Resting-state connectivity reveals a role for sensorimotor systems in vocal emotional processing in children

Ana Isabel Correia [a,d,1], Paulo Branco [a,b,1], Marta Martins [a], Ana Mafalda Reis [c], Nuno Martins [c], São Luís Castro [a,2], César F. Lima [a,d,e,*,2]

[a] Faculty of Psychology and Education Sciences, University of Porto, Porto, Portugal
[b] Faculty of Medicine, University of Porto, Porto, Portugal
[c] Unilabs - Boavista, Porto, Portugal
[d] Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal
[e] Institute of Cognitive Neuroscience, University College London, UK

ARTICLE INFO

ABSTRACT

Voices are a primary source of emotional information in everyday interactions. Being able to process non-verbal vocal emotional cues, namely those embedded in speech prosody, impacts on our behaviour and communication. Extant research has delineated the role of temporal and inferior frontal brain regions for vocal emotional processing. A growing number of studies also suggest the involvement of the motor system, but little is known about such potential involvement. Using resting-state fMRI, we ask if the patterns of motor system intrinsic connectivity play a role in emotional prosody recognition in children. Fifty-five 8-year-old children completed an emotional prosody recognition task and a resting-state scan. Better performance in emotion recognition was predicted by a stronger connectivity between the inferior frontal gyrus (IFG) and motor regions including primary motor, lateral premotor and supplementary motor sites. This is mostly driven by the IFG pars triangularis and cannot be explained by differences in domain-general cognitive abilities. These findings indicate that individual differences in the engagement of sensorimotor systems, and in its coupling with inferior frontal regions, underpin variation in children's emotional speech perception skills. They suggest that sensorimotor and higher-order evaluative processes interact to aid emotion recognition, and have implications for models of vocal emotional communication.

1. Introduction

Voices are central to emotional communication in everyday interactions. Similar to facial expressions, modulations of the *tone of voice* while speaking – emotional prosody – tell us a lot about the emotional states and intentions of others. Prosodic cues include fundamental frequency, amplitude, timing and voice quality variations in speech. Being able to efficiently recognize emotions from these cues relates to psychosocial adjustment both in children and adults (Demopoulos et al., 2016; Leppänen and Hietanen, 2001; McClure and Nowicki Jr., 2001; Nowicki Jr. and Duke, 1992; Ruffman et al., 2010). A frontotemporal network of brain regions plays a crucial role during emotional speech processing, including bilateral auditory cortices (AC), regions along the superior temporal gyrus and sulcus (superior temporal cortex, STC), and

the inferior frontal gyrus (IFG; Dricu and Frühholz, 2016; Frühholz and Grandjean, 2013a, 2013b; Frühholz et al., 2016; Schirmer, 2018; Schirmer and Kotz, 2006).

It has been proposed that vocal emotions are perceived in multiple steps (Schirmer and Kotz, 2006; for a multimodal version, Brück et al., 2011). First, emotionally relevant low-level features of the stimuli are extracted in auditory cortices. Second, regions of the superior temporal gyrus and sulcus are engaged along an auditory ventral pathway, where the processing becomes more integrative and the emotional meaning of the expression is derived. Finally, information is fed into frontal areas, namely the IFG, and made available for higher-order cognitive processes. Both the left and right IFG are suggested to be sensitive to higher-order acoustic information, and to support attentive processes such as the evaluation, categorization and labelling of vocal expressions (Frühholz

and Grandjean, 2013b).

While the role of temporal and inferior frontal regions for emotional prosody is robustly established, the potential involvement of other systems is less understood (e.g., for a recent discussion of the amygdala, Schirmer, 2018). This is particularly the case of the motor system. Extant research discusses its role for speech perception (Scott et al., 2009; Skipper et al., 2017), but little is known about whether and how this extends to nonverbal vocal communication. A growing number of studies with adults is suggestive of a motor involvement in vocal emotional processing. Motor system activation, including in primary motor, lateral premotor and supplementary motor sites, is often seen in response to nonverbal emotional vocalizations (e.g., laughter) (Bestelmeyer et al., 2014; Lima et al., 2015; McGettigan et al., 2015; Warren et al., 2006). Additionally, stimulation of the sensorimotor cortex disrupts the perception of nonverbal vocalizations (Banissy et al., 2010) and linguistic prosody (Sammler et al., 2015). Plausibly, as part of an auditory dorsal pathway, the motor system could support the generation of sensory predictions based on previous sensorimotor experience, that could be flexibly used to optimize perceptual and evaluative processes during auditory processing (Lima et al., 2016). Consistent with this, larger motor system responses during listening to posed and spontaneous laughs were found to correlate with better performance in a laughter authenticity discrimination task (McGettigan et al., 2015).

To date, central questions remain unanswered, though: there is no account linking motor system engagement with facilitated emotional prosody perception, or delineating how it interacts with other nodes of the vocal emotion network. Furthermore, the potential role of the motor system remains especially unknown in childhood. The mechanisms for vocal emotional processing are in place since early in development: 3- to 7-month infants show specialized neural responses to voices in the anterior temporal cortex, similarly to adults (Blasi et al., 2011), and there is suggestive evidence that vocal emotional cues are represented in frontotemporal brain regions within the first year of life (for a recent review, Morningstar et al., 2018). Additionally, 6-month infants discriminate emotions in nonverbal vocalizations and prosodic cues (Flom and Bahrick, 2007; Soderstrom et al., 2017), and children as young as five years are already able to recognize a range of vocal emotions (Allgood and Heaton, 2015; Sauter et al., 2013; but see Aguert et al., 2013). Surprisingly, however, no neuroimaging studies have been conducted on emotional prosody processing in childhood (Morningstar et al., 2018), and motor system activations have been found for laughter perception in adolescents (along with temporal and inferior frontal responses; O'Nions et al., 2017), but the involvement of this system remains to be examined at earlier ages.

In the current study, we investigated for the first time whether emotional prosody recognition abilities in children, indexed by a forced-choice emotion categorization task, relate to patterns of resting-state functional connectivity. Resting-state fMRI can provide unique insights into the role of the motor system in vocal emotions, as it allows us to measure synchronous activations between spatially distinct brain regions in the absence of a task or stimulus (e.g., Lee et al., 2013). A potential motor involvement could thus more confidently be attributed to emotion-specific processes rather than to task-related motor demands, as the task is conducted offline. The degree of covariance in spontaneous fluctuations of the BOLD signal, measured as the strength of intrinsic connectivity between two or more brain regions, has been shown to be associated with behavioural outcomes in adults and children (e.g., Angelides et al., 2017; Koyama et al., 2011; Mollo et al., 2016; Supekar et al., 2013), but no studies so far have examined emotional prosody. Using a hypothesis-driven seed-based correlational approach, we hypothesized that a stronger intrinsic connectivity between the motor system (target) and regions established to play a role for vocal emotional processing – AC, STC and IFG (seeds) – relates to a better ability to recognize prosodic emotions. Among the seeds, results could be particularly clear for the IFG, as our behavioural task emphasized attentive-evaluative processes, a role of this region (Frühholz and Grandjean, 2013b). Furthermore, the anatomical connections between the IFG and the motor cortex are well described (Catani et al., 2012; Vergani et al., 2014). The study was conducted with a large sample of 8-year-old children (N = 55), recruited in the context of an on-going wider project in our lab on development, plasticity and auditory processing. In addition to shedding light on the neural basis of emotional prosody perception during development, investigating children is ideal for an individual differences approach: variability in emotion processing can be seen at any age (e.g., Lima et al., 2015; McGettigan et al., 2015; Neves et al., 2018), but it tends to be larger during development than in adulthood (e.g., Chronaki et al., 2018).

2. Material and methods

2.1. Participants

Sixty-seven children were initially involved in the study. They were all European Portuguese native speakers and had normal hearing. Twelve of them had to be excluded because of excessive head movement during scanning (average motion larger than 1 mm or peak movement larger than 3.5 mm, which corresponds to the in-plane voxel size; $n = 5$), incidental MRI findings ($n = 1$), atypically low full-scale IQ (below 70; American Psychiatric Association, 2013; $n = 5$), or atypically low performance on the emotional prosody recognition task (two standard deviations below the group mean; $n = 1$). The final sample consisted of fifty-five children (23 boys, M age = 8.31 years, range = 7.75–9.25, $SD = 0.32$). Fifty of them were right-handed, 2 were left-handed, and 3 were ambidextrous, according to the Edinburgh Handedness Inventory (Oldfield, 1971). Participants were tested as part of a wider project looking at music training, auditory processing, and brain plasticity in third graders from elementary public schools.

All aspects of the study were approved by the local ethics committee (FPCEUP 2015.1.23), and the work was conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from parents and from local school authorities, and children gave their verbal assent at the start of data collection. Additionally, parents completed a safety form to ensure that the children could be safely scanned.

2.2. Procedure

Behavioural and brain testing occurred over three sessions conducted on different days. In the first session, children completed all subtests of the Portuguese version of the Wechsler Intelligence Scale for Children – 3rd Edition (WISC-III; Wechsler, 2003), which was used as a measure of IQ and administered by an experienced child psychologist. In the second session, they completed the emotional prosody recognition task. Finally, the third session was a scanning session, in which they completed structural and resting-state MRI scans.

2.3. Experimental design and statistical analysis

2.3.1. Auditory stimuli and behavioural task

The experimental stimuli consisted of fifty spoken utterances, recorded by two female speakers. They were semantically neutral utterances (e.g., "O futebol é um desporto", football is a sport) that communicated five emotional states via prosodic cues only – anger, fear, happiness, sadness and neutrality (10 stimuli per emotion). These stimuli were drawn from a perceptually and acoustically validated corpus (Castro and Lima, 2010), and they have been used in previous studies (Lima and Castro, 2011; Lima et al., 2013). Based on the validation data (Castro and Lima, 2010), we ensured that categorization accuracy ($M = 83\%$ correct, $SD = 10$) and duration ($M = 1472$ ms, $SD = 247$) were matched across emotion categories.

After a familiarization phase (three practice trials), the stimuli were presented in a randomized order across two blocks of 25 trials each. On each trial, participants were instructed to perform two consecutive

judgements: a forced-choice identification of the emotional tone (neutral, happy, sad, angry, scared), followed by an intensity judgement, rating how much the expression was present in the stimuli on a scale from 1 (low intensity) to 5 (high intensity; intensity judgements were not used for the current study; task format adapted from Castro and Lima, 2010; Lima and Castro, 2011). The structure of the task is shown schematically in Fig. 1A. Based on pilot testing, we decided to provide trial-based feedback, as this improved children's attention and engagement throughout the task. This did not influence performance in important ways, though: learning effects from Block 1 to Block 2 were low ($M = 8.2\%$; $SD = 12$) and they did not correlate with average performance on the task ($r = 0.12$, $p = 0.37$), that is, individual differences in emotion recognition (the focus of our analyses) were not related to how much children benefited from feedback. Latency data were not considered because the emphasis of the task was on accuracy and children were not instructed to provide a speeded response.

The stimuli were presented via high-quality headphones (Sennheiser HD 202), using SuperLab Version 5.0.3 (Cedrus Corporation, San Pedro, CA), running on an Apple MacBook Air laptop. Responses were collected using a seven-button response pad (Cedrus RB-740). The children were tested in individual sessions lasting about 20 min, conducted in a quiet room at their school.

Performance on the task was examined in terms of the number of correct categorizations per emotion, i.e., number of trials in which the selected category corresponded to the intended expression of the utterance. Raw accuracy rates were corrected for possible response biases using unbiased hit rates, 'Hu' (Wagner, 1993; for a discussion of potential biases in forced-choice tasks, e.g., Isaacowitz et al., 2007). Hu values vary between 0 and 1: when no stimulus of a given category (e.g., happy prosody) is correctly identified, Hu = 0; when all the stimuli of a given category are correctly identified, and the corresponding response category is always correctly used, Hu = 1. Statistical analyses were conducted using JASP Version 0.8.5 (JASP Team, 2017).

### 2.3.2. MRI data acquisition

MRI scanning was completed on a 1.5T Siemens Magnetom Sonata Maestro Class scanner. Resting-state fMRI data were acquired using an EPI sequence with the following parameters: repetition time = 2.5 s, echo time = 60 ms, flip angle = 90°, FOV = 224 x 224 × 134 mm, in-plane resolution = 3.5 × 3.5 mm, slice thickness = 5 mm, and axial slices = 21 (gap size = 29%). One hundred and eighty volumes were collected, in a run lasting 7 min 30s. Children were instructed to keep their eyes open, to think of nothing in particular, and to avoid falling asleep. We confirmed that they stayed awake via visual monitoring throughout the scan. A high-resolution anatomical image was also acquired for registration purposes, using a T1-weighted sequence with the following parameters: repetition time = 1680 ms; echo time = 4.12 ms; flip angle = 8°; FOV = 250 x 250 × 160 mm, in-plane resolution = 1 × 1 mm, slice thickness = 1 mm, and axial slices = 180; acquisition time, 7 min 20 s. A foam headrest and a forehead strap were used to minimize head motion during scanning.

2.3.3. MRI data processing

MRI analyses were conducted using the Oxford Centre for Functional Magnetic Resonance Imaging of the Brain Software Library (FMRIB, Oxford UK; FSL version 5.0.10, https://fsl.fmrib.ox.ac.uk/fsl; RRID:SCR_002823; Jenkinson et al., 2012). Structural images were skull-stripped using BET (Smith, 2002) and segmented into grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF) masks using FAST (Zhang et al., 2001). To ensure no overlap between GM and CSF or WM masks, resulting segmentation masks were adjusted using a binarization threshold of 0.8. Regarding resting-state fMRI data, the first four volumes were discarded to allow for T1 signal stabilization. Motion correction was performed by aligning all volumes to a reference middle time point using MCFLIRT (Jenkinson et al., 2002). Data were skull stripped using BET (Smith, 2002) and spatially smoothed with an FWHM Gaussian kernel of 8 mm. ICA-AROMA (Pruim et al., 2015) was used for data denoising. To further remove residual artefacts from the data, WM and CSF timecourses were estimated from their respective masks and removed from data through multiple regression. Finally, high-pass temporal filtering was performed using an FWHM of 100 s.

Functional scans were coregistered to the structural T1 image by linear registration using FLIRT (Jenkinson et al., 2002) with boundary-based registration (BBR, Greve and Fischl, 2009). Images were then warped to an age-matched standard template (7–11 years MNI template, Fonov et al., 2009, 2011) using FLIRT with 12 degrees of
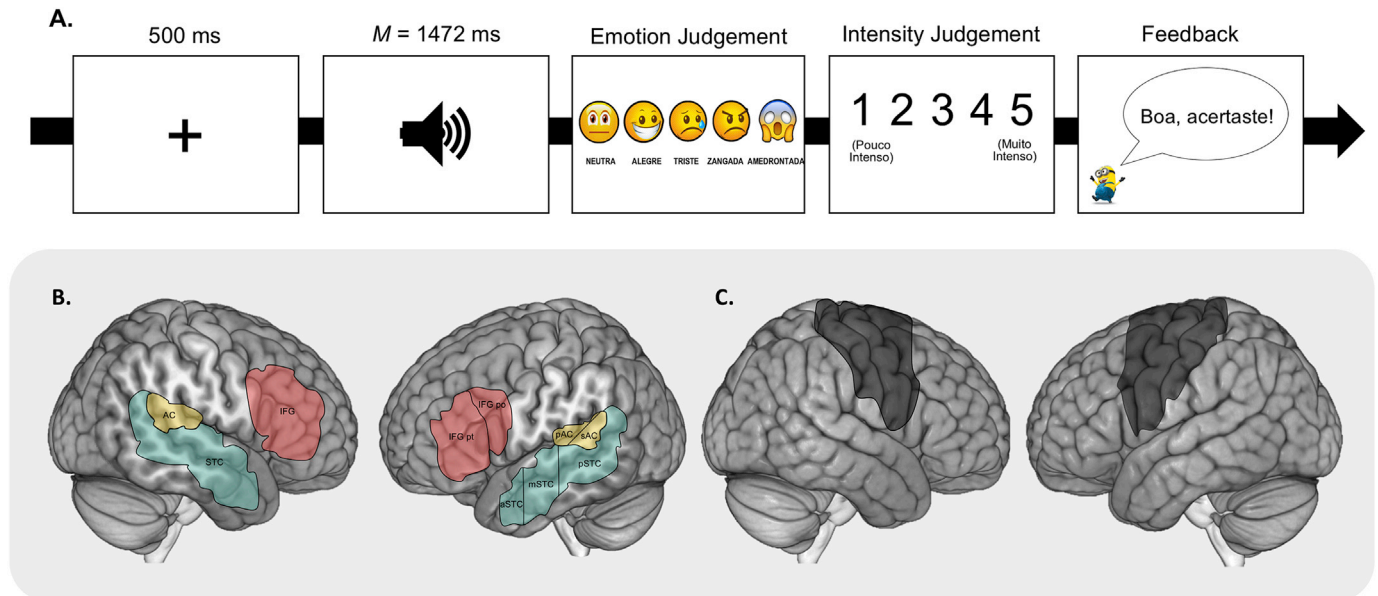
Fig. 1. Structure of the behavioural emotional prosody recognition task (A); regions of interest, ROIs, used as seeds in seed-based correlation analyses (B); and the Human Motor Area Template (HMAT; Mayka et al., 2006), used as target ROI (C). AC, auditory cortex; STC, superior temporal cortex; IFG, inferior frontal gyrus; pAC/sAC: primary and secondary AC, respectively; aSTC/mSTC/pSTC, anterior, middle, and posterior STC, respectively; IFG pt, IFG *pars triangularis*; IFG po, IFG *pars opercularis*.

freedom, and further refined by non-linear registration using FNIRT (Andersson et al., 2007). Data were then inspected for excessive motion. After excluding children with a single movement larger than 3.5 mm or average motion larger than 1 mm, the remaining sample had overall low motion indices (peak movement = 0.86 mm, range = 0.11–3.17; M motion = 0.10, range = 0.03–0.33).

2.3.4. Resting-state fMRI analysis

Resting-state data were analysed using a hypothesis-driven seed-based correlation approach. Three seed regions were selected, namely the AC, STC, and IFG, both on the left and right hemispheres (total of six regions of interest, ROIs). These regions were selected as they form core nodes of the emotional prosody network (e.g., Dricu and Frühholz, 2016; Frühholz and Grandjean, 2013a, 2013b; Frühholz et al., 2016; Kotz and Paulmann, 2011; Sammler et al., 2015; Schirmer and Adolphs, 2017; Schirmer and Kotz, 2006). Masks for the ROIs were devised from the Brodmann template from MRIcron: AC comprised BAs 41 and 42, STC comprised BAs 21 and 22, and IFG comprised BAs 44 and 45 (http://peop le.cas.sc.edu/rorden/mricron/index.html). To examine the potential involvement of the motor system in emotional prosody processing, we selected as target ROI the Human Motor Area Template (HMAT; Mayka et al., 2006). The HMAT is a well-established template, developed based on anatomical and functional information, and it includes the pre-SMA, SMA, dorsal premotor cortex, ventral premotor cortex, primary motor cortex, and primary somatosensory cortex. The seed and target ROIs are illustrated in Fig. 1B and C, respectively.

Seed-based correlation analyses were performed by extracting BOLD timecourses from each seed region and comparing them against each voxel within the HMAT ROI. This was conducted using non-parametric one-sample t-tests with additional covariates, through the FSL *randomise* tool (Winkler et al., 2014). Participants' ability to recognize prosodic emotions (Hu scores) was included in the design matrix as the explanatory variable. Age, sex and IQ were included as covariates of no-interest, to regress out any potential confounding effects related to these variables. Statistics were computed after 5000 permutations and were corrected for multiple comparisons using a threshold-free cluster enhancement (TFCE) cluster correction at $p < 0.05$ significance level. Whenever appropriate, significant results were further examined through multiple regressions in JASP in order to extract r-squared values. This was achieved by extracting mean z-scores within the significant clusters of the main analyses.

As a follow-up to the main analysis, we also conducted correlations focusing on functional subdivisions within the IFG, STC, and AC,

considering recent evidence for their potential differential involvement in emotional prosody (e.g., Frühholz and Grandjean, 2013a, 2013b). The IFG was subdivided into IFG pars triangularis (BA 45) and pars opercularis (BA 44); the STC into anterior, middle and posterior regions, using the same reference coordinates as Frühholz and Grandjean (2013a; aSTC: y > 0, mSTC: −20 > y > 0, and pSTC: y < −20); and the AC was subdivided into primary (BA 41) and secondary (BA 42) subregions (the divisions of the ROIs are illustrated in Fig. 1B).

Finally, for completeness, we also performed an exploratory whole-brain seed-based correlation analysis for each seed ROI. This was to examine whether there were additional regions (others than those included in the HMAT) where connectivity with the seed ROIs related to performance on the emotional prosody recognition task.

2.3.5. Cross-validation and predictive power

As a follow-up to the seed-based correlation analyses, we evaluated the predictive power of our models using a K-fold cross-validation approach. K-fold cross-validation randomly partitions the data into a set number of folds (K), and compares each fold (i.e., validation data) to the predicted linear trend derived from the remaining (training) data. This provides indices on how well random subsets of the data would fit the unbiased predicted model, including the root mean square error (RMSE), mean absolute error (MAE) and predicted r-squared values. A 5-fold repeated cross-validation was performed, using the R Package Caret v. 6.0–78 (Kuhn, 2008). Five folds were randomly assigned and tested (repeated) 10 times against the linear model fit between Hu scores and brain connectivity, after regressing out the variance explained by age, sex and IQ. RMSE, MAE and predicted r-squared were extracted and averaged across folds. For interpretability, RMSE values are reported in their normalised form (NRMSE), by dividing RMSE values by the range of the Hu scores (explanatory variable).

3. Results

3.1. Behavioural results

Participants' accuracy in emotional prosody recognition is shown in Fig. 2A. Children performed generally well on the task, but there were large individual differences: average recognition accuracy ranged between 0.21 and 0.79 ($M = 0.49$; $SD = 0.15$), and this reflects variability that was seen for all emotion categories (Fig. 2A). There were no differences in recognition accuracy across categories (Bonferroni-corrected ps > 0.10), except that happy expressions were better recognized than all
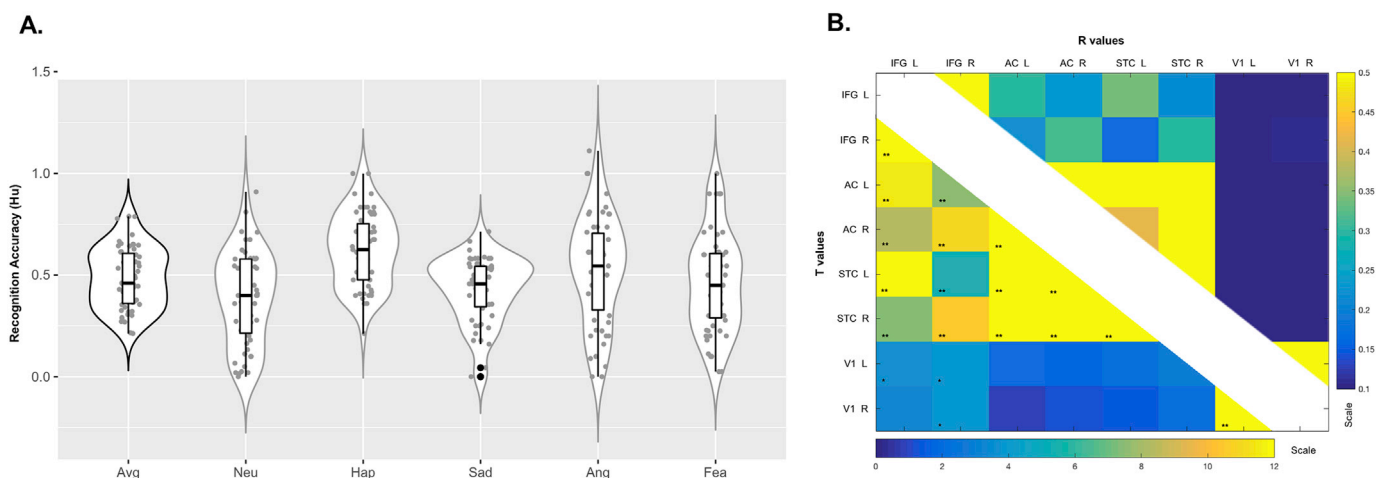
Fig. 2. Individual results, box plots and violin plots depicting emotional prosody recognition accuracy for each category and for the average across emotions (middle horizontal lines represent the medians; boxes indicate the 25th and 75th percentiles; vertical bars denote 95% confidence interval) (A); Intercorrelations between the seed ROIs (IFG, AC, STC) and a control region (V1), *$p < 0.05$, **$p < 0.001$, Bonferroni-corrected (B). Avg, average; Neu, neutrality; Hap, happiness; Sad, sadness; Ang, anger; Fea, fear; L, left; R, right; IFG, inferior frontal gyrus; AC, auditory cortex; STC, superior temporal cortex; V1, primary visual cortex.

the other ones ($ps < 0.03$), and angry expressions were better recognized than neutral ones ($p = 0.02$; main effect of emotion, $F[4, 216] = 14.86$, $p < 0.001$). Performance levels were positively correlated across emotions (see supplementary materials, Fig. S1), and analyses of skewness and kurtosis indicated that there was no substantial departure from normality in emotion recognition scores (skewness, range $= -0.841 - 0.278$; kurtosis, range $= -0.857 - 0.247$; Curran et al., 1996).

3.2. Resting-state fMRI results

3.2.1. Connectivity between seed ROIs

We first asked whether the selected seed ROIs – AC, STC and IFG – were functionally connected with each other, as it could be expected if they formed a network supporting emotional prosody processing. We also examined the specificity of this result by considering connectivity with a control ROI (the left and right primary visual cortex, BA 17). Pairwise correlations across all seed ROIs indicated that they are indeed functionally connected (M $r = 0.41$, range $= 0.17–0.66$, Bonferroni-corrected $ps < .001$), while no substantial evidence of connectivity was found between any of them and the control ROI (M $r = 0.07$, range $= 0.02–0.11$). These results are depicted in Fig. 2B. They add to previous evidence of connectivity between these regions in adults, including structural connectivity studies (e.g., Ethofer et al., 2012; Frühholz et al., 2015; Sammler et al., 2015) and task-based functional connectivity studies on prosody (e.g., Ethofer et al., 2012; Frühholz and

Grandjean, 2012; Leitman et al., 2010).

3.2.2. Connectivity correlates of emotional prosody recognition

Our primary question was whether emotional prosody recognition abilities related to resting-state connectivity between the motor system and the AC, STC and IFG. Seed-based correlation analyses indicated that this was the case for the left and right IFG. Higher average emotion recognition accuracy was associated with greater functional connectivity between the left IFG and two clusters with peaks in the right precentral gyrus and SMA ($R^2 = 0.32$, $p < 0.001$); and between the right IFG and a cluster with peak in the left posterior cingulate gyrus extending to the precentral gyrus ($R^2 = 0.31$, $p < 0.001$). These results are shown in Fig. 3 and listed in Table 1. No suprathreshold associations were found for the remaining seed ROIs, AC ($ps > 0.67$) and STC ($ps > 0.76$). To ensure that the uncovered relationship between emotion recognition and IFG-motor system connectivity is not driven by a single or small subset of emotions, we extracted connectivity values within the motor system clusters found to be connected with the IFG, and conducted follow-up multiple regression analyses separately for each category. Correlations were significant for all emotions (Bonferroni-corrected $ps < 0.04$), except for sadness ($R^2 = 0.04$, $p = 0.68$ for the left IFG, and $R^2 < 0.01$, $p = 1$ for the right IFG; scatterplots for each emotion are presented in supplementary materials, Figs. S2 and S3).

We next asked whether distinct functional subdivisions of our seed ROIs might have contributed differently to the relationship between
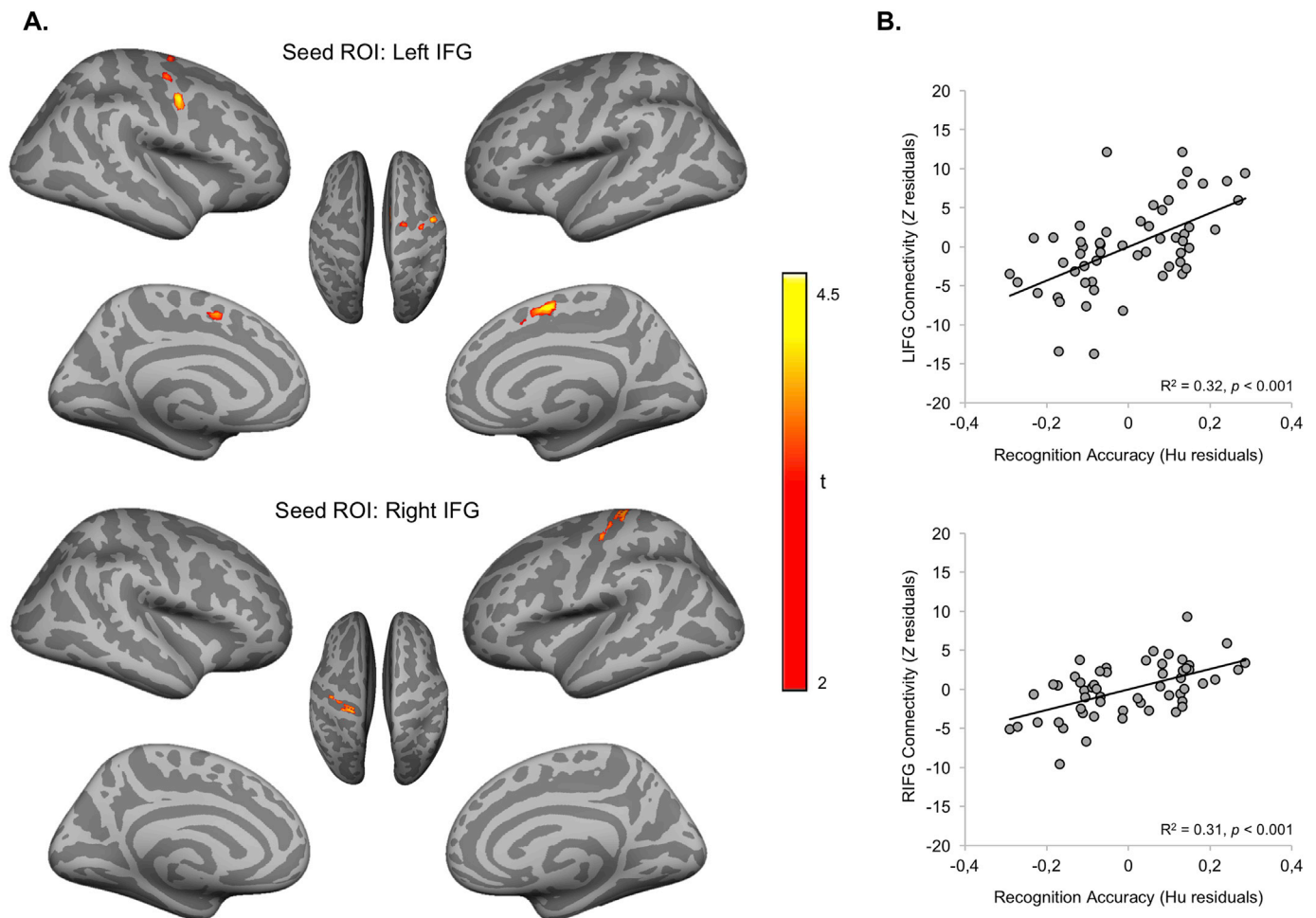
Fig. 3. Maps of motor system regions in which the strength of connectivity with the IFG was associated with emotional prosody recognition ability; thresholded activation maps ($p < 0.05$, TFCE-corrected) are registered to and displayed on a cortical surface using Freesurfer; the cortical surface was generated using an age-appropriate MNI template (see Materials and Methods) (A); scatterplots showing the association between emotional prosody recognition ability and IFG - motor system connectivity (B). IFG, inferior frontal gyrus.

Table 1
Brain regions showing significant association with emotional prosody recognition.

Seed ROIs	BA	Cluster location within target ROI	Cluster size (voxels)	MNI coordinates			t
				x	y	z	
Main Analysis							
L IFG	44/45	R Precentral Gyrus	395	30	−8	68	3.78
		R Precentral Gyrus		42	−4	60	
		R Precentral Gyrus		44	−8	62	
		R Superior Frontal Gyrus		20	−10	60	
		R Precentral Gyrus		46	−18	70	
		SMA	296	8	−2	56	4.41
		Paracingulate Gyrus		4	8	46	
		SMA		−12	4	52	
		R Precentral Gyrus	144	56	−4	46	4.68
R IFG	44/45	L Posterior Cingulate Gyrus	220	−18	−34	58	3.92
		L Posterior Cingulate Gyrus		−26	−30	58	
		L Precentral Gyrus		−30	−22	50	
Follow up analysis							
L IFG pars triangularis	45	R Superior Frontal Gyrus	1117	20	−8	60	4.22
		SMA		8	−2	56	
		Paracingulate Gyrus		6	12	44	
		R Precentral Gyrus		40	−6	58	
		Paracingulate Gyrus		10	12	46	
		SMA		6	2	48	
		R Precentral Gyrus	277	50	−6	36	4.16
		R Precentral Gyrus		56	−6	44	
		L Precentral Gyrus	184	−50	−6	38	4.49
		L Precentral Gyrus	7	−8	−16	72	4.08

Results were thresholded at $p < 0.05$, TFCE-corrected. Labels were taken from the Oxford-Harvard Structural Cortical Atlas (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases; RRID:SCR_001476). ROI, Region of interest; BA, Brodmann area; MNI, Montreal Neurological Institute; L, left; R, right; IFG, inferior frontal gyrus; SMA, supplementary motor area.

connectivity and emotional prosody recognition. Results were particularly clear for the left IFG pars triangularis (BA 45): consistent with the main analysis, higher average emotion recognition accuracy was associated with greater functional connectivity between this subdivision and several regions of the motor system, including the left and right precentral gyrus, SMA, and paracingulate gyrus ($R^2 = 0.31$, $p < 0.001$; see Fig. 4 and Table 1 for details). Such association was also significant for all emotions (Bonferroni-corrected $ps < 0.03$), except for sadness ($R^2 = 0.08$, $p = 0.19$; see scatterplots for each emotion in supplementary materials, Fig. S4). In contrast, only trend-level results were obtained in analyses focused on the right IFG pars triangularis (BA 45; $p = 0.05$), and on the left and right IFG pars opercularis (BA 44; $p = 0.09$ and $p = 0.07$, respectively). No suprathreshold associations were found for subdivisions of the remaining seed ROIs, AC and STC (all $ps > 0.11$).

An important aspect to consider is whether the relationship between emotion recognition and IFG-motor system connectivity reflects emotion-specific processes, or rather domain-general cognitive abilities that might relate to task performance (e.g., working memory). This is relevant given the described contributions of the IFG for domain-general abilities (e.g., Niendam et al., 2012), particularly for working memory processes (e.g., Baldo and Dronkers, 2006; Courtney et al., 1997; McCarthy et al., 1996; Westerberg and Klingberg, 2007). Our models controlled for IQ (see Materials and Methods), but to further address this potential issue, we conducted a post-hoc analysis examining the extent to which the connectivity profiles could be accounted for by working memory abilities, as indexed by the digit span backwards task of WISC-III. A two-stage hierarchical multiple regression was performed, including connectivity between the left IFG and the motor system as the dependent variable. At stage one, digit span performance significantly contributed to the regression model ($F[1,53] = 10.34$, $p = 0.002$), accounting for 16% of the variability in functional connectivity. Importantly, though, entering emotion recognition performance at stage two explained an additional 23% of the variance, and this change in R^2 was statistically significant ($F[1,52] = 19.84$, $p < 0.001$), indicating a role for
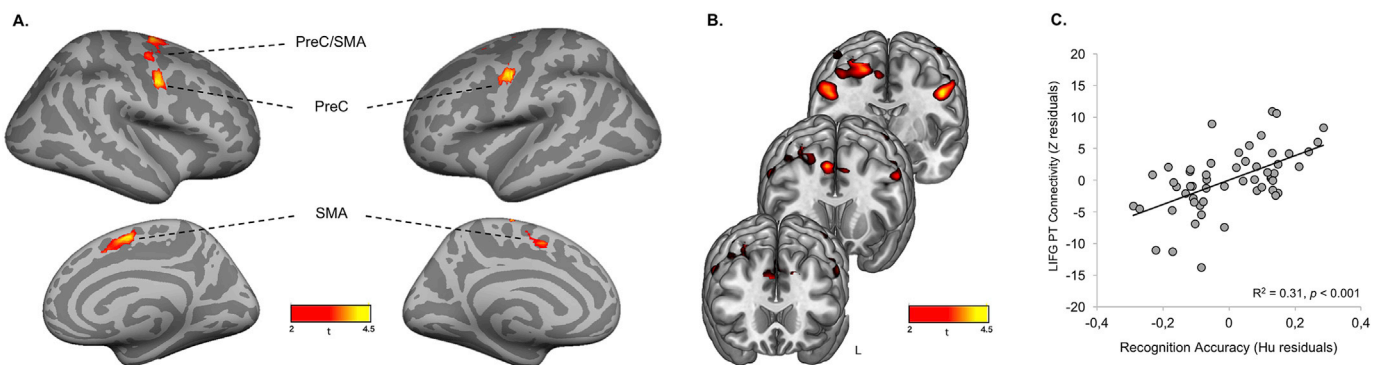
Fig. 4. Maps of motor system regions in which the strength of connectivity with the IFG *pars triangularis* was associated with emotional prosody recognition ability (follow-up analysis); thresholded activation maps ($p < 0.05$, TFCE-corrected) are registered to and displayed on a cortical surface using Freesurfer; the cortical surface was generated using an age-appropriate MNI template (see Materials and Methods) (A); Coronal slices revealing bilateral clusters within the premotor and supplementary motor areas (B); scatterplot denoting the association between the strength of IFG pt - motor system connectivity and emotion recognition accuracy (B). PreC, precentral gyrus; SMA, supplementary motor area.

emotion-specific processes over and above domain-general ones. We repeated this analysis for the right IFG, and found that working memory did not significantly contribute to the regression model ($F[1,53] = 2.58$, $p = 0.11$). However, when we included emotion recognition performance in the model, 27% of the variance in functional connectivity was further explained ($F[1,52] = 20.65$, $p < 0.001$).

Finally, no additional suprathreshold findings emerged in the exploratory whole-brain seed-based correlation analysis, aimed at identifying any other regions where connectivity with the seed ROIs (AC, STC and IFG) could relate to emotional prosody recognition.

3.2.3. Cross-validation and predictive power

We found robust associations between emotional prosody recognition and the connectivity between the IFG and motor system sites, but we also wanted to evaluate the predictive power of these results to assess whether they are likely to be replicated in new, out-of-sample data. A K-fold cross-validation was performed, focussing on the main results of the seed-based correlation analyses: seed regions left IFG, right IFG, and left IFG pars triangularis. The findings are summarized in Table 2. After five-fold repeated cross-validation, predicted r-squared values of 0.35, 0.34 and 0.35 were found for results seeded from the left IFG, the right IFG, and the left IFG pars triangularis, respectively. These values mimic the r-squared values found in the main analyses (r-squared range = 0.31–0.32), indicating that the original model was efficiently modelling the data, and showing minimal or no evidence of overfitting. Average NRMSE values were in the range of 20–23% of the residual variance, also indicating a good fit of the functional connectivity model in predicting emotion recognition scores. Taken together, cross-validation results lend strength to the predictive ability of our findings.

4. Discussion

The role of sensorimotor systems in auditory processing has been widely discussed, but typically regarding speech perception (e.g., Scott et al., 2009; Skipper et al., 2017), auditory imagery (e.g., Lima et al., 2015; Tian et al., 2016) and music perception (e.g., Grahn and Brett, 2007; Krishnan et al., 2018). Less is known about their involvement in vocal emotional communication, particularly regarding emotional speech. Addressing this question is important to further our understanding of the neural underpinnings of vocal emotional processing, and to develop an integrated account on the roles of sensorimotor systems in audition (Lima et al., 2016). Here, 8-year-old children completed a resting-state fMRI scan and an emotional prosody recognition task. First, we found that regions established to be key for emotional prosody processing, namely the AC, STC and IFG, are connected with each other. Second, consistent with our hypothesis, higher emotional prosody recognition skills related to a stronger connectivity between the bilateral IFG and sensorimotor cortices, including primary motor and medial and lateral premotor regions. Third, we established that this effect is

Table 2
Five-fold repeated cross-validation for the three regions of interest with significant association with emotional prosody recognition.

Seed ROIs	NRMSE	R² pred	MAE
L IFG			
Mean ± SD over 5 folds	20.8% ± 3.3%	0.35 ± 0.16	0.10 ± 0.02
Confidence interval (95%)	[19.8%–21.7%]	[0.31–0.40]	[0.10–0.11]
R IFG			
Mean ± SD over 5 folds	20.9% ± 3.1%	0.34 ± 0.20	0.11 ± 0.02
Confidence interval (95%)	[20.0%–21.8%]	[0.29–0.40]	[0.10–0.11]
L IFG pars triangularis			
Mean ± SD over 5 folds	20.9% ± 2.7%	0.35 ± 0.17	0.10 ± 0.02
Confidence interval (95%)	[20.1%–21.7%]	[0.31–0.40]	0.10–0.11

Each fold was repeated 10 times to obtain impartial cross-validation estimates. NRMSE, normalized root mean square error; R² pred, predicted r-squared; MAE, mean absolute error; L, left; R, right; IFG, inferior frontal gyrus.

particularly clear for the IFG pars triangularis, and it cannot be accounted for by differences in domain-general cognitive abilities.

Several studies reveal functional and structural connectivity among the frontotemporal regions that support emotional speech processing. Task-based functional connectivity analyses show that the left and right IFG have interconnections to bilateral voice-sensitive regions of the middle and posterior STC, and to primary and secondary AC; and there are intra- and inter-hemispheric connections among subregions along the STC, again also involving the primary and secondary auditory cortices (Ethofer et al., 2012; Frühholz and Grandjean, 2012; Leitman et al., 2010). Structural connectivity studies reveal connectivity between voice-sensitive STC areas and the IFG in both hemispheres, via pathways such as the superior longitudinal fasciculus, arcuate fasciculus, and inferior longitudinal fasciculus (Ethofer et al., 2012; Frühholz et al., 2015; Sammler et al., 2015). However, this body of work is based on adult samples. We show that similar frontotemporal connectivity extends to 8-year-old children and can be seen at rest, reflecting intrinsic functional coupling among these regions. In both hemispheres, our seeds were intra- and inter-hemispherically connected with each other. No previous fMRI studies have examined emotional speech processing in children, but the idea that this bilateral network is established early in development is consistent with several findings. Children at 5–8 years are already able to recognize positive and negative prosodic emotions (Chronaki et al., 2018; Sauter et al., 2013), a finding that our study corroborates. Furthermore, ERPs evidence indicates enhanced sensory processing of emotional prosody in 7-month-old infants (Grossmann et al., 2005), 3- to 7-month infants show specialized responses to voices in the anterior temporal cortex (Blasi et al., 2011), and brain regions associated with linguistic prosody are similar in children (aged 5–18 years) and adults (Plante et al., 2006). More generally, it is also known that the precursors of the adult cortical network for language are already active early in infancy (evidence from 3-month-old infants, Dehaene-Lambertz et al., 2002), with important aspects of functional selectivity and structural connectivity being established before the age of 10 (for a review of the early development of the speech/language network, Skeide and Friederici, 2016). The literature on emotional prosody in children remains scant, though (Morningstar et al., 2018), and the developmental trajectory of the vocal emotion network needs to be delineated in future studies, for instance by directly comparing participants at distinct developmental stages and by using task-based approaches.

That sensorimotor systems contribute to vocal emotional processing can be hypothesized from studies reporting activations in these systems in response to vocalizations such as laughter, both in adults (Bestelmeyer et al., 2014; Lima et al., 2015; Warren et al., 2006) and adolescents (O'Nions et al., 2017). However, several aspects remained unclear, namely whether such activations are epiphenomenal or functionally relevant for perceptual processes, how sensorimotor systems interact with core regions of the vocal emotion network, and whether their role extends to children and to emotional speech (for distinctions between prosody and nonverbal vocalizations, Pell et al., 2015; Scott et al., 2010). This study provides first evidence that individual differences in emotional prosody recognition can be traced to the strength of connectivity between the IFG and motor system sites in children. We have shown this using theoretically motivated, hypothesis-driven seed-based correlation analyses, and follow-up cross-validation tests that further attested the predictive ability of these findings. This indicates that not only sensorimotor systems are engaged during vocal emotional processing, but also that their involvement (in interaction with other nodes of the network) contributes to behaviour. We thus extend to speech prosody, and to a network approach, previous work with adults showing that suppressing sensorimotor activity impairs performance on a same-different discrimination task of vocalizations (Banissy et al., 2010). Our findings also extend those by McGettigan et al. (2015) showing that sensorimotor activations in response to laughter correlated with the ability to categorize laughter authenticity. Therefore, sensorimotor systems might provide a general mechanism that contributes to process

different vocal emotional signals and different features of such signals (emotion discrimination, authenticity detection, recognition of emotion categories). Importantly, we benefit from a task-free technique, reinforcing the notion that the motor system involvement can be specific to emotional/perceptual processes, rather than a reflection of task-related confounds (e.g., response preparation, button presses). It should be noted that our findings were observed for general emotion recognition performance, as well as for each emotion separately, as indicated by follow-up analyses. The only exception was sadness, for which the association was in the same direction but did not reach significance. There was less individual variation for sadness as compared to the other emotions (see Fig. 2A), and this could possibly explain why results were less clear for this emotion. This finding was unanticipated and warrants further investigation.

The dfunctional coupling between the IFG and the motor system is plausibly supported by anatomical connections between these regions. A direct fibre tract in both hemispheres, the frontal aslant tract, has been identified both in adults and children from the age of 5, and it runs from the IFG pars opercularis and triangularis to the boundary region between pre-SMA and SMA (Broce et al., 2015; Catani et al., 2012; Misaghi et al., 2018; Vergani et al., 2014). Consistent with this, our SMA peaks are located in this boundary region. There are also short U fibres connecting the IFG to lateral motor/premotor regions (Budisavljevic et al., 2017; Sammler et al., 2015) in sites resembling the ones in this study. The medial and lateral peaks that we found are additionally suggestive of a close link between perceptual and production mechanisms in the motor system. We know from the motor literature that there are somatotopically organized maps in SMA and in motor and lateral premotor cortices (e.g., Fontaine et al., 2002; Muakkassa and Strick, 1979; Zeharia et al., 2012), maps that can be seen already in the human preterm period (Dall'Orso et al., 2018). In SMA, the peak that we found corresponds to the site established to control the production of orofacial movements and vocalizations/speech (e.g., Fontaine et al., 2002; Fried et al., 1991; Kirzinger and Jürgens, 1982). Similarly, we found lateral motor system peaks in regions overlapping with those that control mouth movements (Agnew et al., 2011; Krishnan et al., 2018; Warren et al., 2006). Furthermore, combined perceptual and production effects have been reported in these medial and lateral sites for emotional vocalizations (Warren et al., 2006). Such tight perception-production coupling, along with its association with a behavioural advantage, provide support to the view that, during vocal emotional processing, sensorimotor systems mediate the activation of sound-related motor representations, that could guide perceptual and evaluative processes (Lima et al., 2016). They are also consistent with sensorimotor simulation accounts of emotion recognition (Banissy et al., 2010; McGettigan et al., 2015), which have been discussed for adults but also in the context of development (e.g., Pfeifer et al., 2008; Rayson et al., 2016).

The IFG is a central node of the vocal emotion network. According to the model by Schirmer and Kotz (2006), this region integrates emotionally-relevant auditory information provided by the STC and supports explicit evaluative judgements. This is confirmed by a review of the literature (Frühholz and Grandjean, 2013b), though the IFG might also support perception-production links. The posterior IFG, in particular, shows combined perception and production effects in the context of emotional vocalizations (Warren et al., 2006), consistent with its vicinity to primary motor and premotor cortices. In contrast, more anterior IFG regions – for which our results were particularly clear – appear to be preponderant for attentive-evaluative processes. Meta-analytic evidence indicates that the IFG pars triangularis is activated for explicit evaluations of vocal emotions, more for voices vs. faces and for explicit vs. passive/implicit perception (Dricu and Frühholz, 2016). Disrupting activity in the left and right anterior IFG impairs emotional prosody recognition (Hoekert et al., 2010), and learning to control activity in the IFG pars triangularis increases emotional prosody recognition ability (Rota et al., 2011). Here, the task emphasized explicit-evaluative processes. Plausibly, a more efficient integration of sensorimotor

information during the evaluation of prosodic stimuli – as supported by a stronger IFG-motor system coupling – could provide a mechanism to guide and optimize emotion recognition. An important consideration is whether such coupling reflects emotion-specific mechanisms or differences in domain-general cognitive abilities. We excluded this alternative account by including IQ as a covariate in all analyses and by examining the role of working memory in our pattern of results. Our work raises interesting questions for future research. It will be interesting to ask whether the involvement of sensorimotor systems is stronger when emotion recognition is more challenging, for instance by adding background noise. This has been observed for speech (e.g., Scott et al., 2004; Du et al., 2014), but it remains unknown for nonverbal communication. It will also be interesting to examine if and how our findings are modulated by age, that is, whether the IFG-motor system coupling (and its association with prosodic abilities) generalizes to adults, and whether it changes throughout development. Potential modulations related to musical experience might also exist, as musical expertise is associated with improved emotional prosody recognition (Lima and Castro, 2011) and a more efficient access to sensorimotor systems (Krishnan et al., 2018).

5. Conclusions

The current study forms the first demonstration that higher emotional prosody recognition in children relates to a stronger intrinsic connectivity between the IFG and medial and lateral motor system regions. This adds to the literature pointing to a role of sensorimotor activity during vocal emotional processing (e.g., Banissy et al., 2010; McGettigan et al., 2015; Warren et al., 2006), and suggests that sensorimotor and attentive/evaluative mechanisms interact to aid emotion recognition. Our findings also contribute to debates on the development of vocal emotional processing, and on the roles of the motor system for vocal communication (e.g., Sammler et al., 2015; Scott et al., 2009; Skipper et al., 2017) and for auditory processing more generally.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2019.116052.

Declarations of interest

The authors declare no competing financial interests.

References

Agnew, Z.K., McGettigan, C., Scott, S.K., 2011. Discriminating between auditory and motor cortical responses to speech and nonspeech mouth sounds. J. Cogn. Neurosci. 23 (12), 4038–4047. https://doi.org/10.1162/jocn_a_00106.

Aguert, M., Laval, V., Lacroix, A., Gil, S., Le Bigot, L., 2013. Inferring emotions from speech prosody: not so easy at age five. PLoS One 8 (12), e83657. https://doi.org/10.1371/journal.pone.0083657.

Allgood, R., Heaton, P., 2015. Developmental change and cross-domain links in vocal and musical emotion recognition performance in childhood. Dev. Psychol. 33 (3), 398–403. https://doi.org/10.1111/bjdp.12097.

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders, Fifth Ed. (DSM-5). American Psychiatric Association Inc., Washington DC.

Andersson, J.L.R., Jenkinson, M., Smith, S.M., 2007. Non-linear Registration, Aka Spatial Normalization (FMRIB Technical Report TR07JA2). Retrieved from. www.fmrib.ox.ac.uk/analysis/techrep.

Angelides, N.H., Gupta, J., Vickery, T.J., 2017. Associating resting-state connectivity with trait impulsivity. Soc. Cogn. Affect. Neurosci. 12 (6), 1001–1008. https://doi.org/10.1093/scan/nsx031.

Baldo, J.V., Dronkers, N.F., 2006. The role of inferior parietal and inferior frontal cortex in working memory. Neuropsychology 20 (5), 529–538. https://doi.org/10.1037/0894-4105.20.5.529.

Banissy, M.J., Sauter, D.A., Ward, J., Warren, J.E., Walsh, V., Scott, S.K., 2010. Suppressing sensorimotor activity modulates the discrimination of auditory emotions but not speaker identity. J. Neurosci. 30 (41), 13552–13557. https://doi.org/10.1523/JNEUROSCI.0786-10.2010.

Bestelmeyer, P., Maurage, P., Rouger, J., Latinus, M., Belin, P., 2014. Adaptation to vocal expressions reveals multistep perception of auditory emotion. J. Neurosci. 34 (24), 8098–8105. https://doi.org/10.1523/JNEUROSCI.4820-13.2014.

Blasi, A., Mercure, E., Lloyd-Fox, S., Thomson, A., Brammer, M., Sauter, D., Deeley, Q., Barker, G.J., Renvall, V., Deoni, S., Gasston, D., Williams, S.C.R., Johnson, M.H., Simmons, A., Murphy, D.G.M., 2011. Early specialization for voice and emotion processing in the infant brain. Curr. Biol. 21 (14), 1220–1224. https://doi.org/10.1016/j.cub.2011.06.009.

Broce, I., Bernal, B., Altman, N., Tremblay, P., Dick, A.S., 2015. Fiber tracking of the frontal aslant tract and subcomponents of the arcuate fasciculus in 5-8-year-old: relation to speech and language. Brain Lang. 149, 66–76. https://doi.org/10.1016/j.bandl.2015.06.006.

Brück, C., Kreifelts, B., Wildgruber, D., 2011. Emotional voices in context: a neurobiological model of multimodal affective information processing. Phys. Life Rev. 8 (4), 383–403. https://doi.org/10.1016/j.plrev.2011.10.002.

Budisavljevic, S., Dell'Acqua, F., Djordjilovic, V., Miotto, D., Motta, R., Castiello, U., 2017. The role of the frontal aslant tract and premotor connections in visually guided hand movement. Neuroimage 146, 419–428. https://doi.org/10.1016/j.neuroimage.2016.10.051.

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. Behav. Res. Methods 42 (1), 74–81. https://doi.org/10.3758/BRM.42.1.74.

Catani, M., Dell'Acqua, F., Vergani, F., Malik, F., Hodge, H., Roy, P., Valabregue, R., Thiebaut de Schotten, M., 2012. Short frontal lobe connections of the human brain. Cortex 48 (2), 273–291. https://doi.org/10.1016/j.cortex.2011.12.001.

Chronaki, G., Wigelsworth, M., Pell, M.D., Kotz, S.A., 2018. The development of cross-cultural recognition of vocal emotion during childhood and adolescence. Sci. Rep. 8 (8659), 1–17. https://doi.org/10.1038/s41598-018-26889-1.

Courtney, S.M., Ungerleider, L.G., Keil, K., Haxby, J.V., 1997. Transient and sustained activity in a distributed neural system for human working memory. Nature 386 (6625), 608–611. https://doi.org/10.1038/386608a0.

Curran, P.J., West, S.G., Finch, J.F., 1996. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. Psychol. Methods 1 (1), 16–29. http://doi.org/10.1037/1082-989X.1.1.16.

Dall'Orso, S., Steinweg, J., Allievi, A.F., Edwards, A.D., Burdet, E., Arichi, T., 2018. Somatotopic mapping of the developing sensorimotor cortex in the preterm human brain. Cerebr. Cortex 28 (7), 2507–2515. https://doi.org/10.1093/cercor/bhy050.

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

Demopoulos, C., Hopkins, J., Lewine, J.D., 2016. Relations between nonverbal and verbal social cognitive skills and complex social behavior in children and adolescents with autism. J. Abnorm. Child Psychol. 44 (5), 913–921. https://doi.org/10.1007/s10802-015-0082-z.

Dricu, M., Frühholz, S., 2016. Perceiving emotional expressions in others: activation likelihood estimation meta-analyses of explicit evaluation, passive perception and incidental perception of emotions. Neurosci. Biobehav. Rev. 71, 810–828. https://doi.org/10.1016/j.neubiorev.2016.10.020.

Du, Y., Buchsbaum, B.R., Grady, C.L., Alain, C., 2014. Noise differentially impacts phoneme representations in the auditory and speech motor systems. Proc. Natl. Acad. Sci. 111 (19), 7126–7131. https://doi.org/10.1073/pnas.1318738111.

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. Cerebr. Cortex 22 (1), 191–200. https://doi.org/10.1093/cercor/bhr113.

Flom, R., Bahrick, L.E., 2007. The development of infant discrimination of affect in multimodal and unimodal stimulation: the role of intersensory redundancy. Dev. Psychol. 43 (1), 238–252. https://10.1037/0012-1649.43.1.238.

Fonov, V.S., Evans, A.C., McKinstry, R.C., Almli, C.R., Collins, D.L., 2009. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. Neuroimage 47, S102. https://doi.org/10.1016/S1053-8119(09)70884-5.

Fonov, V., Evans, A.C., Botteron, K., Almli, C.R., McKinstry, R.C., Collins, D.L., 2011. Unbiased average age-appropriate atlases for pediatric studies. Neuroimage 54 (1), 313–327. https://doi.org/10.1016/j.neuroimage.2010.07.033.

Fontaine, D., Capelle, L., Duffau, H., 2002. Somatotopy of the supplementary motor area: evidence from correlation of the extent of surgical resection with the clinical patterns of deficit. Neurosurgery 50 (2), 297–305. https://doi.org/10.1097/00006123-200202000-00011.

Fried, I., Katz, A., McCarthy, G., Sass, K.J., Williamson, P., Spencer, S.S., Spencer, D.D., 1991. Functional organization of human supplementary motor cortex studied by electrical stimulation. J. Neurosci. 11 (11), 3656–3666. https://doi.org/10.1523/JNEUROSCI.11-11-03656.1991.

Frühholz, S., Grandjean, D., 2012. Towards a fronto-temporal neural network for the decoding of angry vocal expressions. Neuroimage 62 (3), 1658–1666. https://doi.org/10.1016/j.neuroimage.2012.06.015.

Frühholz, S., Grandjean, D., 2013a. Multiple subregions in superior temporal cortex are differentially sensitive to vocal expressions: a quantitative meta-analysis. Neurosci. Biobehav. Rev. 37 (1), 24–35. https://doi.org/10.1016/j.neubiorev.2012.11.002.

Frühholz, S., Grandjean, D., 2013b. Processing of emotional vocalizations in bilateral inferior frontal cortex. Neurosci. Biobehav. Rev. 37 (10Pt2), 2847–2855. https://doi.org/10.1016/j.neubiorev.2013.10.007.

Frühholz, S., Gschwind, M., Grandjean, D., 2015. Bilateral dorsal and ventral fiber pathways for the processing of affective prosody identified by probabilistic fiber tracking. Neuroimage 109, 27–34. https://doi.org/10.1016/j.neuroimage.2015.01.016.

Frühholz, S., Trost, W., Kotz, S.A., 2016. The sound of emotions — towards a unifying neural network perspective of affective sound processing. Neurosci. Biobehav. Rev. 68, 96–110. https://doi.org/10.1016/j.neubiorev.2016.05.002.

Grahn, J.A., Brett, M., 2007. Rhythm and beat perception in motor areas of the brain. J. Cogn. Neurosci. 19 (5), 893–906. https://doi.org/10.1162/jocn.2007.19.5.893.

Greve, D.N., Fischl, B., 2009. Accurate and robust brain image alignment using boundary-based registration. Neuroimage 48 (1), 63–72. https://doi.org/10.1016/j.neuroimage.2009.06.060.

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.

Hoekert, M., Vingerhoets, G., Aleman, A., 2010. Results of a pilot study on the involvement of bilateral inferior frontal gyri in emotional prosody perception: an rTMS study. BMC Neurosci. 11, 93. https://doi.org/10.1186/1471-2202-11-93.

Isaacowitz, D.M., Löckenhoff, C.E., Lane, R.D., Wright, R., Sechrest, L., Riedel, R., Costa, P.T., 2007. Age differences in recognition of emotion in lexical stimuli and facial expressions. Psychol. Aging 22 (1), 147–159. https://doi.org/10.1037/0882-7974.22.1.147.

JASP Team(2017). JASP (Version 0.8.5)[Computer software]

Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17 (2), 825–841. https://doi.org/10.1006/nimg.2002.1132.

Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., Smith, S.M., 2012. FSL. NeuroImage 62 (2), 782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015.

Kirzinger, A., Jürgens, U., 1982. Cortical lesion effects and vocalization in the squirrel monkey. Brain Res. 233 (2), 299–315. https://doi.org/10.1016/0006-8993(82)91204-5.

Kotz, S.A., Paulmann, S., 2011. Emotion, language, and the brain. Lang. Linguist. Compass. 5, 108–125. https://doi.org/10.1111/j.1749-818X.2010.00267.x.

Koyama, M.S., Di Martino, A., Zuo, X., Kelly, C., Mennes, M., Jutagir, D.R., Castellanos, F.X., Milham, M.P., 2011. Resting-state functional connectivity indexes reading competence in children and adults. J. Neurosci. 31 (23), 8617–8624. https://doi.org/10.1523/JNEUROSCI.4865-10.2011.

Krishnan, S., Lima, C.F., Evans, S., Chen, S., Guldner, S., Manly, T., Scott, S.K., 2018. Beatboxers and guitarists engage sensorimotor fields when listening to the instruments they can play. Cerebr. Cortex 28 (11), 4063–4079. https://doi.org/10.1093/cercor/bhy208.

Kuhn, M., 2008. Building predictive models in R using the caret package. J. Stat. Softw. 28 (5), 1–26. https://doi.org/10.18637/jss.v028.i05.

Lee, M.H., Smyser, C.D., Shimony, J.S., 2013. Resting-state fMRI: a review of methods and clinical applications. AJNR Am. J. Neuroradiol. 34 (10), 1866–1872. https://doi.org/10.3174/ajnr.A3263.

Leitman, D.I., Wolf, D.H., Ragland, J.D., Laukka, P., Loughead, J., Valdez, J.N., Javitt, D.C., Turetsky, B.I., Gur, R.C., 2010. "It's not what you say, but how you say it": a reciprocal temporo-frontal network for affective prosody. Front. Hum. Neurosci. 4 (19), 1–13. https://doi.org/10.3389/fnhum.2010.00019.

Leppänen, J.M., Hietanen, J.K., 2001. Emotion recognition and social adjustment in school-aged girls and boys. Scand. J. Psychol. 42 (5), 429–435. https://doi.org/10.1111/1467-9450.00255.

Lima, C.F., Castro, S.L., 2011. Speaking to the trained ear: musical expertise enhances the recognition of emotional speech prosody. Emotion 11 (5), 1021–1031. https://doi.org/10.1037/a0024521.

Lima, C.F., Garrett, C., Castro, S.L., 2013. Not all sounds sound the same: Parkinson's disease affects differently emotion processing in music and in speech prosody. J. Clin. Exp. Neuropsychol. 35 (4), 373–392. https://doi.org/10.1080/13803395.2013.776518.

Lima, C.F., Lavan, N., Evans, S., Agnew, Z., Halpern, A.R., Shanmugalingam, P., Meekings, S., Boebinger, D., Ostarek, M., McGettigan, C., Warren, J.E., Scott, S.K., 2015. Feel the noise: relating individual differences in auditory imagery to the structure and function of sensorimotor systems. Cerebr. Cortex 25 (11), 4638–4650. https://doi.org/10.1093/cercor/bhv134.

Lima, C.F., Krishnan, S., Scott, S.K., 2016. Roles of supplementary motor areas in auditory processing and auditory imagery. Trends Neurosci. 39 (8), 527–542. https://doi.org/10.1016/j.tins.2016.06.003.

Mayka, M.A., Corcos, D.M., Leurgans, S.E., Vaillancourt, D.E., 2006. Three-dimensional locations and boundaries of motor and premotor cortices as defined by functional brain imaging: a meta-analysis. Neuroimage 31 (4), 1453–1474. https://doi.org/10.1016/j.neuroimage.2006.02.004.

McCarthy, G., Puce, A., Constable, R.T., Krystal, J.H., Gore, J.C., Goldman-Rakic, P., 1996. Activation of human prefrontal cortex during spatial and nonspatial working

memory tasks measured by functional MRI. Cerebr. Cortex 6 (4), 600–611. https://doi.org/10.1093/cercor/6.4.600.

McClure, E.B., Nowicki Jr., S., 2001. Associations between social anxiety and nonverbal processing skill in preadolescent boys and girls. J. Nonverbal Behav. 25 (1), 3–19. https://doi.org/10.1023/A:1006753006870.

McGettigan, C., Walsh, E., Jessop, R., Agnew, Z.K., Sauter, D.A., Warren, J.E., Scott, S.K., 2015. Individual differences in laughter perception reveal roles for mentalizing and sensorimotor systems in the evaluation of emotional authenticity. Cerebr. Cortex 25 (1), 246–257. https://doi.org/10.1093/cercor/bht227.

Misaghi, E., Zhang, Z., Gracco, V.L., Nil, L.F., Beal, D.S., 2018. White matter tractography of the neural network for speech-motor control in children who stutter. Neurosci. Lett. 668, 37–42. https://doi.org/10.1016/j.neulet.2018.01.009.

Mollo, G., Karapanagiotidis, T., Bernhardt, B.C., Murphy, C.E., Smallwood, J., Jefferies, E., 2016. An individual differences analysis of the neurocognitive architecture of the semantic system at rest. Brain Cogn. 109, 112–123. https://doi.org/10.1016/j.bandc.2016.07.003.

Morningstar, M., Nelson, E.E., Dirks, M.A., 2018. Maturation of vocal emotion recognition: insights from the developmental and neuroimaging literature. Neurosci. Biobehav. Rev. 90, 221–230. https://doi.org/10.1016/j.neubiorev.2018.04.019.

Muakkassa, K.F., Strick, P.L., 1979. Frontal lobe inputs to primate motor cortex: evidence for four somatotopically organized 'premotor' areas. Brain Res. 177 (1), 176–182.

Neves, L., Cordeiro, C., Scott, S.K., Castro, S.L., Lima, C.F., 2018. High emotional contagion and empathy are associated with enhanced detection of emotional authenticity in laughter. Q. J. Exp. Psychol. 71 (11), 2355–2363. https://doi.org/10.1177/1747021817741800.

Niendam, T.A., Laird, A.R., Ray, K.L., Dean, Y.M., Glahn, D.C., Carter, C.S., 2012. Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. Cognit. Affect Behav. Neurosci. 12, 241–268. https://doi.org/10.3758/s13415-011-0083-5.

Nowicki Jr., S., Duke, M.P., 1992. The associations of children's nonverbal decoding abilities with their popularity, locus of control, and academic achievement. Res. Theory Human Development 153 (4), 385–393. https://doi.org/10.1080/00221325.1992.10753734.

Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9 (1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4.

O'Nions, E., Lima, C.F., Scott, S.K., Roberts, R., McCrory, E.J., Viding, E., 2017. Reduced laughter contagion in boys at risk for psychopathy. Curr. Biol. 27 (19), 3049–3055 e4. https://doi.org/10.1016/j.cub.2017.08.062.

Pell, M.D., Rothermich, K., Liu, P., Paulmann, S., Sethi, S., Rigoulot, S., 2015. Preferential decoding of emotion from human non-linguistic vocalizations versus speech prosody. Biol. Psychol. 111, 14–25. https://doi.org/10.1016/j.biopsycho.2015.08.008.

Pfeifer, J.H., Iacoboni, M., Mazziotta, J.C., Dapretto, M., 2008. Mirroring others' emotions relates to empathy and interpersonal competence in children. Neuroimage 39 (4), 2076–2085. https://doi.org/10.1016/j.neuroimage.2007.10.032.

Plante, E., Holland, S.K., Schmithorst, V.J., 2006. Prosodic processing by children: an fMRI study. Brain Lang. 97 (3), 332–342. https://dx.doi.org/10.1016/j.bandl.2005.12.004.

Pruim, R.H.R., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J.K., Beckmann, C.F., 2015. ICA-AROMA: a robust ICA-based strategy for removing motion artifacts from fMRI data. Neuroimage 112, 267–277. https://doi.org/10.1016/j.neuroimage.2015.02.064.

Rayson, H., Bonaiuto, J.J., Ferrari, P.F., Murray, L., 2016. Mu desynchronization during observation and execution of facial expressions in 30-month-old children. Devel. Cogn. Neurosci. 19, 279–287. https://doi.org/10.1016/j.dcn.2016.05.003.

Rota, G., Handjaras, G., Sitaram, R., Birbaumer, N., Dogil, G., 2011. Reorganization of functional and effective connectivity during real-time fMRI-BCI modulation of prosody processing. Brain Lang. 117 (3), 123–132. https://doi.org/10.1016/j.bandl.2010.07.008.

Ruffman, T., Murray, J., Halberstadt, J., Taumoepeau, M., 2010. Verbosity and emotion recognition in older adults. Psychol. Aging 25 (2), 492–497. https://doi.org/10.1037/a0018247.

Sammler, D., Grosbras, M.H., Anwander, A., Bestelmeyer, P.E., Belin, P., 2015. Dorsal and ventral pathways for prosody. Curr. Biol. 25 (23), 3079–3085. https://doi.org/10.1016/j.cub.2015.10.009.

Sauter, D.A., Panattoni, C., Happé, F., 2013. Children's recognition of emotions from vocal cues. Br. J. Dev. Psychol. 31 (1), 97–113. https://doi.org/10.1111/j.2044-835X.2012.02081.x.

Schirmer, A., 2018. Is the voice an auditory face? An ALE meta-analysis comparing vocal and facial emotion processing. Soc. Cogn. Affect. Neurosci. 13, 1–13. https://doi.org/10.1093/scan/nsx142.

Schirmer, A., Adolphs, R., 2017. Emotion perception from face, voice, and touch: comparisons and convergence. Trends Cogn. Sci. 21 (3), 216–228. https://doi.org/10.1016/j.tics.2017.01.001.

Schirmer, A., Kotz, S.A., 2006. Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. Trends Cogn. Sci. 10 (1), 24–30. https://doi.org/10.1016/j.tics.2005.11.009.

Scott, S.K., Rosen, S., Wickham, L., Wise, R.J., 2004. A positron emission tomography study of the neural basis of informational and energetic masking effects in speech perception. J. Acoust. Soc. Am. 115 (2), 813–821. https://doi.org/10.1121/1.1639336.

Scott, S.K., McGettigan, C., Eisner, F., 2009. A little more conversation, a little less action – candidate roles for the motor cortex in speech perception. Nat. Rev. Neurosci. 10 (4), 295–302. https://doi.org/10.1038/nrn2603.

Scott, S.K., Sauter, D., McGettigan, C., 2010. Brain mechanisms for processing perceived emotional vocalizations in humans. In: Stefan, M.B. (Ed.), Handbook of Behavioral Neuroscience. Academic Press, London, pp. 187–197.

Skeide, M.A., Friederici, A.D., 2016. The ontogeny of the cortical language network. Nat. Rev. Neurosci. 17, 323–332. https://doi.org/10.1038/nrn.2016.23.

Skipper, J.I., Devlin, J.T., Lametti, D.R., 2017. The hearing ear is always found close to the speaking tongue: review of the role of the motor system in speech perception. Brain Lang. 164, 77–105. https://doi.org/10.1016/j.bandl.2016.10.004.

Smith, S.M., 2002. Fast robust automated brain extraction. Hum. Brain Mapp. 17 (3), 143–155. https://doi.org/10.1002/hbm.10062.

Soderstrom, M., Reimchen, M., Sauter, D., Morgan, J.L., 2017. Do infants discriminate non-linguistic vocal expressions of positive emotions? Cognit. Emot. 31 (2), 298–311. https://doi.org/10.1080/02699931.2015.1108904.

Supekar, K., Swigart, A.G., Tenison, C., Jolles, D.D., Rosenberg-Lee, L.F., Menon, V., 2013. Neural predictors of individual differences in response to math tutoring in primary-grade school children. Proc. Natl. Acad. Sci. U.S.A. 110 (20), 8230–8235. https://doi.org/10.1073/pnas.1222154110.

Tian, X., Zarate, J.M., Poeppel, D., 2016. Mental imagery of speech implicates two mechanisms of perceptual reactivation. Cortex 77, 1–12. https://doi.org/10.1016/j.cortex.2016.01.002.

Vergani, F., Lacerda, L., Martino, J., Attems, J., Morris, C., Mitchell, P., Thiebaut de Schotten, M., Dell'Acqua, F., 2014. White matter connections of the supplementary motor area in humans. J. Neurol. Neurosurg. Psychiatry 85 (12), 1377–1385. https://doi.org/10.1136/jnnp-2013-307492.

Wagner, H.L., 1993. On measuring performance in category judgment studies of nonverbal behavior. J. Nonverbal Behav. 17 (1), 3–28. https://doi.org/10.1007/BF00987006.

Warren, J.E., Sauter, D.A., Eisner, F., Wiland, J., Dresner, M.A., Wise, R.J.S., Rosen, S., Scott, S.K., 2006. Positive emotions preferentially engage an auditory-motor "mirror" system. J. Neurosci. 26 (50), 13067–13075. https://doi.org/10.1523/JNEUROSCI.3907-06.2006.

Wechsler, D., 2003. Escala de Inteligência de Wechsler para Crianças, third ed. (WISC III). CEGOC-TEA, Lisbon.

Westerberg, H., Klingberg, T., 2007. Changes in cortical activity after training of working memory – a single-subject analysis. Physiol. Behav. 92 (1–2), 186–192. https://doi.org/10.1016/j.physbeh.2007.05.041.

Winkler, A.M., Ridgway, G.R., Webster, M.A., Smith, S.M., Nichols, T.E., 2014. Permutation inference for the general linear model. Neuroimage 92, 381–397. https://doi.org/10.1016/j.neuroimage.2014.01.060.

Zeharia, N., Hertz, U., Flash, T., Amedi, A., 2012. Negative blood oxygenation level dependent homunculus and somatotopic information in primary motor cortex and supplementary motor area. Proc. Natl. Acad. Sci. 109 (45), 18565–18570. https://doi.org/10.1073/pnas.1119125109.

Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imaging 20 (1), 45–57. https://doi.org/10.1109/42.906424.