditions and languages were not significant (all  $p > 0.1$ ). Speech samples were produced by one native female speaker of Chinese and one native female speaker of Portuguese. The mean sound intensity was balanced between conditions.

Two independent groups of native Chinese speakers ( $n = 20$ , 10 males, age =  $23.0 \pm 2.1$  years) and native Portuguese speakers ( $n = 17$ , 9 males, age =  $22.3 \pm 1.8$  years) were invited to identify the type of emotion conveyed by the four samples and to rate emotional intensity on a Lickert scale from 1 (weak) to 9 (very strong; see Table 1). Recognition rates from native Chinese speakers were significantly higher for Chinese stimuli than Portuguese stimuli ( $F(1,19) = 156$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.891$ ;  $93.3 \pm 12.0\%$  vs.  $67.0 \pm 24.6\%$ ). There was also a significant interaction between language and emotion on the recognition rate ( $F(3,57) = 26.8$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.585$ ), indicating that while recognition rates were highly comparable across emotion types in Chinese (angry =  $93.6 \pm 15.7\%$ , fearful =  $92.4 \pm 10.2\%$ , happy =  $92.0 \pm 10.8\%$ , and neutral =  $95.0 \pm 11.0\%$ ), they differed significantly between conditions in Portuguese: angry ( $81.6 \pm 11.5\%$ ) and neutral prosodies ( $87.5 \pm 17.7\%$ ) were identified with better accuracy than fearful ( $58.0 \pm 15.7\%$ ;  $ps < 0.001$ ) and happy prosodies ( $41.0 \pm 18.9\%$ ;  $ps < 0.001$ ), and fearful prosody was

**TABLE 1** Recognition rate and emotional intensity of materials rated by native Chinese adults ( $n = 20$ ) and native Portuguese adults ( $n = 17$ )

Material	Recognition rate		Emotional intensity	
	Native Chinese	Native Portuguese	Native Chinese	Native Portuguese
Chinese prosodies				
Fearful	92%	51%	5.40	5.29
Angry	94%	80%	6.60	5.73
Happy	92%	65%	6.05	5.88
Neutral	95%	86%		
Portuguese prosodies				
Fearful	58%	87%	5.15	6.05
Angry	82%	90%	5.95	6.50
Happy	41%	89%	5.70	6.40
Neutral	88%	92%		

The recognition rate was measured by selecting one emotion label from anger, happiness, fear and neutral. The emotional intensity was measured using a 9-point scale.

recognized more accurately than happy prosody ( $p = 0.039$ ). Ratings for the emotional intensity were significantly higher for Chinese than Portuguese ( $F(1,19) = 8.28$ ,  $p = 0.010$ ,  $\eta_p^2 = 0.303$ ;  $6.02 \pm 1.46$  vs.  $5.60 \pm 1.22$ ). There was no interaction between language and emotion type on emotional intensity ( $F < 1$ ).

Recognition rates from native Portuguese speakers were significantly higher for Portuguese stimuli than Chinese ones ( $F(1,16) = 80.3$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.834$ ;  $89.8 \pm 11.6\%$  vs.  $70.6 \pm 19.3\%$ ). There was also a significant interaction between language and emotion on recognition rate ( $F(3,48) = 10.1$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.387$ ), indicating that while recognition rates were comparable across emotion types in Portuguese (angry =  $90.5 \pm 9.7\%$ , fearful =  $87.1 \pm 16.7\%$ , happy =  $89.1 \pm 10.6\%$ , and neutral =  $92.5 \pm 7.8\%$ ), they differed significantly between conditions in Chinese: angry ( $80.4 \pm 11.1\%$ ) and neutral prosodies ( $86.2 \pm 14.7\%$ ) were identified with better accuracy than fearful ( $50.7 \pm 13.7\%$ ;  $ps < 0.001$ ) and happy prosodies ( $65.1 \pm 14.6\%$ ;  $ps \leq 0.005$ ), and happy prosody was recognized more accurately than fearful prosody ( $p = 0.029$ ). Ratings for the emotional intensity were significantly higher for Portuguese than Chinese ( $F(1,16) = 23.8$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.598$ ;  $6.32 \pm 1.28$  vs.  $5.63 \pm 0.99$ ). There was no interaction between language and emotion type on emotional intensity ( $F < 1$ ).

It is important to note that Chinese is a tonal language while Portuguese is not. The current experiment manipulated vocal emotions through variations in speech prosody which, as rated by both native Chinese and native Portuguese speakers, are comparable between the two languages. Lexical tones in Chinese are abstract lexical frames functioning as prosodic cues to distinguish the meaning of Chinese characters rather than emotions. Therefore, the tones of individual Chinese characters would not interfere with the potential emotional effects evoked by speech prosodies.

### 2.3 | Experimental design and procedure

The experiment was conducted in the neonatal ward of Peking University First Hospital, Beijing, China. Sounds were presented passively (see Cheng et al., 2012; Zhang et al., 2014) through a pair of loudspeakers (EDIFIER R26T, Shenzhen, China) placed 10 cm away from the neonates' left and right ears (see Figure 1), at a sound pressure level of 55 to 60 dB. Mean background noise intensity level was 30 dB. Neonates were randomly assigned to the Chinese or the Portuguese version of the experiment and matched in terms of the number of participants and gender. NIRS recording was carried out when the infants were in a quiet state of alert or in a state of natural active sleep (see Cheng et al., 2012; Gómez et al., 2014). We recruited a total number of 30 neonates in the Chinese experiment and 30 in the Portuguese experiment. Neonates who started crying were not included in the analyses, resulting in dataset of 27 (14 boys) participants in the Chinese experiment and 24 (11 boys) in the Portuguese experiment being analyzed. Each 15-s stimulus was repeated 10 times per affective condition, resulting in 40 trials presented in a random order. The interstimulus intervals were silent and varied randomly between 14 and 16 s.



**FIGURE 1** Picture of one neonate with optodes placed upon the head

## 2.4 | NIRS data recording

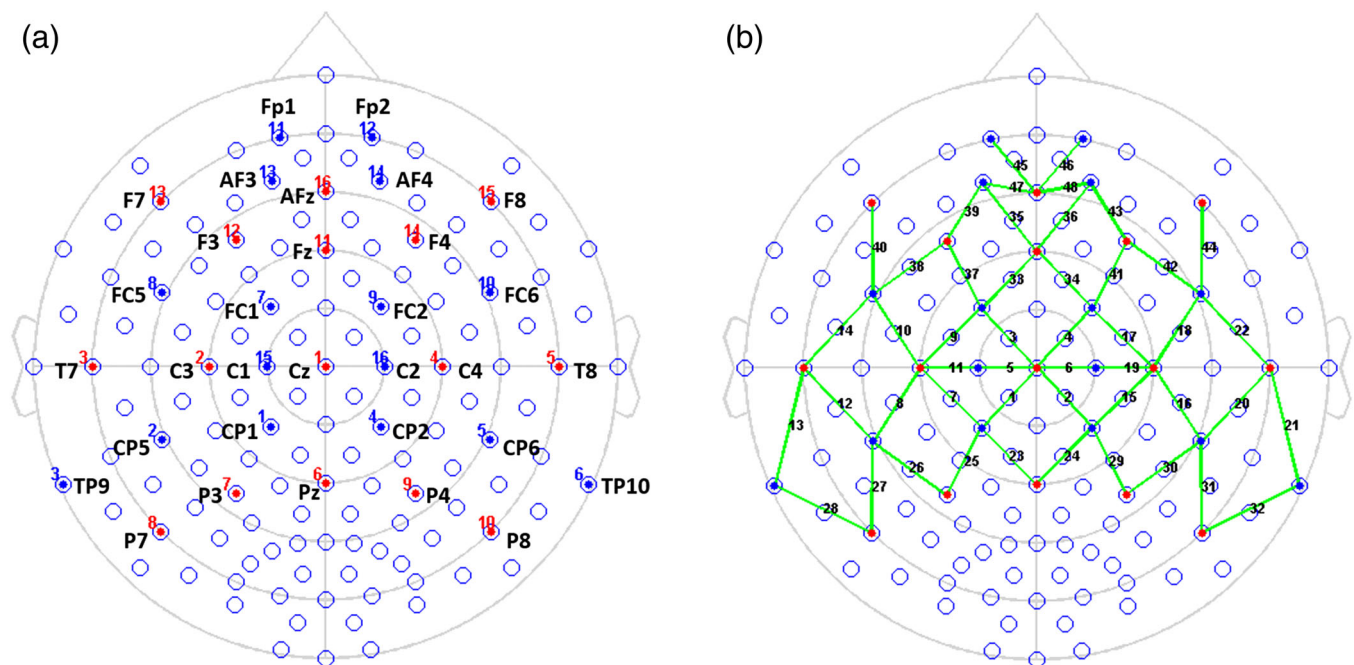
NIRS data were recorded in a continuous-wave mode using the NIRScout 1,624 system (NIRx Medical Technologies, LLC. Los Angeles, CA), which consisted of 16 LED emitters (intensity = 5 mW/wavelength) and 16 detectors at two wavelengths (760 and 850 nm). Based on previous studies in infants (e.g., Benavides-Varela, Gómez, & Mehler, 2011; Cheng et al., 2012; Minagawa-Kawai et al., 2011; Saito et al., 2007; Sato et al., 2012; Taga & Asakawa, 2007; Zhang et al., 2014) and adults (e.g., Brück et al., 2011; Frühholz, Trost, & Kotz, 2016), we placed the optodes over temporal, frontal, and central regions of the brain, using a NIRS-EEG compatible cap of 32 cm diameter (EASYCAP, Herrsching, Germany) in accordance with

the international 10/10 system. There were 48 useful channels (24 per hemisphere), where source and detector were at a mean distance of 2.5 cm (Figure 2; see also Altwater-Mackensen & Grossmann, 2016; Bennett, Bolling, Anderson, Pelphrey, & Kaiser, 2014; Obrig et al., 2017; Quaresima, Bisconti, & Ferrari, 2012; Telkemeyer et al., 2009). The distance between source and detector in each channel is shown in Table 2. The data were recorded continuously at a sampling rate of 4 Hz.

## 2.5 | NIRS data preprocessing, modeling and statistical analyses

The data were processed within the nirsLAB analysis package (v2016.05, NIRx Medical Technologies, LLC.). NIRS data were screened manually. Detector saturation never occurred during any of the recordings. NIRS data contain two main types of artifacts: transient spikes and abrupt discontinuities. First, spikes were manually detected and replaced by linear interpolation with nearest data points. Second, discontinuities (or “jumps”) were automatically detected and corrected using the procedure implemented in nirsLAB (std threshold = 5). Intensity data were then converted into optical density changes ( $\Delta OD$ ), and the  $\Delta OD$  of both measured wavelengths were transformed into relative concentration changes of oxyhemoglobin and deoxyhemoglobin ( $\Delta[HbO]$  and  $\Delta[Hb]$ ) based on the modified Beer-Lambert law (Cope & Delpy, 1988). The differential path length factor was assumed to be 7.25 for a wavelength of 760 nm and 6.38 for a wavelength of 850 nm (Essenpreis et al., 1993).

Statistical evaluation of concentration changes was based on a general linear model of the hemodynamic response function (HRF). The HRF measured in full-term neonates (Arichi et al., 2012) is smaller



**FIGURE 2** Locations of optodes and channels with respect to the EEG 10/10 system. (a) Sixteen LED emitters were placed on positions F7-F8, F3-F4, T7-T8, C3-C4, P7-P8, P3-P4, AFz, Fz, Cz, and Pz (red dots), while sixteen detectors were placed on Fp1-Fp2, AF3-AF4, FC5-FC6, FC1-FC2, C1-C2, TP9-TP10, CP5-CP6, and CP1-CP2 (blue dots). (b) the 16 × 16 optodes constitute 48 channels of interest (green lines) [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

in amplitude and deferred in time-course as compared to that of adults (parameters in nirsLAB = [7 16 1 1 2 0 32]). When estimating beta, the data were pre-whitened with the AR( $n$ ) model ( $1 < n \leq 30$ ), based on the autoregressive iteratively reweighted least-squares method introduced by Barker, Aarabi, and Huppert (2013). The details of the implementation of the algorithm can be found in Huppert (2016). Although both  $\Delta[\text{HbO}]$  and  $\Delta[\text{Hb}]$  were derived, we elected to perform statistical analyses on  $\Delta[\text{HbO}]$  due to its superior sensitivity in the evaluation of functional activity (Sato et al., 2012). When estimating beta, nirsLAB used a SPM-based algorithm (restricted maximum likelihood) to compute a least-squares solution to an overdetermined system of linear equations.

A one-way ANOVA with *emotional category* (neutral, fearful, angry, and happy) as a within-subject variable was performed separately for the Chinese and the Portuguese experiment, to avoid confounds due to cross-language perceptual variances. The tests resulted in two threshold-corrected ( $p < 0.05$ )  $F$ -statistic maps. Follow-up analyses involved pairwise comparisons between the four emotional conditions, focusing on the significant channels revealed by the threshold-corrected  $F$ -statistic maps. The statistical results in individual channels were corrected for multiple comparisons across channels by the false discovery rate (FDR), following the Storey and Tibshirani (2003) procedure implemented in Matlab (v2015b, the Mathworks, Inc., Natick, MA).

## 2.6 | Spatial registration for NIRS channels

NIRS channel locations were defined as the central zone of the light path between each adjacent source-detector pair. To determine the cortical structures underlying NIRS channel positions, a neonate head model (Brigadoi et al., 2014) was used to identify corresponding MNI coordinates of the channel center according to the EEG 10/10 system. We then applied automated anatomical labeling (Tzourio-Mazoyer et al., 2002) to the MNI coordinates based on an infant template (Shi et al., 2011; Table 2).

In addition to the infant head template, we also retrieved corresponding channel locations on a adult head template. For this purpose, a Matlab toolbox NFRI (<http://brain.job.affrc.go.jp/tools/>; Singh, Okamoto, Dan, Jurcak, & Dan, 2005) was used to estimate the NMI coordinates of the channel center. Then the MNI coordinates were transformed into Talairach space (Lancaster et al., 2007; Laird et al., 2010) and mapped onto adult brain atlases (Lancaster et al., 2000; Tzourio-Mazoyer et al., 2002). It is generally assumed in fNIRS studies that macroanatomical structures of infant (Hill et al., 2010; Matsui et al., 2014) and child cortices (Burgund et al., 2002; Watanabe et al., 2013) are similar to those of adults. Therefore, we expect that the probabilistic spatial registration in the adult template (Table 3) would provide meaningful anatomical information for NIRS channels when interpreting neonatal data.

## 3 | RESULTS

In the Chinese version of the experiment, a one-way repeated measure ANOVA with affective prosody (happy, fearful, angry, and

neutral) as the within-group factor showed a main effect of emotional category on fNIRS amplitude on 6 channels (16, 20, 22, 26, 35, and 42). The  $F$ -statistic map using a threshold of  $q < 0.05$  is shown in Figure 3 and  $F$  values are summarized in Table 4. Follow-up analyses involving pairwise comparisons between individual affective prosody conditions and the neutral control condition showed that (a) happy prosody elicited significantly higher activations than neutral prosody on five channels (20, 22, 26, 35, and 42) corresponding to the right temporal and bilateral frontal regions; (b) fearful prosody elicited significantly higher activations than neutral prosody at two channels (16 and 20) corresponding to right temporal and parietal regions; and (c) angry prosody elicited significantly higher activations at channel 20, which corresponds to the right STG (Table 5). Moreover, pairwise comparisons between the three experimental conditions showed that happy prosodies elicited significantly higher activations at channel 22 (the right STG) and channel 42 (the right IFG) than fearful and angry prosodies, respectively. There was no significant difference between fearful and angry prosodies.

In the Portuguese version of the experiment, the one-way repeated measures ANOVA showed a marginal main effect of affective prosody at channel 14 (left STG;  $q = 0.050$ ; see Table 4). Post hoc pairwise comparisons showed that the main effect of emotional category was driven by significantly higher activation for angry than neutral prosody (Table 5). No other significant differences were found when comparing emotional conditions and neutral prosody or emotional conditions with one another ( $q > 0.1$ ). The waveforms of  $\Delta[\text{HbO}]$  and  $\Delta[\text{Hb}]$  at two representative channels (14 and 22) are plotted in Figure 4.

## 4 | DISCUSSION

The current study examined the neural perception of affective prosody in human neonates. Our functional NIRS (fNIRS) findings showed that emotional (i.e., happy, fearful, and angry) prosody, compared to neutral prosody, enhanced activation levels in the posterior portion of the right STG (i.e., Channel 20). Known as the "emotional voice area" (Ethofer et al., 2012), the STG is the essential neural substrate for auditory processing and language development (Anand et al., 2005; Bigler et al., 2007), and has strong implications in social cognition and emotional processing (Goulden et al., 2012; Schaefer, Putnam, Benca, & Davidson, 2006). Adult studies have shown that the anterior (Bach et al., 2008; Fecteau, Belin, Joannette, & Armony, 2007), middle (Ethofer et al., 2009; Wiethoff et al., 2008), and posterior portions of the right STG (Brück et al., 2011; Wittman, Van Heuven, & Schiller, 2012) respond to vocal expressions of emotions. Compared to healthy individuals, STG volumes are larger in children and adolescents with generalized anxiety disorder (De Bellis et al., 2002), and smaller in patients with depression (Fitzgerald, Laird, Maller, & Daskalakis, 2008) and schizophrenia (Honea, Crow, Passingham, & Mackay, 2005; Shenton et al., 1992). Abnormal activities in the right STG during emotional processing of facial expression and speech prosody have been observed in children with autism spectrum disorder (ASD; Kana, Patriquin, Black, Channell, & Wicker, 2016), alexithymia (Goerlich-Dobre et al., 2014; Reker et al., 2010), and social anhedonia (Germine,

**TABLE 2** Spatial registration for NIRS channels (neonatal template)

	Channel	Source-detector distance (cm)	MNI coordinate			Anatomic label (AAL)
			x	y	z	
1	Cz-CP1	2.5	-11	-27	61	Paracentral lobule left
2	Cz-CP2	2.5	11	-27	61	Postcentral gyrus right
3	Cz-FC1	3.0	-9	-4	58	Supplementary motor area left
4	Cz-FC2	3.0	10	-4	58	Supplementary motor area right
5	Cz-C1	2.3	-11	-16	61	Precentral gyrus left
6	Cz-C2	2.3	11	-16	61	Precentral gyrus right
7	C3-CP1	3.5	-30	-27	51	Postcentral gyrus left
8	C3-CP5	2.5	-42	-26	34	Supramarginal gyrus left
9	C3-FC1	3.0	-29	-4	48	Precentral gyrus left
10	C3-FC5	2.5	-40	-6	32	Postcentral gyrus left
11	C3-C1	2.3	-31	-15	52	Precentral gyrus left
12	T7-CP5	2.8	-43	-26	12	Superior temporal gyrus left
13	T7-TP9	2.3	-41	-23	-5	Inferior temporal gyrus left
14	T7-FC5	3.0	-42	-6	10	Superior temporal gyrus left
15	C4-CP2	3.5	31	-28	52	Postcentral gyrus right
16	C4-CP6	2.5	41	-25	35	Supramarginal gyrus right
17	C4-FC2	3.0	29	-4.5	49	Middle frontal gyrus right
18	C4-FC6	2.5	39	-6.4	33	Postcentral gyrus right
19	C4-C2	2.3	30	-16	52	Precentral gyrus right
20	T8-CP6	2.8	43	-25	13	Superior temporal gyrus right
21	T8-TP10	2.3	40	-23	-4	Inferior temporal gyrus right
22	T8-FC6	3.0	41	-6	11	Superior temporal gyrus right
23	Pz-CP1	2.8	-11	-50	55	Precuneus left
24	Pz-CP2	2.8	10	-50	55	Superior parietal gyrus right
25	P3-CP1	3.0	-27	-48	47	Superior parietal gyrus left
26	P3-CP5	2.5	-38	-47	29	Angular gyrus left
27	P7-CP5	2.7	-40	-43	13	Middle temporal gyrus left
28	P7-TP9	3.0	-38	-41	-4	Inferior temporal gyrus left
29	P4-CP2	3.0	26	-48	48	Superior parietal gyrus right
30	P4-CP6	2.5	37	-46	31	Angular gyrus right
31	P8-CP6	2.7	41	-42	14	Middle temporal gyrus right
32	P8-TP10	3.0	37	-40	-3	Inferior temporal gyrus right
33	Fz-FC1	2.8	-10	17	49	Superior frontal gyrus (dorsal) left
34	Fz-FC2	2.8	10	17	49	Superior frontal gyrus (dorsal) right
35	Fz-AF3	2.7	-11	34	32	Superior frontal gyrus (dorsal) left
36	Fz-AF4	2.7	10	33	33	Superior frontal gyrus (dorsal) right
37	F3-FC1	2.8	-25	16	41	Middle frontal gyrus left
38	F3-FC5	2.3	-36	14	25	Inferior frontal gyrus (triangular) left
39	F3-AF3	3.0	-26	32	25	Middle frontal gyrus left
40	F7-FC5	2.3	-39	10	9	Inferior frontal gyrus (triangular) left
41	F4-FC2	2.8	25	15	43	Middle frontal gyrus right
42	F4-FC6	2.3	35	13	27	Inferior frontal gyrus (triangular) right
43	F4-AF4	3.0	25	31	27	Middle frontal gyrus right
44	F8-FC6	2.3	39	10	11	Inferior frontal gyrus (triangular) right
45	AFz-Fp1	2.6	-9	42	13	Superior frontal gyrus (medial) left
46	AFz-Fp2	2.6	8	41	14	Superior frontal gyrus (medial) right
47	AFz-AF3	2.4	-11	41	22	Superior frontal gyrus (dorsal) left
48	AFz-AF4	2.4	10	40	24	Superior frontal gyrus (dorsal) right

A neonate head model (Brigadoi, Aljabar, Kuklisova-Murgasova, Arridge, & Cooper, 2014) was used to find the MNI coordinates of the channel center. Then the MNI coordinates were transformed to brain labels according to an infant AAL template (Shi et al., 2011).

**TABLE 3** Spatial registration for NIRS channels (adult template)

	Channel	Brodmann area (Talairach daemon; percentage of overlap) <sup>a</sup>	AAL (percentage of overlap) <sup>a</sup>
1	Cz-CP1	4–primary motor cortex (0.57) 6–pre-motor and supplementary motor cortex (0.31)	Paracentral_Lobule_L (0.55) Postcentral_L (0.44)
2	Cz-CP2	4–primary motor cortex (0.58) 6–pre-motor and supplementary motor cortex (0.38)	Postcentral_R (0.44) Precentral_R (0.37)
3	Cz-FC1	6–pre-motor and supplementary motor cortex (1)	Frontal_Sup_L (0.57) Supp_Motor_Area_L (0.43)
4	Cz-FC2	6–pre-motor and supplementary motor cortex (1)	Frontal_Sup_R (0.57) Supp_Motor_Area_R (0.43)
5	Cz-C1	6–pre-motor and supplementary motor cortex (1)	Paracentral_Lobule_L (0.40) Precentral_L (0.33)
6	Cz-C2	6–pre-motor and supplementary motor cortex (1)	Frontal_Sup_R (0.41) Precentral_R (0.40)
7	C3-CP1	3–primary somatosensory cortex (0.37)	Postcentral_L (0.91)
8	C3-CP5	40–Supramarginal gyrus, part of Wernicke's area (0.51)	SupraMarginal_L (0.77)
9	C3-FC1	6–pre-motor and supplementary motor cortex (1)	Precentral_L (0.84)
10	C3-FC5	6–pre-motor and supplementary motor cortex (0.96)	Postcentral_L (0.59) Precentral_L (0.41)
11	C3-C1	6–pre-motor and supplementary motor cortex (0.62)	Precentral_L (0.78)
12	T7-CP5	22–superior temporal Gyrus (0.51) 42–primary and auditory association cortex (0.39)	Temporal_Mid_L (0.59) Temporal_Sup_L (0.39)
13	T7-TP9	20–inferior temporal gyrus (0.72)	Temporal_Inf_L (1)
14	T7-FC5	22–superior temporal Gyrus (0.64)	Temporal_Sup_L (0.49) Rolandic_Oper_L (0.32)
15	C4-CP2	3–primary somatosensory cortex (0.37)	Postcentral_R (0.82)
16	C4-CP6	40–supramarginal gyrus part of Wernicke's area (0.53)	SupraMarginal_R (0.99)
17	C4-FC2	6–pre-motor and supplementary motor cortex (1)	Frontal_Mid_R (0.81)
18	C4-FC6	6–pre-motor and supplementary motor cortex (0.97)	Postcentral_R (0.56) Precentral_R (0.44)
19	C4-C2	6–pre-motor and supplementary motor cortex (0.64)	Precentral_R (0.86)
20	T8-CP6	22–superior temporal gyrus (0.62) 42–primary and auditory association cortex (0.37)	Temporal_Sup_R (0.80)
21	T8-TP10	20–inferior temporal gyrus (0.76)	Temporal_Inf_R (0.94)
22	T8-FC6	22–superior temporal gyrus (0.63)	Temporal_Sup_R (0.68)
23	Pz-CP1	7–somatosensory association cortex (1)	Precuneus_L (0.60) Parietal_Sup_L (0.40)
24	Pz-CP2	7–somatosensory association cortex (1)	Parietal_Sup_R (0.67) Precuneus_R (0.30)
25	P3-CP1	7–somatosensory association cortex (0.59) 40–supramarginal gyrus, part of Wernicke's area (0.36)	Parietal_Sup_L (0.60) Parietal_Inf_L (0.40)
26	P3-CP5	40–supramarginal gyrus, part of Wernicke's area (0.86)	Angular_L (0.52) Parietal_Inf_L (0.39)
27	P7-CP5	22–superior temporal gyrus (0.59)	Temporal_Mid_L (0.77)
28	P7-TP9	37–fusiform gyrus (0.65) 20–inferior temporal gyrus (0.32)	Temporal_Inf_L (0.79)
29	P4-CP2	7–somatosensory association cortex (0.83)	Parietal_Sup_R (0.84)
30	P4-CP6	40–supramarginal gyrus, part of Wernicke's area (0.81)	Angular_R (0.50) Parietal_Inf_R (0.50)
31	P8-CP6	22–superior temporal gyrus (0.62)	Temporal_Mid_R (0.79)
32	P8-TP10	37–fusiform gyrus (0.65) 20–inferior temporal gyrus (0.30)	Temporal_Inf_R (0.80)

(Continues)

TABLE 3 (Continued)

	Channel	Brodman area (Talairach daemon; percentage of overlap) <sup>a</sup>	AAL (percentage of overlap) <sup>a</sup>
33	Fz-FC1	8–includes frontal eye fields (0.54) 6–pre-motor and supplementary motor cortex (0.46)	Frontal_Sup_L (0.64) Frontal_Sup_Medial_L (0.35)
34	Fz-FC2	8–includes frontal eye fields (0.52) 6–pre-motor and supplementary motor cortex (0.48)	Frontal_Sup_R (0.52) Frontal_Sup_Medial_R (0.45)
35	Fz-AF3	9–dorsolateral prefrontal cortex (0.72)	Frontal_Sup_L (0.83)
36	Fz-AF4	9–dorsolateral prefrontal cortex (0.83)	Frontal_Sup_R (0.86)
37	F3-FC1	8–includes frontal eye fields (0.98)	Frontal_Mid_L (1)
38	F3-FC5	46–dorsolateral prefrontal cortex (0.76)	Frontal_Inf_Tri_L (0.81)
39	F3-AF3	10–frontopolar area (0.94)	Frontal_Mid_L (0.87)
40	F7-FC5	45–pars triangularis, part of Broca's area (0.61) 47–inferior prefrontal gyrus (0.33)	Frontal_Inf_Tri_L (0.91)
41	F4-FC2	8–includes frontal eye fields (1)	Frontal_Mid_R (0.89)
42	F4-FC6	46–dorsolateral prefrontal cortex (0.58)	Frontal_Inf_Tri_R (0.70)
43	F4-AF4	10–frontopolar area (0.89)	Frontal_Mid_R (0.98)
44	F8-FC6	45–pars triangularis, part of Broca's area (0.64) 47–inferior prefrontal gyrus (0.30)	Frontal_Inf_Tri_R (0.73)
45	AFz-Fp1	10–frontopolar area (1)	Frontal_Sup_Medial_L (0.53) Frontal_Sup_L (0.47)
46	AFz-Fp2	10–frontopolar area (1)	Frontal_Sup_Medial_R (0.65) Frontal_Sup_R (0.35)
47	AFz-AF3	10–frontopolar area (1)	Frontal_Sup_L (0.67) Frontal_Sup_Medial_L (0.33)
48	AFz-AF4	10–frontopolar area (1)	Frontal_Sup_R (0.61) Frontal_Sup_Medial_R (0.35)

The matlab toolbox NFRI (Singh et al., 2005) was used to estimate the NMI coordinates of the channel center and transform the coordinates to brain labels.

<sup>a</sup> One NIRS channel may be associated with several Brodman areas. For the sake of brevity, here we only report the Brodman areas with a percentage of overlap >0.30.

Garrido, Bruce, & Hooker, 2011). In 7-month-old normally developed infants, happy and angry prosodies increased neural activation in the right superior temporal cortex (STC) as compared to neutral prosody (Grossmann et al., 2010). The current finding, therefore, shows that the right STG is critically involved in emotion processing from birth in humans. Given that the right STG is classically associated with pitch processing (Boemio, Fromm, Braun, & Poeppel, 2005; Patterson, Uppenkamp, Johnsrude, & Griffiths, 2002; Zatorre & Belin, 2001), emotional and neutral prosody might be discriminated through neural responses to pitch variations, since emotional prosody has a distinctively higher pitch as compared to neutral prosody (see acoustic measures of the materials in Liu & Pell, 2012).

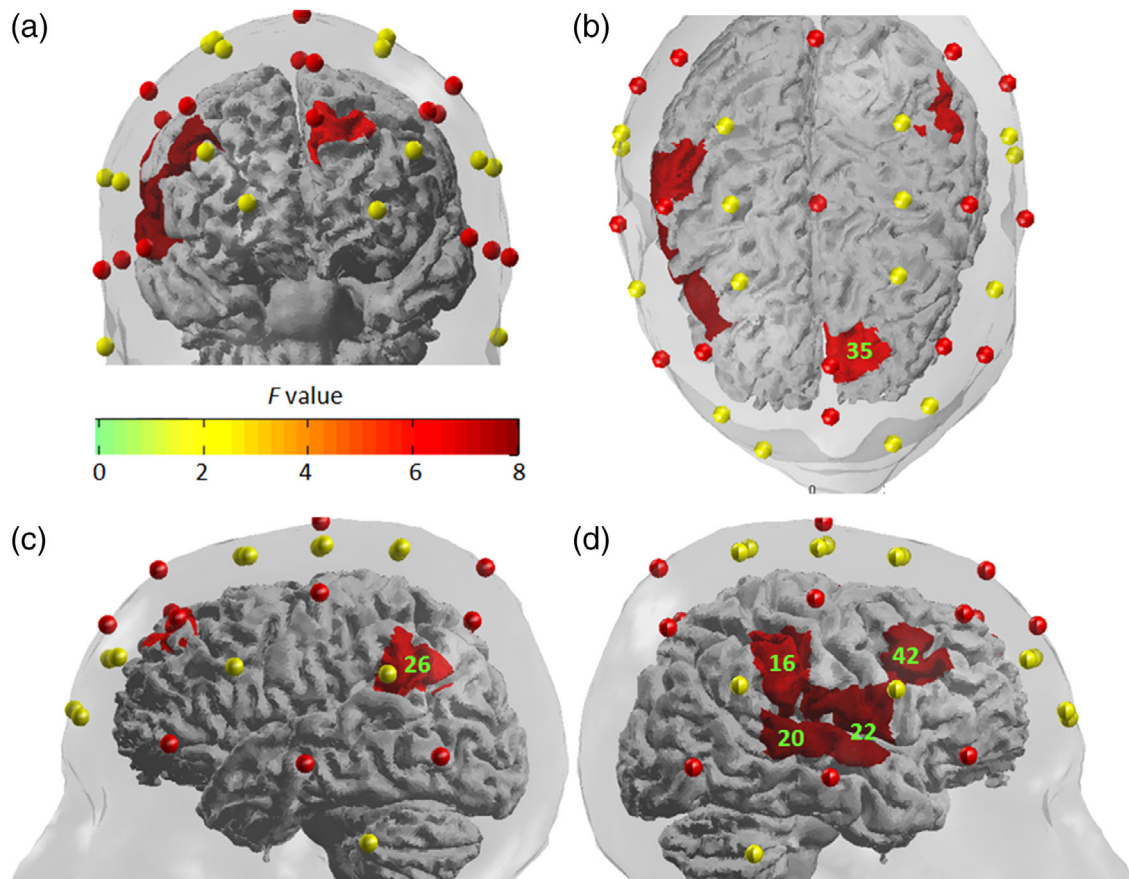
#### 4.1 | Unique neural characteristics in neonates

Interestingly, activities in the right STS failed to discriminate any category of emotional prosody from the baseline (i.e., neutral prosody) in neonates, while adult studies have established its involvement in the processing of vocal emotions (Beaucousin et al., 2007; Belin et al., 2000; Ethofer et al., 2009; Fecteau et al., 2007; Grandjean et al., 2005; Kreifelts, Ethofer, Huberle, Grodd, & Wildgruber, 2010; Sander et al., 2005; Wildgruber, Ackermann, Kreifelts, & Ethofer, 2006). This discrepancy suggests that the right STS may mature at a later stage in life and is involved in the processing of more complex aspects of

affective speech prosody (i.e., beyond pitch perception). It is also possible, however, that fNIRS lacks the sensitivity to detect hemodynamic changes in brain sulci to the same extent as in the cortex surface (Lloyd-Fox et al., 2014).

Unlike the emotional effects observed in the right STG, an area of the brain that responds to all three affective prosody conditions, happy, but not fearful or angry prosody, activated frontal regions relative to the neutral prosody. The frontal regions included the left superior frontal gyrus (SFG, part of the dorsalis), left angular gyrus (AG), and right IFG (part of the triangularis). The right IFG has been implicated in the evaluation and modulation of emotion as an integration hub of the underlying neural network (Kirby & Robinson, 2017; Wager, Davidson, Hughes, Lindquist, & Ochsner, 2008); in adults, disrupted connectivity between the amygdala and the right IFG has been associated with mood disorders (Johnstone, van Reekum, Urry, Kalin, & Davidson, 2007; Townsend et al., 2013); in children with ASD, reduced IFG activation has been found to be associated with impaired social functioning (Dapretto et al., 2006; Kita et al., 2011). Grossmann et al. (2010) showed enhanced activation in the right IFG for happy prosody, as compared to neutral prosody, in 7-month-old infants. The current finding, therefore, suggests that the involvement of the right IFG in the processing of affective prosody is established in the very beginning of life.





**FIGURE 3** The  $F$ -statistic map showing brain regions that had significantly different activations among the four conditions (neutral, fearful, angry, and happy prosody). Reported  $F$  values are thresholded by  $q < 0.05$  (corrected for multiple comparisons using FDR). (a) Front view. (b) Top view. (c) Left view. (d) Right view. Green labels denote the number of channels [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

Further, studies on adults and infants have not observed activations in the left SFG and AG during the perception of vocal emotions, despite evidence of the involvement of these brain areas in language processing (Nair et al., 2015; Price, 2010), such as humor comprehension (Azim, Mobbs, Jo, Menon, & Reiss, 2005; Chan et al., 2013; Mobbs, Greicius, Abdel-Azim, Menon, & Reiss, 2003) and expression (Fried, Wilson, MacDonald, & Behnke, 1998). One possibility is that the left SFG and AG constitute the primitive neural mechanisms for processing affective (i.e., happy) prosody, and the function specificity in higher-level verbal emotion is the result of neural specialization of the language network.

Fearful, but not happy or angry prosody, activated the right supramarginal gyrus (SMG) from the baseline. The right SMG has been implicated in phonological processing (Hartwigsen et al., 2010; McDermott, Petersen, Watson, & Ojemann, 2003), emotional perception (Adolphs, Damasio, Tranel, Cooper, & Damasio, 2000), and empathic judgments (Silani, Lamm, Ruff, & Singer, 2013). Interestingly, while the right SMG has been shown to be sensitive to highly aroused vocal stimuli (Aryani, Hsu, & Jacobs, 2018), Köchel, Schöngassner, and Schienle (2013) found that nonverbal, fearful sounds (e.g., screams of fear and pain) enhanced activation levels in the bilateral SMG in adults. During the prenatal stage, neonates are less likely to have

**TABLE 4** fNIRS channels (indicated by one-way ANOVA) showing different activation patterns across experimental conditions (neutral, fearful, angry, and happy prosody)

Channel	Anatomic label (AAL)	$F$ value	$p$ value	$q$ value <sup>a</sup>
Chinese prosodies				
16 C4-CP6	Supramarginal gyrus right	6.75	0.015	0.043
20 T8-CP6	Superior temporal gyrus right	8.44	0.007	0.038
22 T8-FC6	Superior temporal gyrus right	9.20	0.005	0.038
26 P3-CP5	Angular gyrus left	6.54	0.017	0.043
35 Fz-AF3	Superior frontal gyrus (dorsal) left	6.11	0.020	0.043
42 F4-FC6	Inferior frontal gyrus (triangular) right	7.98	0.009	0.038
Portuguese prosodies				
14 T7-FC5	Superior temporal gyrus left	8.70	0.007	0.050

<sup>a</sup> Produced by FDR procedure ( $q$  threshold = 0.05).

**TABLE 5** Follow-up pairwise comparisons between different emotions

Channel	AAL	t value	p value	q value*
Chinese prosodies				
Happiness > neutral				
20 T8-CP6	Superior temporal gyrus right	3.40	0.002	0.016
22 T8-FC6	Superior temporal gyrus right	3.86	<0.001	0.010
26 P3-CP5	Angular gyrus left	3.01	0.006	0.026
35 Fz-AF3	Superior frontal gyrus (dorsal) left	2.99	0.006	0.026
42 F4-FC6	Inferior frontal gyrus (triangular) right	3.72	<0.001	0.010
Fear > neutral				
16 C4-CP6	Supramarginal gyrus right	3.63	0.001	0.032
20 T8-CP6	Superior temporal gyrus right	3.28	0.003	0.039
Anger > neutral				
20 T8-CP6	Superior temporal gyrus right	3.28	0.003	0.035
Happiness > fear				
22 T8-FC6	Superior temporal gyrus right	3.79	<0.001	0.039
Happiness > anger				
42 F4-FC6	Inferior frontal gyrus (triangular) right	3.72	<0.001	0.046
Portuguese prosodies				
Anger > neutral				
14 T7-FC5	Superior temporal gyrus left	3.56	0.002	0.046

\*produced by FDR procedure ( $q$  threshold = 0.05).

experience with nonverbal emotional sounds in comparison to speech prosody. Therefore, we speculated that the right SMG is the neural substrate of fear perception starting at birth, and that experience with a greater range of affective stimuli might induce neuroplastic changes leading to bilateral activation of the SMG in adulthood.

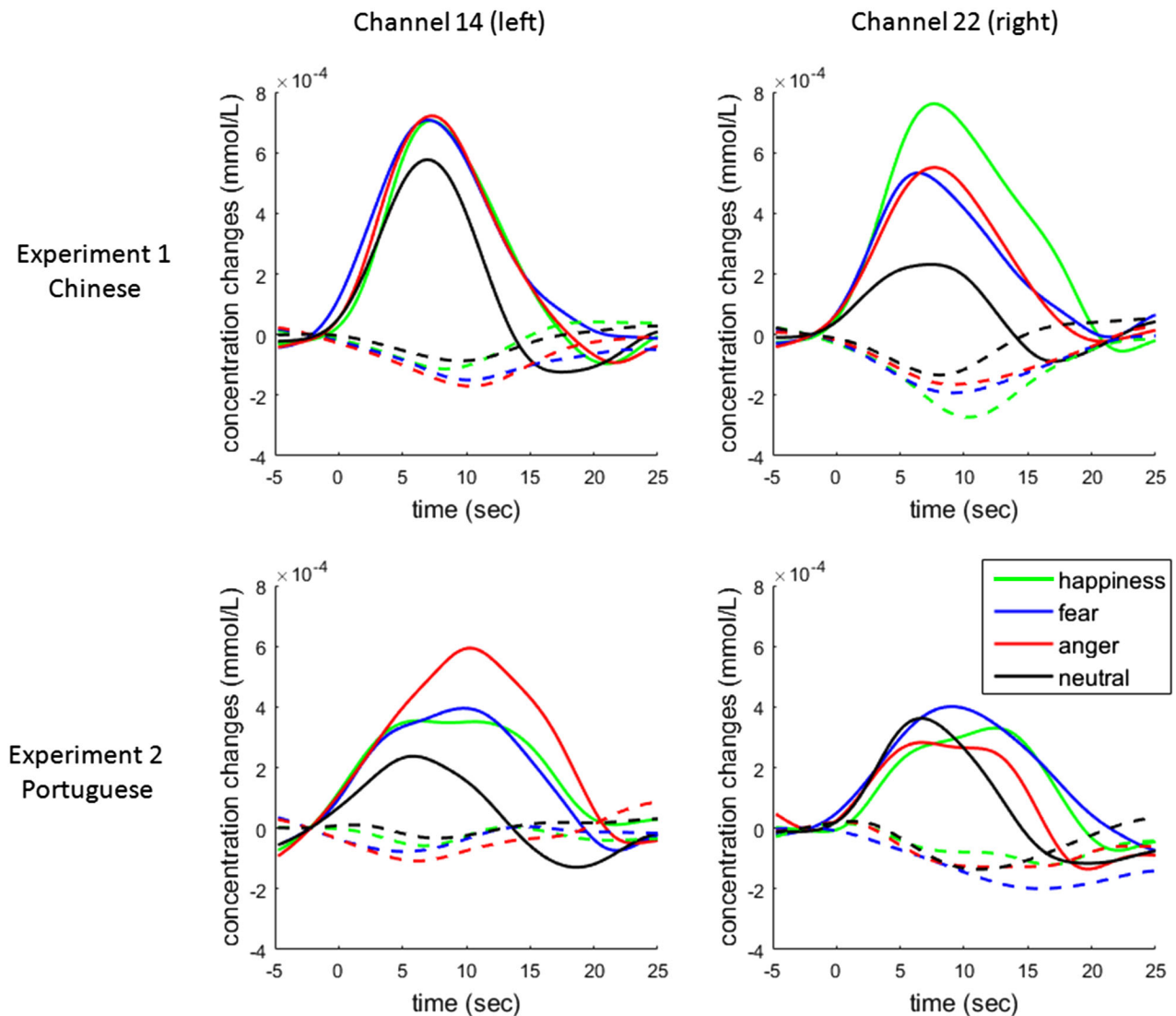
## 4.2 | Preferences for positive prosody

Comparisons between positive and negative prosodies revealed an early preference for positive emotion in the neonate's brain; relative to neutral prosody, happy prosody activated a more widespread neural network as compared to either fearful or angry prosody (Table 5). Moreover, happy prosody elicited increased neural responses in the right STG (middle portion) and the right IFG as compared to fearful and angry prosodies. According to the hierarchical model of prosodic emotion comprehension (Bach et al., 2008; Brück et al., 2011; Ethofer et al., 2006), prosodic information is extracted by voice-sensitive brain structures in the auditory cortex and the middle STC (including STG, STS, and the middle temporal gyrus). While the right STC is associated with the identification of emotional categories, the right IFG appears to be involved in in-depth processing and detailed evaluation of vocally expressed emotions (Schirmer & Kotz, 2006). Therefore, our findings suggest that the extraction of voice features in happy prosody involves a more complex neural integration process than fearful prosody, indicating preferential processing for positive prosody at the earliest phase of vocal perception.

In addition, activation levels of the IFG (Channel 42) suggest that the evaluation of the emotional contents elicited more in-depth processing for happy as compared to angry prosody, a pattern of variation consistent with the literature on adults. For instance, activation levels in the right IFG have been shown to differentiate perception of the vocal emotions of happiness, anger, sadness, and surprise (Kotz,

Kalberlah, Bahlmann, Friederici, & Haynes, 2013). Positive stimuli elicited greater activation than negative stimuli in the right IFG in healthy adults following administration of antidepressants (Norbury, Mackay, Cowen, Goodwin, & Harmer, 2008).

The current study, however, did not find evidence for a negative bias as reported by Cheng et al. (2012). When compared to happy prosody, fearful, and angry prosodies failed to increase neural activity in any of the brain areas covered by fNIRS channels. The discrepancy might be due to differences in the stimuli used (syllables vs. segments of sentences) or the measure of interest used (EEG vs. NIRS) between studies. However, preferential responses to positive stimuli, rather than negative bias, might have an evolutionary advantage in the very beginning of human life. As shown by Mastropieri and Turkewitz (1999), happy prosodies attract more attention than negative (i.e., angry and sad) prosodies and induce more eye-opening responses in neonates. Presumably, it is more critical for neonates to learn to associate positive cues with social care than negative cues with potential threats (see Kirita, & Mitsuo, 1995 for the phenomenon of the "happy face advantage" and Zhang et al., 2018 for the latest evidence at behavioral and neuroimaging levels). Processing priority to positive information is not only because the level of threat for a neonate is rather limited, but also because their attitude is naturally passive; orientation to the positive stimulus might result in more or better care, but bias toward negative stimuli offers no obvious advantage. Negative bias as an adaptive behavior becomes more important when infants start to actively explore the world around them and actively memorize painful experiences as they develop their ability to read signs of danger. This result is, therefore, not inconsistent with the prevailing view that negative bias is characteristic of 6-month-olds and over (e.g., Grossmann et al., 2005, 2010; Hoehl & Striano, 2010; Peltola et al., 2009).



**FIGURE 4** The time-course of  $\Delta[\text{HbO}]$  and  $\Delta[\text{Hb}]$  in response to the four prosodies in two experiments. The four subplots display the waveforms at Channels 14 and 22 (located in left and right superior temporal gyrus). Full lines indicate the waveforms of  $\Delta[\text{HbO}]$  while dashed lines indicate the waveforms of  $\Delta[\text{Hb}]$  [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

It is important to note that the speech prosody used in the current study included one positive (happy) and two negative (fearful and angry) categories. This imbalance is due to the limited choices of emotional categories of neonates, who are most likely to perceive and respond to basic emotions such as happy, fearful, and angry. To examine the negative bias and positive preference hypothesis, it is recommended that future studies balance positive and negative prosodic emotions to avoid possible confounding effects due to a mismatch between them.

### 4.3 | Language-dependency of prosodic emotional processing

The results suggest that the perception of prosodic emotions may be language-dependent in neonates, as they distinguish three types of emotional prosody in Chinese, but only angry from neutral prosody in Portuguese. The native (i.e., Chinese) and foreign (i.e., Portuguese) language are defined as the language to which the mother was exposed or not exposed during pregnancy, suggesting that differential

responses to emotional prosody between languages are likely to arise from prenatal exposure to prosodic variations in a familiar language (Kisilevsky et al., 2009; Mastropieri & Turkewitz, 1999; Sato et al., 2012). These findings are consistent with previous studies showing preferential processing of prosodic emotions in a native language as compared to foreign languages (e.g., Mastropieri & Turkewitz, 1999) and are also in line with adult studies on voice identification (Goggin, Strube, & Simental, 1991; Perrachione & Wong, 2007; Thompson, 1987). Other studies have also shown that prosodic cues and intonation patterns attracted greater attention from newborns when presented in the native as compared to foreign languages (Mehler et al., 1988; Moon et al., 1993). Future studies will further examine the language-dependency of processing prosodic emotions in neonates with more diversity of language background.

According to Mastropieri and Turkewitz (1999), the prenatal learning of emotional perception in speech is achieved through *intra-uterio* association of acoustic properties of speech with maternal physiological changes. Contrary to the current findings, none of the

emotional prosodies in Mastropieri and Turkewitz (1999), including angry prosody, evoked significant eye-opening responses when presented in a foreign language. One possibility is that the distinctive response to angry prosody observed here is due to overlaps in prosodic features between Chinese and Portuguese. Indeed, monolingual Chinese adults tested on Portuguese speech samples had relatively higher recognition rates for angry than fearful and happy emotional prosodies (see Table 1). Another possible reason for the divergent finding is that fNIRS is a more sensitive measurement as compared to eye-opening responses used in Mastropieri and Turkewitz (1999).

Another interesting observation is that angry prosody differentially activated the right STG when presented in Chinese and the left STG when presented in Portuguese. Research on speech perception has shown that left temporal cortices are most responsive to rapid spectral transition during phonemic perception. Further, the relatively gradual changes in prosody tend to elicit right STG activations in neonates (Arimitsu et al., 2011; Peña et al., 2003; Telkemeyer et al., 2009), infants (Dehaene-Lambertz, Dehaene, & Hertz-Pannier, 2002; Minagawa-Kawai, Mori, Naoi, & Kojima, 2007), and adults (Boemio et al., 2005; Grandjean et al., 2005; Wiethoff et al., 2008). Neural activations in the left STG, therefore, are likely to be associated with phonemic variations rather than responses to the affective contents of speech prosody in Portuguese. There are 32 phonemes in Chinese and 37 phonemes in Portuguese. The angry prosody used in the Chinese experiment contained 248 phonemes, but 279 in the Portuguese experiment. As an alternative explanation, studies on adults have shown that pitch processing in the native language activates the right STG, while activations in the left STG are specific to extracting pitch patterns in a foreign language (Wang, Sereno, Jongman, & Hirsch, 2003; Wong, Perrachione, & Parrish, 2007). This explanation would be consistent with our view that pitch processing is the critical component for emotional perception of speech prosody in neonates.

## 5 | CONCLUSION

In sum, the current study provides the first neurophysiological evidence for selective affective prosody processing in human neonates. Consistent with observations in infants and adults, emotional prosodies, as compared to neutral prosody, enhanced activation levels in the right STG in neonates. However, emotional prosody failed to activate the STS, a key brain region for infants and adults, but raised activation levels in the left SFG and the left angular gyrus, specifying neural reorganization processes involved in the perception of emotional prosody during neonatal development. While there is no evidence for a negativity bias, distinct neural responses to happy prosody were observed in the right IFG, suggesting that a preference for positive stimuli might be a critical evolutionary characteristic of the neonate brain. Finally, neonates mainly discriminated emotional from neutral prosody when presented within the native, but not foreign speech, indicating that affective prosody is acquired through prenatal associative learning in the language context of the family. In our view, these findings may help improve understanding of neurodevelopmental disorders such as autism, which involves impaired perception of emotional tones in speech (Hobson et al., 1989; Rutherford,

Baron-Cohen, & Wheelwright, 2002; Van Lancker, Cornelius, & Kreiman, 1989). One limitation of the current study is that the cognitive (i.e., positive preferences) and neural characteristics involved in the perception of emotional prosody were examined at only one point in time (i.e., 0–4 days). While neural mechanistic differences between neonates examined in the current study and older participants (i.e., infants and adults) from previous studies suggest that critical specializations might take place during early infancy, it is necessary that future studies adopt a longitudinal approach to better reveal the development of neural substrates in this process.

## CONFLICT OF INTEREST

The Authors declare no conflicts of interest in relation to the subject of this study.

## ORCID

Dandan Zhang  <https://orcid.org/0000-0003-1825-7114>

## REFERENCES

- Adolphs, R., Damasio, H., Tranel, D., Cooper, G., & Damasio, A. R. (2000). A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping. *The Journal of Neuroscience*, *20*, 2683–2690.
- Altwater-Mackensen, N., & Grossmann, T. (2016). The role of left inferior frontal cortex during audiovisual speech perception in infants. *NeuroImage*, *133*, 14–20.
- Anand, A., Li, Y., Wang, Y., Wu, J., Gao, S., Bukhari, L., ... Lowe, M. J. (2005). Activity and connectivity of brain mood regulating circuit in depression: A functional magnetic resonance study. *Biological Psychiatry*, *57*, 1079–1088.
- Arichi, T., Fagiolo, G., Varela, M., Melendez-Calderon, A., Allievi, A., Merchant, N., ... Edwards, A. D. (2012). Development of BOLD signal hemodynamic responses in the human brain. *NeuroImage*, *63*, 663–673.
- Arimitsu, T., Uchida-Ota, M., Yagihashi, T., Kojima, S., Watanabe, S., Hokuto, I., et al. (2011). Functional hemispheric specialization in processing phonemic and prosodic auditory changes in neonates. *Frontiers in Psychology*, *2*, 202.
- Aryani, A., Hsu, C. T., & Jacobs, A. M. (2018). The sound of words evokes affective brain responses. *Brain Sciences*, *8*, E94.
- Azim, E., Mobbs, D., Jo, B., Menon, V., & Reiss, A. L. (2005). Sex differences in brain activation elicited by humor. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 16496–16501.
- Bach, D. R., Grandjean, D., Sander, D., Herdener, M., Strik, W. K., & Seifritz, E. (2008). The effect of appraisal level on processing of emotional prosody in meaningless speech. *NeuroImage*, *42*, 919–927.
- Barker, J. W., Aarabi, A., & Huppert, T. J. (2013). Autoregressive model based algorithm for correcting motion and serially correlated errors in fNIRS. *Biomedical Optics Express*, *4*, 1366–1379.
- Belyk, M., & Brown, S. (2014). Perception of affective and linguistic prosody: An ALE meta-analysis of neuroimaging studies. *Social Cognitive and Affective Neuroscience*, *9*, 1395–1403.
- Benavides-Varela, S., Gómez, D. M., & Mehler, J. (2011). Studying neonates' language and memory capacities with functional near-infrared spectroscopy. *Frontiers in Psychology*, *2*, 64.
- Beaucousin, V., Lacheret, A., Turbelin, M. R., Morel, M., Mazoyer, B., & Tzourio-Mazoyer, N. (2007). fMRI study of emotional speech comprehension. *Cerebral Cortex*, *17*, 339–352.
- Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, *403*, 309–312.
- Bennett, R. H., Bolling, D. Z., Anderson, L. C., Pelphrey, K. A., & Kaiser, M. D. (2014). fNIRS detects temporal lobe response to affective touch. *Social Cognitive and Affective Neuroscience*, *9*, 470–476.

- Bigler, E. D., Mortensen, S., Neeley, E. S., Ozonoff, S., Krasny, L., Johnson, M., ... Lainhart, J. E. (2007). Superior temporal gyrus, language function, and autism. *Developmental Neuropsychology*, 31, 217–238.
- 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, 1220–1224.
- Boemio, A., Fromm, S., Braun, A., & Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nature Neuroscience*, 8, 389–395.
- Brigadói, S., Aljabar, P., Kuklisova-Murgasova, M., Arridge, S. R., & Cooper, R. J. (2014). A 4D neonatal head model for diffuse optical imaging of pre-term to term infants. *NeuroImage*, 100, 385–394.
- Brück, C., Kreifelts, B., & Wildgruber, D. (2011). Emotional voices in context: A neurobiological model of multimodal affective information processing. *Physics of Life Reviews*, 8, 383–403.
- Burgund, E. D., Kang, H. C., Kelly, J. E., Buckner, R. L., Snyder, A. Z., Petersen, S. E., & Schlaggar, B. L. (2002). The feasibility of a common stereotactic space for children and adults in fMRI studies of development. *NeuroImage*, 17, 184–200.
- Castro, S. L., & Lima, C. F. (2010). Recognizing emotions in spoken language: A validated set of Portuguese sentences and pseudosentences for research on emotional prosody. *Behavior Research Methods*, 42, 74–81.
- Chan, Y. C., Chou, T. L., Chen, H. C., Yeh, Y. C., Lavalley, J. P., Liang, K. C., & Chang, K. E. (2013). Towards a neural circuit model of verbal humor processing: An fMRI study of the neural substrates of incongruity detection and resolution. *NeuroImage*, 66, 169–176.
- Cheng, Y., Lee, S. Y., Chen, H. Y., Wang, P. Y., & Decety, J. (2012). Voice and emotion processing in the human neonatal brain. *Journal of Cognitive Neuroscience*, 24, 1411–1419.
- Cope, M., & Delpy, D. T. (1988). System for long-term measurement of cerebral blood and tissue oxygenation on newborn infants by near infra-red transillumination. *Medical & Biological Engineering & Computing*, 26, 289–294.
- Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. A., Sigman, M., Bookheimer, S. Y., & Iacoboni, M. (2006). Understanding emotions in others: Mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience*, 9, 28–30.
- De Bellis, M. D., Keshavan, M. S., Shifflett, H., Iyengar, S., Dahl, R. E., Axelson, D. A., et al. (2002). Superior temporal gyrus volumes in pediatric generalized anxiety disorder. *Biological Psychiatry*, 51, 553–562.
- Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science*, 298, 2013–2015.
- Ekman, P., Friesen, W., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., et al. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology*, 53, 712–717.
- Essenpreis, M., Elwell, C. E., Cope, M., van der Zee, P., Arridge, S. R., & Delpy, D. T. (1993). Spectral dependence of temporal point spread functions in human tissues. *Applied Optics*, 32, 418–425.
- 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. *Cerebral Cortex*, 22, 191–200.
- Ethofer, T., Kreifelts, B., Wiethoff, S., Wolf, J., Grodd, W., Vuilleumier, P., & Wildgruber, D. (2009). Differential influences of emotion, task, and novelty on brain regions underlying the processing of speech melody. *Journal of Cognitive Neuroscience*, 21, 1255–1268.
- 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, 580–587.
- Farroni, T., Menon, E., Rigato, S., & Johnson, M. H. (2007). The perception of facial expressions in newborns. *European Journal of Developmental Psychology*, 4, 2–13.
- Fecteau, S., Belin, P., Joanette, Y., & Armony, J. L. (2007). Amygdala responses to nonlinguistic emotional vocalizations. *NeuroImage*, 36, 480–487.
- Fernald, A. (1993). Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development*, 64, 657–674.
- Fitzgerald, P. B., Laird, A. R., Maller, J., & Daskalakis, Z. J. (2008). A meta-analytic study of changes in brain activation in depression. *Hum. Brain Mapp.*, 29, 683–695.
- 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, 238–252.
- Fried, I., Wilson, C., MacDonald, K., & Behnke, E. (1998). Electric current stimulates laughter. *Nature*, 391, 650.
- Frühholz, S., & Grandjean, D. (2013a). Multiple subregions in superior temporal cortex are differentially sensitive to vocal expressions: A quantitative meta-analysis. *Neuroscience and Biobehavioral Reviews*, 37, 24–35.
- Frühholz, S., Trost, W., & Kotz, S. A. (2016). The sound of emotions - towards a unifying neural network perspective of affective sound processing. *Neuroscience and Biobehavioral Reviews*, 68, 96–110.
- Germiné, L. T., Garrido, L., Bruce, L., & Hooker, C. (2011). Social anhedonia is associated with neural abnormalities during face emotion processing. *NeuroImage*, 58, 935–945.
- Goerlich-Dobre, K. S., Wittman, J., Schiller, N. O., van Heuven, V. J., Aleman, A., & Martens, S. (2014). Blunted feelings: Alexithymia is associated with a diminished neural response to speech prosody. *Social Cognitive and Affective Neuroscience*, 9, 1108–1117.
- Goggin, J. P., Strube, G., & Simental, L. R. (1991). The role of language familiarity in voice identification. *Memory & Cognition*, 19, 448–458.
- Gómez, D. M., Berent, I., Benavides-Varela, S., Bion, R. A., Cattarossi, L., Nespor, M., et al. (2014). Language universals at birth. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 5837–5841.
- Goulden, N., McKie, S., Thomas, E. J., Downey, D., Juhasz, G., Williams, S. R., ... Elliott, R. (2012). Reversed frontotemporal connectivity during emotional face processing in remitted depression. *Biological Psychiatry*, 72, 604–611.
- 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, 145–146.
- Grossmann, T., Oberecker, R., Koch, S. P., & Friederici, A. D. (2010). The developmental origins of voice processing in the human brain. *Neuron*, 65, 852–858.
- Grossmann, T., Striano, T., & Friederici, A. D. (2005). Infants' electric brain responses to emotional prosody. *Neuroreport*, 16, 1825–1828.
- Hartwiges, G., Baumgaertner, A., Price, C. J., Koehnke, M., Ulmer, S., & Siebner, H. R. (2010). Phonological decisions require both the left and right supramarginal gyri. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 16494–16499.
- Häuser, K., & Domahs, F. (2014). Functional lateralization of lexical stress representation: A systematic review of patient data. *Frontiers in Psychology*, 5, 317.
- Hill, J., Dierker, D., Neil, J., Inder, T., Knutsen, A., Harwell, J., ... van Essen, D. (2010). A surface-based analysis of hemispheric asymmetries and folding of cerebral cortex in term-born human infants. *The Journal of Neuroscience*, 30, 2268–2276.
- Hobson, R. P., Ouston, J., & Lee, A. (1988). Emotion recognition in autism: coordinating faces and voices. *Psychol Med.*, 18, 911–923.
- Hoehl, S., & Striano, T. (2010). The development of emotional face and eye gaze processing. *Developmental Science*, 13, 813–825.
- Honea, R., Crow, T. J., Passingham, D., & Mackay, C. E. (2005). Regional deficits in brain volume in schizophrenia: a meta-analysis of voxel-based morphometry studies. *Am J Psychiatry*, 162, 2233–2245.
- Huppert, T. J. (2016). Commentary on the statistical properties of noise and its implication on general linear models in functional near-infrared spectroscopy. *Neurophotonics*, 3, 010401.
- 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, 887–900.
- Johnstone, T., van Reekum, C. M., Urry, H. L., Kalin, N. H., & Davidson, R. J. (2007). Failure to regulate: Counterproductive recruitment of top-down prefrontal-subcortical circuitry in major depression. *The Journal of Neuroscience*, 27, 8877–8884.

- Kana, R. K., Patriquin, M. A., Black, B. S., Channell, M. M., & Wicker, B. (2016). Altered medial frontal and superior temporal response to implicit processing of emotions in autism. *Autism Research, 9*, 55–66.
- Kirby, L. A. J., & Robinson, J. L. (2017). Affective mapping: An activation likelihood estimation (ALE) meta-analysis. *Brain and Cognition, 118*, 137–148.
- Kirita, T., & Mitsuo, E. (1995). Happy face advantage in recognizing facial expressions. *Acta psychologica, 89*, 149–163.
- Kisilevsky, B. S., Hains, S. M., Brown, C. A., Lee, C. T., Cowperthwaite, B., Stutzman, S. S., et al. (2009). Fetal sensitivity to properties of maternal speech and language. *Infant Behavior & Development, 32*, 59–71.
- Kita, Y., Gunji, A., Inoue, Y., Goto, T., Sakihara, K., Kaga, M., ... Hosokawa, T. (2011). Self-face recognition in children with autism spectrum disorders: A near-infrared spectroscopy study. *Brain & Development, 33*, 494–503.
- Köchel, A., Schöngassner, F., & Schienle, A. (2013). Cortical activation during auditory elicitation of fear and disgust: A near-infrared spectroscopy (NIRS) study. *Neuroscience Letters, 549*, 197–200.
- Kotz, S. A., Kalberlah, C., Bahlmann, J., Friederici, A. D., & Haynes, J. D. (2013). Predicting vocal emotion expressions from the human brain. *Human Brain Mapping, 34*, 1971–1981.
- Kreifelts, B., Ethofer, T., Huberle, E., Grodd, W., & Wildgruber, D. (2010). Association of trait emotional intelligence and individual fMRI-activation patterns during the perception of social signals from voice and face. *Human Brain Mapping, 31*, 979–991.
- Laird, A. R., Robinson, J. L., McMillan, K. M., Tordesillas-Gutiérrez, D., Moran, S. T., Gonzales, S. M., et al. (2010). Comparison of the disparity between Talairach and MNI coordinates in functional neuroimaging data: Validation of the Lancaster transform. *NeuroImage, 51*, 677–683.
- Lancaster, J. L., Tordesillas-Gutiérrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., ... Fox, P. T. (2007). Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. *Human Brain Mapping, 28*, 1194–1205.
- Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., ... Fox, P. T. (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping, 10*, 120–131.
- Liu, P., & Pell, M. D. (2012). Recognizing vocal emotions in mandarin Chinese: A validated database of Chinese vocal emotional stimuli. *Behavior Research Methods, 44*, 1042–1051.
- Lloyd-Fox, S., Richards, J. E., Blasi, A., Murphy, D. G., Elwell, C. E., & Johnson, M. H. (2014). Coregistering functional near-infrared spectroscopy with underlying cortical areas in infants. *Neurophotonics, 1*, 025006.
- Mastropieri, D., & Turkewitz, G. (1999). Prenatal experience and neonatal responsiveness to vocal expressions of emotion. *Developmental Psychobiology, 35*, 204–214.
- Matsui, M., Homae, F., Tsuzuki, D., Watanabe, H., Katagiri, M., Uda, S., ... Taga, G. (2014). Referential framework for transcranial anatomical correspondence for fNIRS based on manually traced sulci and gyri of an infant brain. *Neuroscience Research, 80*, 55–68.
- McDermott, K. B., Petersen, S. E., Watson, J. M., & Ojemann, J. G. (2003). A procedure for identifying regions preferentially activated by attention to semantic and phonological relations using functional magnetic resonance imaging. *Neuropsychologia, 41*, 293–303.
- Mehler, J., Juczyk, P., Lambertz, G., Halsted, N., Bertoni, J., & Amiel-Tison, C. (1988). A precursor of language acquisition. *Cognition, 29*, 143–178.
- Meng, X., Yuan, J., & Li, H. (2009). Automatic processing of valence differences in emotionally negative stimuli: Evidence from an ERP study. *Neuroscience Letters, 464*, 228–232.
- Minagawa-Kawai, Y., Mori, K., Naoi, N., & Kojima, S. (2007). Neural attunement processes in infants during the acquisition of a language-specific phonemic contrast. *The Journal of Neuroscience, 27*, 315–321.
- 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. *Cerebral Cortex, 21*, 254–261.
- Mobbs, D., Greicius, M. D., Abdel-Aziz, E., Menon, V., & Reiss, A. L. (2003). Humor modulates the mesolimbic reward centers. *Neuron, 40*, 1041–1048.
- Moon, C., Cooper, R. P., & Fifer, W. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development, 16*, 495–500.
- Nair, V. A., Young, B. M., La, C., Reiter, P., Nadkarni, T. N., Song, J., et al. (2015). Functional connectivity changes in the language network during stroke recovery. *Annals of Clinical Translational Neurology, 2*, 185–195.
- Norbury, R., Mackay, C. E., Cowen, P. J., Goodwin, G. M., & Harmer, C. J. (2008). The effects of reboxetine on emotional processing in healthy volunteers: An fMRI study. *Molecular Psychiatry, 13*, 1011–1020.
- Obrig, H., Mock, J., Stephan, F., Richter, M., Vignotto, M., & Rossi, S. (2017). Impact of associative word learning on phonotactic processing in 6-month-old infants: A combined EEG and fNIRS study. *Developmental Cognitive Neuroscience, 25*, 185–197.
- Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., & Griffiths, T. D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron, 36*, 767–776.
- Pell, M. D., Monetta, L., Paulmann, S., & Kotz, S. A. (2009). Recognizing emotions in a foreign language. *Journal of Nonverbal Behavior, 33*, 107–120.
- Peltola, M. J., Leppänen, J. M., Mäki, S., & Hietanen, J. K. (2009). Emergence of enhanced attention to fearful faces between 5 and 7 months of age. *Social Cognitive and Affective Neuroscience, 4*, 134–142.
- Peña, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., et al. (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*, 11702–11705.
- Perrachione, T. K., & Wong, P. C. M. (2007). Learning to recognize speakers of a non-native language: Implications for the functional organization of human auditory cortex. *Neuropsychologia, 45*, 1899–1910.
- Price, C. J. (2010). The anatomy of language: A review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences, 1191*, 62–88.
- Quaresima, V., Bisconti, S., & Ferrari, M. (2012). A brief review on the use of functional near-infrared spectroscopy (fNIRS) for language imaging studies in human newborns and adults. *Brain and Language, 121*, 79–89.
- Reker, M., Ohrmann, P., Rauch, A. V., Kugel, H., Bauer, J., Dannlowski, U., ... Suslow, T. (2010). Individual differences in alexithymia and brain response to masked emotion faces. *Cortex, 46*, 658–667.
- Rigato, S., Farroni, T., & Johnson, M. H. (2010). The shared signal hypothesis and neural responses to expressions and gaze in infants and adults. *Social Cognitive and Affective Neuroscience, 5*, 88–97.
- Rutherford, M. D., Baron-Cohen, S., & Wheelwright, S. (2002). Reading the mind in the voice: a study with normal adults and adults with Asperger syndrome and high functioning autism. *J Autism Dev. Disord., 32*, 189–194.
- Saito, Y., Kondo, T., Aoyama, S., Fukumoto, R., Konishi, N., Nakamura, K., ... Toshima, T. (2007). The function of the frontal lobe in neonates for response to a prosodic voice. *Early Human Development, 83*, 225–230.
- Sambeth, A., Ruohio, K., Alku, P., Fellman, V., & Huotilainen, M. (2008). Sleeping newborns extract prosody from continuous speech. *Clinical Neurophysiology, 119*, 332–341.
- Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). Emotion and attention interactions in social cognition: Brain regions involved in processing anger prosody. *NeuroImage, 28*, 848–858.
- Sato, H., Hirabayashi, Y., Tsubokura, H., Kanai, M., Ashida, T., Konishi, I., ... Maki, A. (2012). Cerebral hemodynamics in newborn infants exposed to speech sounds: A whole-head optical topography study. *Human Brain Mapping, 33*, 2092–2103.
- Schaefer, H. S., Putnam, K. M., Benca, R. M., & Davidson, R. J. (2006). Event-related functional magnetic resonance imaging measures of neural activity to positive social stimuli in pre- and post-treatment depression. *Biological Psychiatry, 60*, 974–986.
- Schirmer, A., & Kotz, S. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences, 10*, 24–30.
- Shenton, M. E., Kikinis, R., Jolesz, F. A., Pollak, S. D., LeMay, M., Wible, C. G., ... Coleman, M. (1992). Abnormalities of the left temporal lobe

- and thought disorder in schizophrenia. A quantitative magnetic resonance imaging study. *N Engl J Med.*, 327, 604–612.
- Shi, F., Yap, P. T., Wu, G., Jia, H., Gilmore, J. H., Lin, W., & Shen, D. (2011). Infant brain atlases from neonates to 1- and 2-year-olds. *PLoS One*, 6, e18746.
- Silani, G., Lamm, C., Ruff, C. C., & Singer, T. (2013). Right supramarginal gyrus is crucial to overcome emotional egocentricity bias in social judgments. *The Journal of Neuroscience*, 33, 15466–15476.
- Singh, A. K., Okamoto, M., Dan, H., Jurcak, V., & Dan, I. (2005). Spatial registration of multichannel multi-subject fNIRS data to MNI space without MRI. *NeuroImage*, 27, 842–851.
- Storey, J. D., & Tibshirani, R. (2003). Statistical significance for genome wide studies. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 9440–9445.
- 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, 1246–1252.
- Telkemeyer, S., Rossi, S., Koch, S. P., Nierhaus, T., Steinbrink, J., Poeppel, D., ... Wartenburger, I. (2009). Sensitivity of newborn auditory cortex to the temporal structure of sounds. *The Journal of Neuroscience*, 29, 14726–14733.
- Thompson, C. P. (1987). A language effect in voice identification. *Applied Cognitive Psychology*, 1, 121–131.
- Thompson, W., & Balkwill, L. L. (2006). Decoding speech prosody in five languages. *Semiotica*, 158, 407–424.
- Townsend, J. D., Torrisi, S. J., Lieberman, M. D., Sugar, C. A., Bookheimer, S. Y., & Althuler, L. L. (2013). Frontal-amygdala connectivity alterations during emotion downregulation in bipolar I disorder. *Biological Psychiatry*, 73, 127–135.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., ... Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, 15, 273–289.
- Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, 134, 383–403.
- Vaish, A., & Striano, T. (2004). Is visual reference necessary? Vocal versus facial cues in social referencing. *Developmental Science*, 7, 261–269.
- Van Lancker, D., Cornelius, C., & Kreiman, J. (1989). Recognition of emotional-prosodic meanings in speech by autistic, schizophrenic, and normal children. *Developmental Neuropsychology*, 5, 207–226.
- Van Lancker Sidtis, D., Pachana, N., Cummings, J. L., & Sidtis, J. J. (2006). Dysprosodic speech following basal ganglia insult: Toward a conceptual framework for the study of the cerebral representation of prosody. *Brain and Language*, 97, 135–153.
- Vannasing, P., Florea, O., González-Frankenberger, B., Tremblay, J., Paquette, N., Safi, D., ... Gallagher, A. (2016). Distinct hemispheric specializations for native and non-native languages in one-day-old newborns identified by fNIRS. *Neuropsychologia*, 84, 63–69.
- Wager, T. D., Davidson, M. L., Hughes, B. L., Lindquist, M. A., & Ochsner, K. N. (2008). Prefrontal-subcortical pathways mediating successful emotion regulation. *Neuron*, 59, 1037–1050.
- Wang, Y., Sereno, J. A., Jongman, A., & Hirsch, J. (2003). fMRI evidence for cortical modification during learning of mandarin lexical tone. *Journal of Cognitive Neuroscience*, 15, 1019–1027.
- Watanabe, H., Homae, F., Nakano, T., Tsuzuki, D., Enkhur, L., Nemoto, K., ... Taga, G. (2013). Effect of auditory input on activations in infant diverse cortical regions during audiovisual processing. *Human Brain Mapping*, 34, 543–565.
- Wiethoff, S., Wildgruber, D., Kreifelts, B., Becker, H., Herbert, C., Grodd, W., & Ethofer, T. (2008). Cerebral processing of emotional prosody—Influence of acoustic parameters and arousal. *NeuroImage*, 39, 885–893.
- Wildgruber, D., Ackermann, H., Kreifelts, B., & Ethofer, T. (2006). Cerebral processing of linguistic and emotional prosody: FMRI studies. *Prog. Brain Res.*, 4, 249–268.
- Wildgruber, D., Ethofer, T., Grandjean, D., & Kreifelts, B. (2009). A cerebral network model of speech prosody comprehension. *International Journal of Speech-Language Pathology*, 11, 277–281.
- Witteman, J., Van Heuven, V. J., & Schiller, N. O. (2012). Hearing feelings: a quantitative meta-analysis on the neuroimaging literature of emotional prosody perception. *Neuropsychologia*, 50, 2752–2763.
- Wong, P. C., Perrachione, T. K., & Parrish, T. B. (2007). Neural characteristics of successful and less successful speech and word learning in adults. *Human Brain Mapping*, 28, 995–1006.
- Yuan, J., Luo, Y., Yan, J. H., Meng, X., Yu, F., & Li, H. (2009). Neural correlates of the females' susceptibility to negative emotions: An insight into gender-related prevalence of affective disturbances. *Human Brain Mapping*, 30, 3676–3686.
- Yuan, J., Meng, X., Yang, J., Hu, L., & Yuan, H. (2012). The valence strength of unpleasant emotion modulates brain processing of behavioral inhibitory control: Neural correlates. *Biological Psychology*, 89, 240–251.
- Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, 11, 946–953.
- Zhang, D., Liu, Y., Hou, X., Sun, G., Cheng, Y., & Luo, Y. (2014). Discrimination of fearful and angry emotional voices in sleeping human neonates: A study of the mismatch brain responses. *Frontiers in Behavioral Neuroscience*, 8, 422.
- Zhang, H., Chen, X., Chen, S., Li, Y., Chen, C., Long, Q., & Yuan, J. 2018. Facial Expression Enhances Emotion Perception Compared to Vocal Prosody: Behavioral and fMRI Studies. *Neurosci Bull.*, 2018 May 9. <https://doi.org/10.1007/s12264-018-0231-9>. [Epub ahead of print]

**How to cite this article:** Zhang D, Chen Y, Hou X, Wu YJ. Near-infrared spectroscopy reveals neural perception of vocal emotions in human neonates. *Hum Brain Mapp.* 2019;40: 2434–2448. <https://doi.org/10.1002/hbm.24534>