

RESEARCH ARTICLE

Intonation processing increases task-specific fronto-temporal connectivity in tonal language speakers

Pei-Ju Chien^{1,2,3,4}  | Angela D. Friederici⁴  | Gesa Hartwigsen³  | Daniela Sammler² 

¹International Max Planck Research School NeuroCom, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

²Otto Hahn Group 'Neural Bases of Intonation in Speech and Music', Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

³Lise Meitner Research Group 'Cognition and Plasticity', Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

⁴Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

Correspondence

Pei-Ju Chien, International Max Planck Research School NeuroCom, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1a, 04103 Leipzig, Germany.

Email: chien@cbs.mpg.de

Abstract

Language comprehension depends on tight functional interactions between distributed brain regions. While these interactions are established for semantic and syntactic processes, the functional network of speech intonation – the linguistic variation of pitch – has been scarcely defined. Particularly little is known about intonation in tonal languages, in which pitch not only serves intonation but also expresses meaning via lexical tones. The present study used psychophysiological interaction analyses of functional magnetic resonance imaging data to characterise the neural networks underlying intonation and tone processing in native Mandarin Chinese speakers. Participants categorised either intonation or tone of monosyllabic Mandarin words that gradually varied between statement and question and between Tone 2 and Tone 4. Intonation processing induced bilateral fronto-temporal activity and increased functional connectivity between left inferior frontal gyrus and bilateral temporal regions, likely linking auditory perception and labelling of intonation categories in a phonological network. Tone processing induced bilateral temporal activity, associated with the auditory representation of tonal (phonemic) categories. Together, the present data demonstrate the breadth of the functional intonation network in a tonal language including higher-level phonological processes in addition to auditory representations common to both intonation and tone.

KEYWORDS

auditory categories, functional connectivity, lexical tone, phonology, pitch, prosody

1 | INTRODUCTION

Language processing is a complex cognitive ability unique to humans. Numerous functional and structural neuroimaging studies have consistently shown that language comprehension is supported by

interactions within large-scale fronto-temporo-parietal networks (Saur et al., 2008; for reviews, see Friederici, 2011; Hagoort, 2014; Hickok & Poeppel, 2007; Price, 2012). However, most of this work has focused primarily on semantic, phonological or syntactic aspects of language comprehension (Przeździk, Haak, Beckman, & Bartsch, 2019; Vigneau et al., 2006), while other linguistic functions have received relatively little attention. Here, we will focus on one of these aspects that are

Gesa Hartwigsen and Daniela Sammler shared senior authorship.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2020 The Authors. *Human Brain Mapping* published by Wiley Periodicals LLC.

important for successful language comprehension and shapes everyday communication: speech intonation.

Speech intonation, characterised by the pitch of the voice that changes over time, modulates linguistic information in numerous ways (Cole, 2015; Cutler, Dahan, & van Donselaar, 1997; Wagner & Watson, 2010). Apart from conveying a speaker's emotion or communicative intention, sometimes even at the single word level (e.g. Hellbernd & Sammler, 2016), intonational cues are crucial for parsing syntactic boundaries at the sentence level (Li & Yang, 2009; Snedeker & Trueswell, 2003; van der Burght, Goucha, Friederici, Kreitewolf, & Hartwigsen, 2019) and for understanding which aspects of a message are in focus (e.g. Dahan, Tanenhaus, & Chambers, 2002; Grice, Ritter, Niemann, & Roettger, 2017; Kristensen, Wang, Petersson, & Hagoort, 2013). At the neural level, linguistic intonation processing is characterised by increased activity in distributed fronto-temporo-parietal areas (Ischebeck, Friederici, & Alter, 2008; Kreitewolf, Friederici, & von Kriegstein, 2014; Meyer, Alter, Friederici, Lohmann, & von Cramon, 2002) forming a large-scale network supported by anatomical links in the right hemisphere (Sammler, Grosbras, Anwander, Bestelmeyer, & Belin, 2015).

However, the majority of previous studies focused on intonation processing in non-tonal languages (e.g. English or German). While some early studies have sought to identify brain areas for intonation perception in tonal languages such as Mandarin Chinese (hereafter Mandarin) (e.g. Gandour et al., 2003; Tong et al., 2005), the functional interactions between these regions remain largely unclear. More generally, the limited number of studies investigating intonation in tonal languages is surprising, given that pitch information has more versatile linguistic functions in tonal than non-tonal languages. The present study aims to fill this gap by investigating functional interactions in the neural network for intonation processing in Mandarin.

Crucially, in tonal languages, not only intonation but also lexical tone (hereafter tone) is conveyed by pitch information. In Mandarin, four tones differing in pitch height and contour are employed to contrast lexical meanings: A high level tone (Tone 1), a high rising tone (Tone 2, hereafter T2), a low falling-rising tone (Tone 3), and a high falling tone (Tone 4, hereafter T4) (Chao, 1968; Ladefoged & Johnson, 2011). Previous functional magnetic resonance imaging (fMRI) studies on tone processing have identified bilateral fronto-temporal (Kwok et al., 2017; Kwok, Dan, Yakpo, Matthews, & Tan, 2016) and fronto-parietal regions (Gandour et al., 2003, 2004) reflecting phonological and semantic processing of tone. More recent evidence suggests functional interactions between left frontal and temporal regions during tone processing in Mandarin sentences, associated with the semantic and phonological information in tonal pitch (Ge et al., 2015). Yet, unlike the current understanding of tone, the links between brain regions underlying intonation processing in tonal languages remain poorly understood. In particular, since tonal pitch may influence intonation processing, previous findings from intonation studies in non-tonal languages may not translate to tonal languages, and functional interactions between neural key regions may be different.

Previous work has often emphasised cognitive similarities between intonation and tone processing given that they both rely on pitch (Gandour et al., 2003, 2004; Liu, Chen, & Schiller, 2016; Yuan, 2011). At the neural level, both domains show overlap in left fronto-parietal regions, likely reflecting phonological processing of linguistic pitch contours (Gandour et al., 2003, 2004). Importantly, in a recent study with Mandarin speakers, we demonstrated the overlap of both processes in left fronto-parietal areas, namely, left inferior frontal gyrus (IFG) and supramarginal gyrus (SMG), but found a dissociation in right frontal regions (Chien, Friederici, Hartwigsen, & Sammler, 2020). More precisely, intonation involved right IFG more strongly than tone, likely reflecting the labelling of prosodic categories specific to intonation. What remained unclear in this study, however, was whether and how left and right frontal regions interact with temporal and parietal areas during processing of intonation and tone, that is, areas involved in general auditory processing of pitch, the gradual emergence of abstract phonemic (Levy & Wilson, 2020) or prosodic categories (Sammler et al., 2015), and phonological processing.

The present fMRI study was designed to address this question. Specifically, we explored task-related functional connectivity of frontal regions involved in intonation and tone processing in Mandarin speakers. To this end, we combined psychophysiological interaction (PPI) analyses (Friston et al., 1997) with a novel analysis of the fMRI data acquired in our previous study (Chien et al., 2020; gender task excluded). This study employed an audio-morphing paradigm that allows comparing brain activity during perception of stimuli with typical intonation or tone (i.e. *clear* exemplars) and stimuli with *ambiguous* pitch contours generated by means of pair-wise audio-morphing between the clear exemplars. This approach not only circumvents the need of a control task known to influence results (e.g. Kreitewolf et al., 2014; Luks, Nusbaum, & Levy, 1998). Furthermore, clear and ambiguous stimuli are fully matched in their average acoustic properties because they are derived from the same originals. Consequently, brain areas involved more strongly in clear than ambiguous intonation or tone should optimally reflect the processing of intonation categories and tones, controlled for low-level acoustics (for a similar approach with non-tonal language materials, see Hellbernd & Sammler, 2018). Based on these results, we conducted PPI analyses with seeds in bilateral IFG to delineate and compare task-related functional links in the fronto-temporo-parietal network involved in intonation and tone processing.

Our hypotheses were as follows. First, we expected that the comparison of clear > ambiguous intonation or tone should identify core regions of intonation and tone processing in IFG, superior/middle temporal gyrus (STG/MTG) and SMG, which should be bilateral for intonation (e.g. Chien et al., 2020) and left-lateralised for tone (e.g. Gandour et al., 2003). Second, we expected increased functional coupling between inferior frontal and temporo-parietal regions in clear compared to ambiguous stimuli, which should be bilateral or right-lateralised for intonation (e.g. Sammler et al., 2015), and left-lateralised for tone (e.g. Liang & Du, 2018).

2 | METHODS

2.1 | Participants

We present data from 24 healthy native Mandarin speakers (8 males, mean age 25.4 years, age range 21–31) who took part in the fMRI experiment reported in Chien et al. (2020). All participants were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971) (mean laterality quotient: 86.91, $SD = 14.78$), had normal hearing and normal or corrected-to-normal vision, and no history of neurological or psychiatric disorders or any contraindications against MRI. All participants gave written informed consent to participate in the experiment approved by the Ethics Committee of the University of Leipzig (126/18-ek).

2.2 | Experimental procedure and stimuli

A full description of the experimental procedure and stimuli can be found in our previous study (Chien et al., 2020). In short, in the fMRI experiment, participants were instructed to categorise stimuli in terms of intonation or tone in separate blocks in two-alternative forced-choice tasks (see Figure 1a). In intonation blocks, participants judged via button press whether the stimuli were spoken as statement or question (irrespective of whether they were spoken with T2 or T4). In tone blocks, they judged whether the stimuli were spoken with T2 or T4 (irrespective of whether they were spoken as statement or question). Block order was counterbalanced across participants. A short training session (10 min) prior to scanning was used to familiarise

participants with the tasks. Participants responded with their right index and middle finger. The two task-relevant categories were presented on the left and right side of the screen (e.g. statement and question in the intonation task) corresponding to the button assignment that was counterbalanced across participants. Each block (2 intonation, 2 tone) had 120 trials grouped into 8 mini-blocks with 15 trials each (see Figure 1a). Stimuli were presented with a jittered stimulus onset asynchrony (SOA) from 2.5 to 3.5 s (mean SOA = 3 s) in a pseudo-random order such that each morph step followed each morph step with similar probability. The first mini-block in each block started with a 5-s visual task instruction. After each mini-block, a 15-s pause was implemented during which only a fixation cross was presented. Stimulus presentation and response registration was controlled with Presentation software (Version 19.0, Neurobehavioural Systems, Inc., Berkeley, CA). The fMRI experiment lasted about 50 min. Stimuli were presented via MR-compatible headphones (MR confon GmbH, Magdeburg, Germany).

Experimental stimuli consisted of the monosyllabic Mandarin word 'bi' [International Phonetic Alphabet (pi:)] that varied along continua either in intonation (ranging from statement to question in five steps) or tone (ranging from T2 to T4, also in five steps; see Figure 1(b)). Details of stimulus recording and preparation (i.e. creation of stimulus continua with STRAIGHT; Kawahara, 2006) as well as the acoustic properties of the stimuli can be found in our previous study (Chien et al., 2020). Importantly, stimuli at the outer ends of the continua (i.e. Morph steps 1 and 5) were *clear* exemplars of the respective sound category (i.e. statement/question or T2/T4). Conversely, stimuli around the centre of the continua (i.e. Morph steps 2, 3 and 4) were *ambiguous* because they contained pitch information from both sides

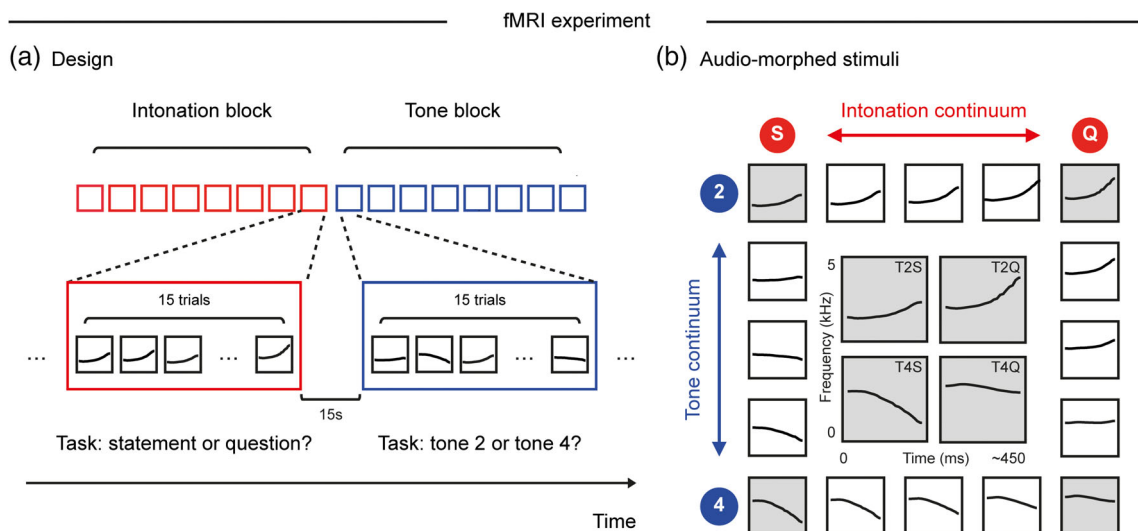


FIGURE 1 Experimental design and stimuli. (a) Participants performed an intonation (statement or question) and a tone task (T2 or T4) in separate blocks. Each block contained eight mini-blocks (see red and blue boxes) with 15 trials. Mini-blocks were separated by 15 s breaks. The scanning session lasted approximately 50 minutes. (b) Mandarin syllable 'bi' (IPA: [pi:]) spoken with Tone 2 (meaning 'nose') and with Tone 4 (meaning 'arm') was recorded as statement or question (4 central panels in grey). These stimuli were used to construct 5-step continua along two dimensions: Intonation (horizontal, red) and tone (vertical, blue). Each box in the continua illustrates the pitch contour of one morph step. Stimuli of the tone continua were randomly presented in the tone blocks; stimuli of the intonation continua in the intonation blocks. Q, question; S, statement; T2, Tone 2; T4, Tone 4

of the continua. Figure 1b shows the pitch contours of clear and ambiguous stimuli in the intonation and tone continua of one female speaker (extracted with Praat 6.0.53; Boersma & Weenink, 2019). We presented 16 continua from four speakers (2 females, 2 males) during the experiment.

2.3 | Data acquisition

Functional images were acquired with a 3T Siemens Magnetom Prisma scanner (Siemens AG, Erlangen, Germany) with a multi-band echo-planar imaging sequence (TR = 2000 ms, TE = 23.2 ms, multi-band acceleration factor = 3, 60 slices in axial direction and interleaved order, thickness = 2.5 mm, 10% inter-slice gap, field of view = 192 mm, voxel size = $2 \times 2 \times 2.5$ mm, flip angle = 90°) (Feinberg et al., 2010; Moeller et al., 2010) and a 32-channel head coil. Anatomical T1-weighted images were taken from the brain database of the Max Planck Institute for Human Cognitive and Brain Sciences Leipzig, or additionally acquired using a standard magnetisation-prepared rapid acquisition gradient echo (MPRAGE) sequence in sagittal orientation (whole brain coverage, voxel size = 1 mm isotropic, field of view = 256 mm, TR = 2,300 ms, TE = 2.98 ms, flip angle = 9°).

2.4 | Behavioural data analysis

Participants' average response times (RTs) and response consistencies for clear (i.e. Morph steps 1 and 5) and ambiguous stimuli (i.e. Morph steps 2, 3, and 4) were analysed. Response consistencies were calculated as the mean absolute deviation of 'statement' or 'question', and 'T2' or 'T4' response proportions from 50%, separately for clear and ambiguous morph steps in each participant (50% corresponds to the lowest, 100% to the highest possible response consistency in a two-alternative forced-choice task). Note that we did not analyse slopes or other parameters of psychometric curves fitted to the data (cf. Chien et al., 2020) because such analysis does not capture behavioural differences between clear and ambiguous stimuli. Using paired-samples *t* tests with Bonferroni-correction, RTs and response consistency data were compared between clear and ambiguous stimuli for intonation and tone tasks separately. Statistical analyses were conducted using SPSS (PASW) Statistics 21.0 (SPSS Inc., Chicago, IL) and customised scripts in MatlabR2019 (The MathWorks, Inc., Natick, MA).

2.5 | fMRI data analysis

fMRI data were preprocessed and analysed with SPM 12 (Wellcome Trust Centre for Neuroimaging, London, UK). Preprocessing steps included slice timing correction, realignment, segmentation, coregistration of the functional and anatomical images, normalisation into the Montreal Neurological Institute (MNI) stereotactic space, and smoothing using a Gaussian kernel of 8 mm full width at half maximum.

A general linear model (GLM) was estimated for each participant. In the GLM, onset regressors representing each morph step in the tasks (i.e. five morph steps in intonation and in tone) were convolved with a canonical hemodynamic response function. RTs were included as a duration-modulated parametric regressor orthogonalised to the stimulus onset regressors to account for differences in task difficulty (Grinband, Wager, Lindquist, Ferrera, & Hirsch, 2008). Finally, onsets of visual task instructions and six motion parameters were modelled as nuisance regressors. Four T-contrasts for comparisons of interest (i.e. clear > ambiguous intonation, ambiguous > clear intonation, clear > ambiguous tone, ambiguous > clear tone) were calculated for each participant and subjected to one-sample *t* tests at the second level. All comparisons were thresholded using a cluster-forming threshold of $p < .001$ uncorrected at the voxel level and a family-wise error (FWE) correction of $p < .05$ at the cluster level. Anatomical locations were identified based on the Jülich probabilistic cytoarchitectonic maps in the SPM Anatomy Toolbox 2.2b (Eickhoff et al., 2005) and the anatomical Harvard-Oxford atlas in FSL (Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012). fMRI results were visualised using Mango (Research Imaging Institute, UT Health Science Center at San Antonio, TX; <http://ric.uthscsa.edu/mango/>) with the ch2better template from MRICron (Rorden & Brett, 2000). Percent signal change in 5 mm spherical regions of interest around selected peak voxels were extracted using rfxplot (Gläscher, 2009) for illustration purposes.

2.6 | PPI analyses

Four PPI analyses were performed to identify stimulus-dependent changes in functional coupling between left or right IFG and the rest of the brain during the processing of intonation and tone. Left and right IFG were defined as seed regions because they showed increased activity during intonation and tone processing, in line with previous studies (e.g. Chien et al., 2020; Gandour et al., 2003, 2004).

To define the IFG seed regions (VOIs, volumes of interest), an anatomical mask combining premotor cortex (PMC) and the larger IFG region (pars opercularis and pars triangularis) was created with the Harvard-Oxford atlas in FSL (Jenkinson et al., 2012). This mask was used to confine the centre of VOIs. Within the mask, spheres with 6 mm radius were drawn around the nearest local peak of each participant relative to the group peak voxels in clear > ambiguous intonation (left IFG: $x, y, z = -34, 12, 30$; right IFG: $x, y, z = 40, 10, 28$) and in ambiguous > clear tone (left IFG: $x, y, z = -54, 18, 22$; right IFG: $x, y, z = 50, 14, 24$). We chose coordinates from two different contrasts because bilateral IFG activity differed between intonation (clear > ambiguous) and tone (ambiguous > clear) tasks. Voxels within the 6 mm spheres that showed activity at a threshold of $p < .05$ (uncorrected) were included in the VOIs. Participants were excluded if their local peak voxel was not included in the anatomical mask. Accordingly, the resulting sample size for the PPI in the intonation

task was $n = 21$ for left and $n = 18$ for right IFG, while the PPI in the tone task included $n = 22$ for left and $n = 21$ participants for right IFG.

In each participant, the first eigenvariate of the fMRI signal change in the VOIs was extracted, and their mean time course was multiplied with the time course of the respective experimental condition (i.e. clear > ambiguous intonation or ambiguous > clear tone). This interaction term of VOI signal time course and experimental condition was modelled as the regressor of interest in the PPI. Additionally, the mean deconvolved VOI time course and regressor of experimental condition were included in the model as covariates of no interest. In each PPI, a T-contrast for the psychophysiological interaction term was calculated for each participant and subjected to one-sample *t*-tests at the second level. Results were thresholded using a cluster-forming threshold of $p < .001$ uncorrected at the voxel level and a FWE correction of $p < .05$ at the cluster level. Anatomical labelling and visualisation of results was done in the same way as in the fMRI whole-brain analysis.

3 | RESULTS

3.1 | Behavioural data

Group-averaged RTs and response consistencies are depicted in Figure 2. For RTs, we found that participants processed clear stimuli significantly faster than ambiguous stimuli in both the intonation [clear: mean \pm SEM = 982 \pm 121 ms, ambiguous: 1107 \pm 135 ms, $t(23) = -6.949$, $p < .001$] and the tone task [clear: 858 \pm 126 ms, ambiguous: 1040 \pm 147 ms, $t(23) = -12.200$, $p < .001$]. Similar patterns were found for response consistency, indicated by the mean absolute deviation of 'statement', 'question', 'T2' and 'T4' response proportions from 50% (see dashed lines in Figure 2b). We found that response consistency was higher for clear than ambiguous stimuli in both the intonation [clear: 90.1 \pm 6.2%, ambiguous: 81.5 \pm 5.6%, $t(23) = 8.586$, $p < .001$] and the tone task [clear: 98 \pm 3.8%, ambiguous: 93 \pm 4.2%, $t(23) = 9.001$, $p < .001$].

3.2 | fMRI data

In intonation (Figure 3a, Table 1), the contrast of clear > ambiguous stimuli revealed a stronger involvement of bilateral fronto-temporal regions, including IFG/PMC/middle frontal gyrus (MFG), Heschl's gyrus (HG)/planum temporale (PT), aSTG and pSTG. In the right hemisphere, activity extended into the aMTG. Furthermore, increased activity was found in the bilateral cerebellum, basal ganglia (caudate, pallidum) and right thalamus. Finally, stronger activity was also identified in the precuneus (PCun)/posterior cingulate cortex (PCC), right angular gyrus (AG), left postcentral gyrus (PoCG), midline calcarine gyrus/cuneus and lingual gyrus, and bilateral middle occipital gyrus. The opposite contrast, ambiguous > clear intonation, did not show significant activity.

In tone (Figure 3b, Table 2), the comparison of clear > ambiguous stimuli showed stronger activity in bilateral temporal regions, including HG/PT and right pSTG. Bilateral cerebellum as well as the PCun/PCC, right AG, left PoCG, midline cuneus/occipital pole (OP) and lingual gyrus were also activated. The opposite contrast (Figure 4, Table 3), ambiguous > clear tone, revealed stronger activity in bilateral IFG/MFG, and left anterior insula.

3.3 | PPI data

The PPI analyses with seeds in bilateral IFG showed significant increases in functional connectivity during the processing of clear (compared to ambiguous) intonation stimuli. Specifically, the left IFG seed showed significantly increased functional coupling with bilateral HG/PT, bilateral pSTG, right aSTG, pre-supplementary motor area (preSMA), PCun, middle cingulate cortex, PoCG and fusiform gyrus (Figure 5, Table 4). The right IFG seed showed significantly increased connectivity only with the preSMA (at a more liberal p cluster = .060, FWE-corrected) (Figure 5, Table 4). The opposite comparison of ambiguous > clear intonation did not reveal any significant results, even after decreasing the threshold to an exploratory $p < .001$.

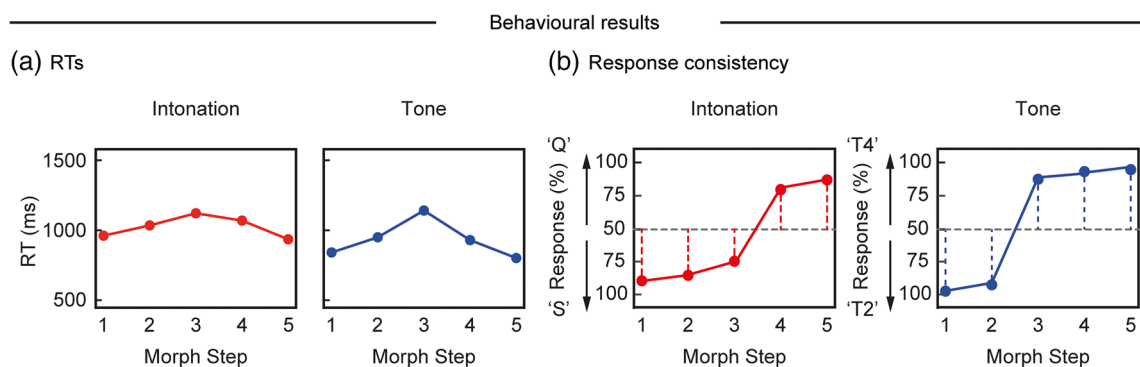


FIGURE 2 Behavioural results. (a) Group-averaged response times (RTs) and (b) proportions of question/statement and T2/T4 responses relative to 50% in the intonation (red) and tone task (blue) showed faster RTs and higher response consistency in categorisation of clear stimuli (Morph steps 1 and 5) compared to ambiguous stimuli (Steps 2–4). Q, question; S, statement

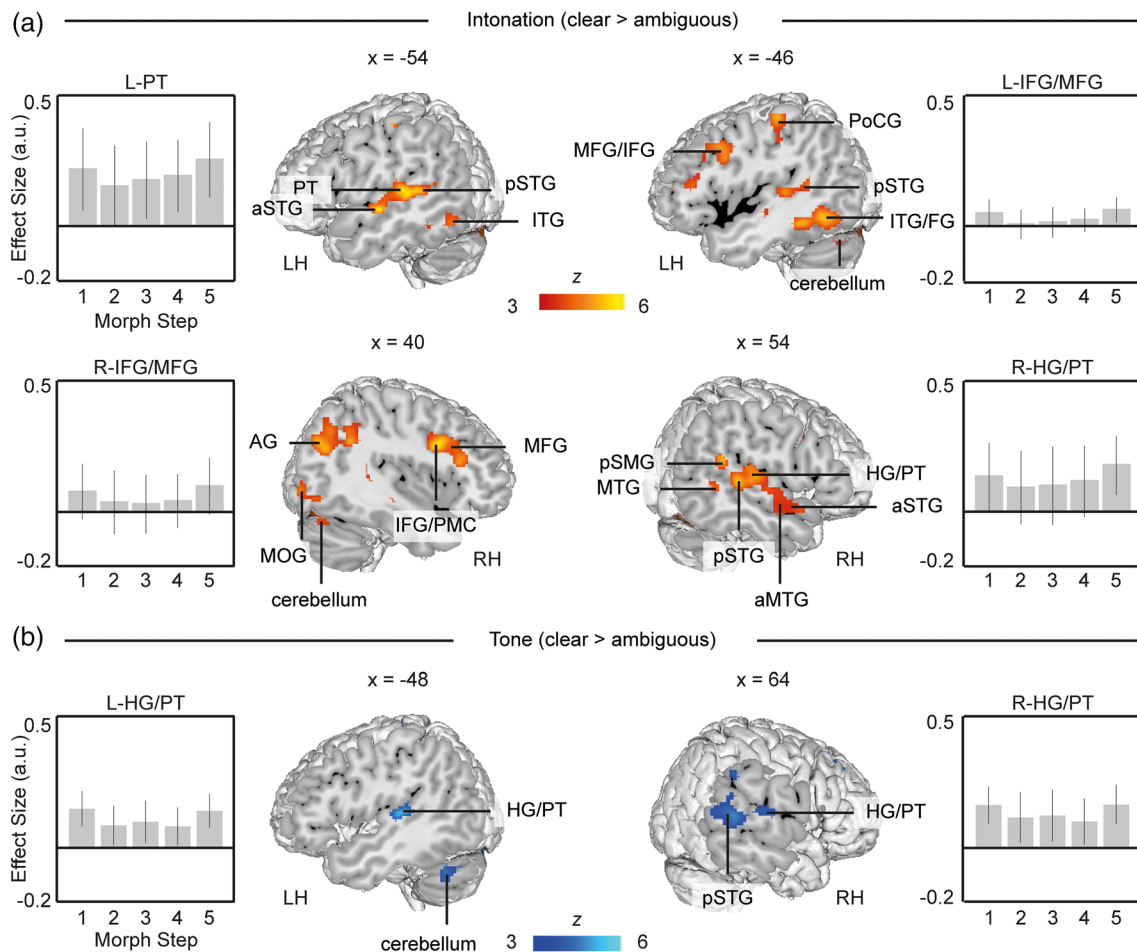


FIGURE 3 Comparison of clear > ambiguous stimuli in (a) the intonation task, and (b) the tone task. All clusters are thresholded at p cluster $<.05$ (FWE-corrected). Parameter estimates are shown for selected frontotemporal clusters. Error bars represent ± 1 SEM. FWE, family-wise error; LH, left hemisphere; RH, right hemisphere

uncorrected. No significant changes in functional coupling were found for either seed region in the tone task.

4 | DISCUSSION

The present study investigated functional interactions in the neural networks for intonation and tone processing in tonal language speakers. We first contrasted brain activity in Mandarin speakers during the categorisation of stimuli with clear versus ambiguous intonation or tone pitch contours, and then performed PPI analyses with bilateral IFG as seeds in both domains. In intonation (clear relative to ambiguous), we found bilateral fronto-temporal activity and increased functional connectivity between left IFG and bilateral temporal regions. This neural network may link emerging auditory representations of intonation categories in the temporal lobe with explicit prosodic labels in the frontal lobe, both forming the phonological processing network in Mandarin Chinese. Increased functional coupling during clear intonation was also found between bilateral IFG and preSMA, which may reflect verbal response planning. In tone (clear

relative to ambiguous), activity was limited to bilateral temporal regions, likely reflecting auditory (phonemic) representations of tone categories. Together, the present data provide novel insights into the neural bases of intonation and tone processing in Mandarin speakers by showing (a) similar auditory (categorical) perception in both domains in superior temporal regions, and (b) contribution of higher-level phonological and verbal response preparation processes in intonation involving fronto-temporal and lateral-medial frontal neural coupling.

4.1 | Fronto-temporal interactions during processing of clear intonation stimuli

The perception of intonation (clear relative to ambiguous) in Mandarin increased (a) activity in fronto-temporal regions, including bilateral IFG/PMC (extending to MFG), HG/PT, pSTG and aSTG, and (b) functional connectivity between left IFG and bilateral temporal areas.

Superior temporal areas have frequently been linked to the acoustic processing of prosody in non-tonal languages (for reviews,

TABLE 1 Comparison of clear > ambiguous intonation

Region	BA	k	z value	MNI coordinates		
				x	y	z
R precuneus (PCun)	7	18,988	5.35	14	-60	34
R posterior cingulate cortex (PCC)	31		4.75	16	-44	36
L intraparietal sulcus	39		4.86	-28	-58	38
L superior parietal lobule	7		4.72	-6	-68	56
L calcarine gyrus	17		4.82	-8	-82	12
R lingual gyrus	18		4.82	22	-86	-2
L angular gyrus (AG)	39		4.91	-30	-64	34
L postcentral gyrus (PoCG)	4		4.64	-34	-26	64
L precentral gyrus (PrCG)	4		4.63	-34	-22	60
L cerebellum (VIIa)	-		4.86	-32	-68	-46
L caudate	-	4,560	5.09	-6	10	0
L thalamus (Thal)	-		4.78	-22	-28	6
R thalamus (Thal)	-		4.34	18	-28	8
R pallidum	-		4.67	24	-4	6
L planum temporale (PT)	41		4.73	-58	-28	10
L posterior superior temporal gyrus (pSTG)	22		4.27	-54	-36	8
L anterior superior temporal gyrus (aSTG)	22		4.51	-54	-10	-4
R posterior superior temporal gyrus (pSTG)	22	1803	4.97	64	-28	4
R Heschl's gyrus (HG)	41		4.26	52	-28	4
R planum temporale (PT)	41		4.41	62	-16	10
R anterior superior temporal gyrus (aSTG)	22		3.69	60	-6	-2
R temporal pole	38		3.66	58	10	-12
R anterior middle temporal gyrus (aMTG)	22		3.32	50	-6	-16
R posterior supramarginal gyrus (pSMG)	22		4.57	50	-40	18
R middle temporal gyrus (MTG)	22		3.67	56	-44	0
L inferior frontal gyrus (pars opercularis) (IFG op.)/middle frontal gyrus (MFG)	9	1,264	5.25	-34	12	30
L middle frontal gyrus (MFG)	9		3.97	-46	14	38
L inferior frontal gyrus (pars opercularis) (IFG op.)/premotor cortex (PMC)/middle frontal gyrus (MFG)	9		3.85	-48	10	32
L Rolandic operculum	-		3.80	-40	2	18
L inferior frontal gyrus (pars triangularis) (IFG tri.)/frontal pole	46		4.57	-40	36	10
L frontal pole	10		3.41	-38	48	10
R inferior frontal gyrus (pars opercularis) (IFG op.)/premotor cortex (PMC)	6	711	4.42	40	10	28
R inferior frontal gyrus (pars opercularis) (IFG op.)/middle frontal gyrus (MFG)	9		3.89	46	20	34

Note: Peak voxels in clusters are highlighted in bold. All clusters are thresholded at p cluster < .05, FWE-corrected. Anatomical labelling for the cerebellum was based on the Jülich probabilistic cytoarchitectonic maps in the SPM Anatomy Toolbox 2.2b (Eickhoff et al., 2005).

Abbreviations: BA, Brodmann area; k, cluster size (number of voxels); L, left hemisphere; R, right hemisphere.

see Baum & Pell, 1999; Belyk & Brown, 2014; Schirmer & Kotz, 2006), and are in line with recent descriptions of a pathway for prosody along the (right) temporal lobe (Sammler et al., 2015). Notably, the average acoustic properties of clear and ambiguous stimuli did not differ in our study. This leads us to infer that the observed activity

pattern reflects representations of intonation categories instead of low-level processing of acoustic features. This interpretation is supported by previous studies associating these regions with category-selective responses for natural complex sounds (Leaver & Rauschecker, 2010), spoken language (Norman-Haignere,

TABLE 2 Comparison of clear > ambiguous tone

Region	BA	k	z value	MNI coordinates		
				x	y	z
R precuneus (PCun)	7	5,704	5.73	4	-52	68
L precuneus (PCun)	7		3.94	-6	-68	56
R posterior cingulate cortex (PCC)	31		4.83	6	-40	36
L posterior cingulate cortex (PCC)	31		4.50	-10	-34	44
R occipital pole (OP)/cuneus	18		4.02	2	-90	12
R fusiform gyrus (FG)	37	2,531	4.59	30	-32	-24
L fusiform gyrus (FG)	37		3.73	-26	-38	-22
R cerebellum (VIIa crus II)	-		3.89	6	-78	-34
L cerebellum (IX)	-		4.43	-14	-44	-44
L cerebellum (VII)	-		3.99	-30	-38	-40
Mid cerebellum (VI)	-		3.95	0	-82	-18
Mid cerebellum (VIIa crus II)	-		3.75	0	-80	-32
L cerebellum (VIIa crus I)	-		3.72	-28	-80	-22
R lingual gyrus	-		3.71	6	-74	-10
R posterior supramarginal gyrus (pSMG)	22	1,214	4.53	52	-40	20
	40		3.73	60	-38	28
R Heschl's gyrus (HG)	41		4.40	50	-18	4
R planum temporale (PT)	42		3.56	64	-14	10
R posterior superior temporal gyrus (pSTG)	42		3.97	70	-28	8
	22		3.89	64	-36	8
R middle temporal gyrus (MTG)/angular gyrus (AG)	39		3.51	60	-50	12
R middle frontal gyrus (MFG)	8	711	4.38	32	34	44
R superior frontal gyrus	6		3.79	24	18	46
L Heschl's gyrus (HG)	41	530	4.96	-42	-24	6
L planum temporale (PT)	42		4.25	-62	-26	14
L postcentral gyrus (PoCG)	4	504	4.09	-36	-24	64
	2		3.93	-32	-32	70
R angular gyrus (AG)	39	438	4.42	42	-70	32
L middle frontal gyrus^a (MFG)	9	236	3.89	-30	36	42
L superior frontal gyrus	10		3.47	-30	44	34

Note: Peak voxels in clusters are highlighted in bold. . Anatomical labelling for the cerebellum was based on the Jülich probabilistic cytoarchitectonic maps in the SPM Anatomy Toolbox 2.2b (Eickhoff et al., 2005).

Abbreviations: BA, Brodmann area; k, cluster size (number of voxels); L, left hemisphere; R, right hemisphere.

^aAll clusters are thresholded at p cluster < .05, FWE-corrected, except for the cluster of L middle frontal gyrus (p cluster = .097, FWE-corrected).

Kanwisher, & McDermott, 2015; Obleser & Eisner, 2009), as well as intonation in non-tonal languages (Hellbernd & Sammler, 2018; for review, see Schirmer & Kotz, 2006).

Bilateral IFG/PMC are other regions frequently reported in intonation research, both with non-tonal (Kreitewolf et al., 2014) and tonal language materials (Chien et al., 2020; Gandour et al., 2003, 2004). They have been associated with explicit labelling of prosodic categories (Sammler et al., 2015) and phonological processing of linguistic pitch contours (Chien et al., 2020; for review, see also Belyk & Brown, 2014), in line with previous work on phonological processing in the posterior IFG (Hartwigsen et al., 2016).

Importantly, we found strong functional interactions between left IFG and bilateral temporal regions during the processing of clear intonation. This finding suggests dynamic interactions between auditory perception of intonation categories and higher order phonological processes. Our data represent the first characterisation of a phonological intonation network in a tonal language. It mirrors connectivity reported in previous studies with non-tonal language materials (Saur et al., 2010; Tyler & Marslen-Wilson, 2008; Xiang, Fonteijn, Norris, & Hagoort, 2010). Structurally, these connections may rest on large intra- and interhemispheric fibre bundles such as the arcuate/superior longitudinal fascicle previously associated with phonological processes

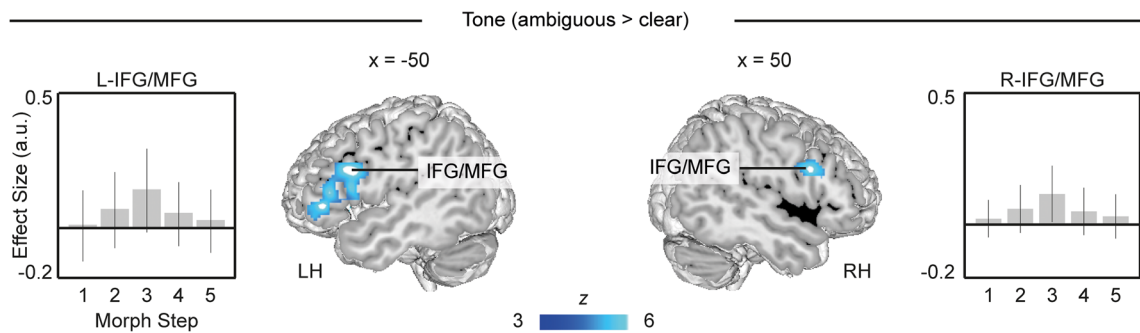


FIGURE 4 Comparison of ambiguous > clear stimuli in the tone task. All clusters are thresholded at p cluster < .05 (FWE-corrected). Parameter estimates are shown for left and right IFG/MFG. Error bars represent ± 1 SEM. FWE, family-wise error; IFG, inferior frontal gyrus; MFG, middle frontal gyrus

TABLE 3 Comparison of ambiguous > clear tone

Region	BA	k	z value	MNI coordinates		
				x	y	z
L inferior frontal gyrus (pars opercularis) (IFG op.)/middle frontal gyrus (MFG)	9	1726	4.89	-54	18	22
	44		4.01	-50	16	12
L inferior frontal gyrus (pars triangularis) (IFG tri.)	45		4.54	-48	34	0
L inferior frontal gyrus (pars orbitalis) (IFG orb.)	47		4.40	-44	32	-6
L anterior insula	13		4.87	-30	20	-2
R inferior frontal gyrus (pars opercularis) (IFG op.)/middle frontal gyrus (MFG)	9	309	4.19	50	14	24

Note: Peak voxels in clusters are highlighted in bold. All clusters are thresholded at p -cluster < .05, FWE-corrected. Abbreviations: BA, Brodmann area; k, cluster size (number of voxels); L, left hemisphere; R, right hemisphere.

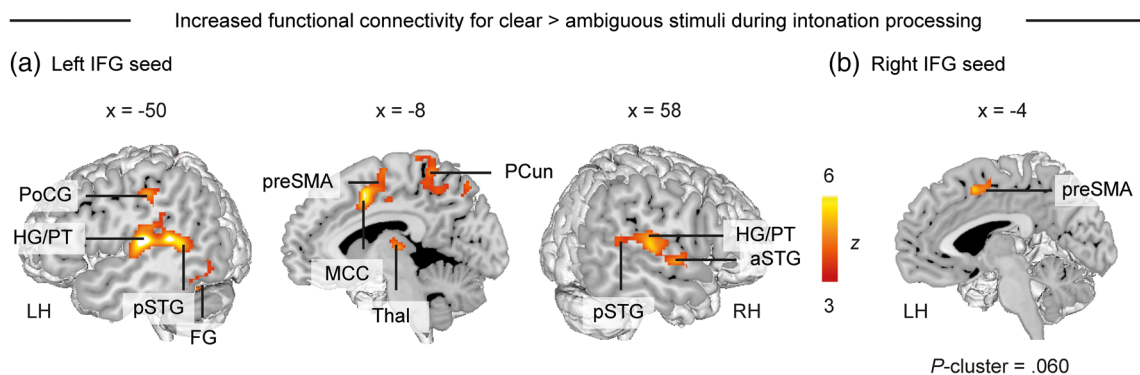


FIGURE 5 Increased functional connectivity of (a) left and (b) right IFG during the processing of clear > ambiguous intonation stimuli. All clusters are thresholded at p -cluster < .05 (FWE-corrected) unless otherwise indicated. FWE, family-wise error; IFG, inferior frontal gyrus

(Glasser & Rilling, 2008; Saur et al., 2008), and the posterior corpus callosum interconnecting left and right temporal lobes (Friederici & Alter, 2004; Friederici, von Cramon, & Kotz, 2007; Sammler et al., 2018; Sammler, Kotz, Eckstein, Ott, & Friederici, 2010).

Surprisingly, we found no increase of functional connectivity between right IFG and temporal regions, even though the right IFG showed a significant activity increase for clear compared to ambiguous stimuli. This finding suggests subtle functional differences

between left and right IFG in intonation processing revealed through their distinct functional connectivity profiles, not their activity patterns. Overall, this lateralisation pattern is in line with functional lateralisation hypotheses of prosody perception (van Lancker, 1980; Wildgruber et al., 2004). According to these models, lateralisation in an overall bilateral system (Chien et al., 2020; Kreitewolf et al., 2014; see also Belyk & Brown, 2014) is modulated by the *linguistic function* of pitch. Correspondingly, stronger connectivity of the left IFG for

Region	BA	k	z value	MNI coordinates		
				x	y	z
Left IFG seed						
L Heschl's gyrus (HG)	41	5,044	5.18	-38	-30	8
L planum temporale (PT)	22		4.77	-50	-38	10
L posterior superior temporal gyrus (pSTG)	22		4.02	-62	-32	10
L pre-supplementary motor area (preSMA)	6		4.74	-4	6	48
R pre-supplementary motor area (preSMA)	6		4.56	4	-4	48
	6		4.12	10	2	56
R Heschl's gyrus (HG)	41	1,572	4.58	52	-20	8
R planum temporale (PT)	41		3.91	50	-34	12
R anterior superior temporal gyrus (aSTG)	22		3.90	58	-2	-4
R posterior superior temporal gyrus (pSTG)	21		3.78	54	-40	10
R thalamus (Thal)	-		4.06	20	-26	12
L thalamus (Thal)	-		3.88	-8	-18	6
L precuneus (PCun)	7	1,370	4.53	-20	-56	44
R precuneus (PCun)	7		3.80	2	-68	54
L posterior cingulate cortex (PCC)	31		3.63	-16	-40	40
L postcentral gyrus (PoCG)/precuneus (PCun)	7		3.87	-10	-48	66
L postcentral gyrus (PoCG)	4		3.48	-8	-32	78
L fusiform gyrus (FG)	20	349	3.93	-42	-40	-22
L inferior occipital gyrus	19		3.26	-52	-64	-6
L inferior temporal gyrus (ITG)	37		3.24	-48	-62	-14
L anterior insula^a	13	187	4.30	-30	18	6
R fusiform gyrus^a (FG)	37	165	4.62	40	-42	-26
R cerebellum (VIIa crus I)	-		3.29	46	-54	-28
Right IFG seed						
L pre-supplementary motor area^a (preSMA)	6	210	4.03	-4	10	50

TABLE 4 Functional connectivity of left and right IFG in clear > ambiguous intonation

Note: Peak voxels in clusters are highlighted in bold. Clusters are thresholded at p -cluster < .05, FWE-corrected. Anatomical labelling for the cerebellum was based on the Jülich probabilistic cytoarchitectonic maps in the SPM Anatomy Toolbox 2.2b (Eickhoff et al., 2005).

^aClusters reported with lower FWE-correction: L anterior insula: $p = .062$, R fusiform gyrus: $p = .091$, and L pre-supplementary motor area: $p = .060$.

Abbreviations: BA, Brodmann area; k, cluster size (number of voxels); L, left hemisphere; R, right hemisphere.

clear than ambiguous stimuli may emerge because only clear question and statement intonations are part of the phonological inventory of Mandarin, and may fulfil a syntactic role. In contrast, a more general role of the right-hemispheric network in processing pitch contours irrespective of their linguistic relevance (van der Burght et al., 2019), including music (e.g. Bianco et al., 2016), may account for the absence of right fronto-temporal connectivity differences between clear and ambiguous intonation stimuli. Notably, the modulation of hemispheric lateralisation by the linguistic function of intonation has originally been described for non-tonal languages (Perkins, Baran, & Gandour, 1996; van Lancker, 1980; Wildgruber et al., 2004), suggesting neural similarities of intonation processing across tonal and non-tonal languages (Chien et al., 2020; Gandour et al., 2003, 2004).

Beyond these asymmetric fronto-temporal interactions, both left and right IFG showed increased connectivity to preSMA when intonation was clear. This coupling cannot be due to greater processing effort known to intensify crosstalk between IFG and preSMA (Hampshire, Chamberlain, Monti, Duncan, & Owen, 2010; Hellbernd & Sammler, 2018) because the behavioural data show that participants were able to categorise clear intonation more easily than ambiguous stimuli. In fact, the preSMA is an interface region between prefrontal and motor systems that is thought to facilitate motor responses to perceived (speech) sounds, such as sub-vocal articulation (Lima, Krishnan, & Scott, 2016). Furthermore, connectivity between IFG and preSMA via the frontal aslant tract has been associated with verbal fluency and speech initiation (Catani et al., 2013; Dragoy et al., 2020), leading us to suggest that our result

reflects the preparation of a verbal response. Future studies should further explore the role of the IFG-preSMA interaction in intonation perception and production, and its generalisability across tonal and non-tonal languages.

4.2 | Temporal contributions to tone processing

We found increased temporal activity in clear (relative to ambiguous) tone, while frontal activity was stronger in ambiguous (relative to clear) tone, in the absence of fronto-temporal interactions.

The perception of clear (relative to ambiguous) tone increased activity in temporal but not frontal areas, including bilateral HG and pSTG. Superior temporal activity is in line with previous studies that associated these regions with the acoustic and phonological processing of tonal pitch (Kwok et al., 2016; Si, Zhou, & Hong, 2017; Zhang et al., 2011; for meta-analyses, see Kwok et al., 2017; Liang & Du, 2018). Importantly, the tight acoustic control of our clear > ambiguous contrast (see Section 4.1) further suggests that these temporal regions support the representation of tonal categories beyond low-level acoustic features of speech sounds. The absence of frontal activity in clear tone was unexpected and contrasts with previous findings in the tone literature (Kwok et al., 2016; Si et al., 2017; for meta-analyses, see Kwok et al., 2017; Liang & Du, 2018). This apparent discrepancy probably emerges from the use of different task designs and contrasts. Accordingly, in the present study, frontal activity for clear tone processing may have been subtracted out in the direct comparison with ambiguous tone, which might also rely on frontal resources.

Indeed, we found stronger frontal lobe involvement in the opposite contrast. Comparing ambiguous against clear tone revealed strong activity in bilateral IFG/MFG and the left anterior insula. Similar activity patterns have been previously associated with increased ambiguity in prosodic (Feng, Gan, Wang, Wong, & Chandrasekaran, 2018; Hellbernd & Sammler, 2018; Leitman et al., 2010) or phonetic categories (Blumstein, Myers, & Rissman, 2005), possibly reflecting greater cognitive control in case of conflicting acoustic cues. Conflict may have been particularly high in our ambiguous tone stimuli with flat contour (e.g. step 3 in Figure 1b) because they could be perceived as yet another tone category in Mandarin (i.e. Tone 1, a high level tone) but had to be labelled as either Tone 2 or Tone 4 in our study. This interpretation would also fit with models that ascribe domain-general processes to IFG, especially when the experimental manipulations are complex and cognitively demanding (e.g. Duncan, 2010; Fedorenko & Blank, 2020; Novick, Trueswell, & Thompson-Schill, 2005). More generally, the opposite response profile of frontal regions in our tone and intonation tasks fits well with a recent proposal suggesting language-specific versus domain-general functional subdivisions in IFG (Fedorenko & Blank, 2020). Together, the present data allow to refine our previous findings (Chien et al., 2020) and suggest that the observed overlap of tone and intonation in left frontal regions may have included different cognitive processes instead of completely shared phonological processing.

4.3 | Parietal activity for processing intonation and tone

Apart from fronto-temporal areas, our analyses also showed stronger activity in parietal brain regions, including PCun/PCC, AG and left PrCG/PoCG, when contrasting clear > ambiguous stimuli, in both intonation and tone. Effects in the PCun/PCC and AG were driven by less deactivation during clear than ambiguous stimuli. The observed deactivation likely reflects less effortful processing for clear stimuli in regions associated with the default mode network, known to reflect the modulation of mental resources required by a goal-directed task (for reviews, see Buckner, Andrews-Hanna, & Schacter, 2008; Hartwigsen, 2018).

The absence of activation in the left SMG was unexpected in comparison to our earlier findings (Chien et al., 2020). Previous work highlights the role of the inferior parietal cortex in phonological processing (e.g. Gandour et al., 2004; Hartwigsen et al., 2017, 2016; Kreitewolf et al., 2014). These apparent inconsistencies may be attributed to the types of contrasts employed across studies. Comparing intonation or tone categorisation versus gender categorisation (Chien et al., 2020), intonation discrimination vs. speaker discrimination (Kreitewolf et al., 2014), and intonation or tone discrimination vs. passive listening (Gandour et al., 2004) may primarily tax attention to phonological vs. non-phonological aspects of the stimuli. However, attention to phonology may be comparably high when categorising clear and ambiguous intonation or tone such that the parietal activity was contrasted out in the direct comparison.

5 | CONCLUSIONS

In sum, our findings show that intonation processing in Mandarin speakers involved bilateral fronto-temporal areas and functional coupling between left IFG and bilateral temporal regions that may bind auditory intonation perception with prosodic category labelling in a phonological processing network. Furthermore, intonation processing enhanced functional interactions between bilateral IFG and preSMA that may reflect the automatic preparation of a verbal response. Tonal processing was limited to temporal regions in the present study, likely reflecting the auditory representation of tone categories.

In conclusion, this study extends our current understanding of the functional dynamics of intonation processing in tonal language speakers by showing that intonation processing includes higher-level phonological processes and verbal response preparation in fronto-temporal and lateral-medial frontal networks, together with the categorical perception in temporal regions also involved when processing tone. Future studies should employ dynamic causal modelling to investigate the direction of information flow between fronto-temporal and different frontal regions during intonation processing.

ACKNOWLEDGEMENTS

The authors would like to thank all the participants, and Manuela Hofmann, Nicole Pampus, Domenica Wilfling, and Simone Wipper for

MRI data acquisition. This study was funded by the Max Planck Society.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon request. Authors can confirm that all relevant data are included in the article.

ETHICS STATEMENT

The experiment was approved by the Ethics Committee of the University of Leipzig (126/18-ek).

ORCID

Pei-Ju Chien  <https://orcid.org/0000-0001-7810-107X>

Angela D. Friederici  <https://orcid.org/0000-0002-6328-865X>

Gesa Hartwigsen  <https://orcid.org/0000-0002-8084-1330>

Daniela Sammler  <https://orcid.org/0000-0001-7458-0229>

REFERENCES

- Baum, S. R., & Pell, M. D. (1999). The neural bases of prosody: Insights from lesion studies and neuroimaging. *Aphasiology*, 13(8), 581–608. <https://doi.org/10.1080/026870399401957>
- 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. <https://doi.org/10.1093/scan/nst124>
- Bianco, R., Novembre, G., Keller, P. E., Kim, S. G., Scharf, F., Friederici, A. D., ... Sammler, D. (2016). Neural networks for harmonic structure in music perception and action. *NeuroImage*, 142, 454–464. <https://doi.org/10.1016/j.neuroimage.2016.08.025>
- Blumstein, S. E., Myers, E. B., & Rissman, J. (2005). The perception of voice onset time: An fMRI investigation of phonetic category structure. *Journal of Cognitive Neuroscience*, 17(9), 1353–1366. <https://doi.org/10.1162/0898929054985473>
- Boersma, P., & Weenink, D. (2019). Praat: doing phonetics by computer [Computer program]. Version 6.0.53, Retrieved from <http://www.praat.org/>.
- Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, 1124, 1–38. <https://doi.org/10.1196/annals.1440.011>
- Catani, M., Mesulam, M. M., Jakobsen, E., Malik, F., Martersteck, A., Wieneke, C., ... Rogalski, E. (2013). A novel frontal pathway underlies verbal fluency in primary progressive aphasia. *Brain*, 136, 2619–2628. <https://doi.org/10.1093/brain/awt163>
- Chao, Y. R. (1968). *A grammar of spoken Chinese*. Berkeley, CA: University of California Press.
- Chien, P.-J., Friederici, A. D., Hartwigsen, G., & Sammler, D. (2020). Neural correlates of intonation and lexical tone in tonal and non-tonal language speakers. *Human Brain Mapping*, 41(7), 1842–1858.
- Cole, J. (2015). Prosody in context: A review. *Language, Cognition and Neuroscience*, 30, 1–31. <https://doi.org/10.1080/23273798.2014.963130>
- Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40(2), 141–201. <https://doi.org/10.1177/002383099704000203>
- Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language*, 47, 292–314. [https://doi.org/10.1016/S0749-596X\(02\)00001-3](https://doi.org/10.1016/S0749-596X(02)00001-3)
- Dragoy, O., Zyryanov, A., Bronov, O., Gordeyeva, E., Gronskaya, N., Kryuchkova, O., ... Zuev, A. (2020). Functional linguistic specificity of the left frontal aslant tract for spontaneous speech fluency: Evidence from intraoperative language mapping. *Brain and Language*, 208, 104836. <https://doi.org/10.1016/j.bandl.2020.104836>
- Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, 14(4), 172–179. <https://doi.org/10.1016/j.tics.2010.01.004>
- Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, 25(4), 1325–1335. <https://doi.org/10.1016/j.neuroimage.2004.12.034>
- Fedorenko, E., & Blank, I. A. (2020). Broca's area is not a natural kind. *Trends in Cognitive Sciences*, 24(4), 270–284. <https://doi.org/10.1016/j.tics.2020.01.001>
- Feinberg, D. A., Moeller, S., Smith, S. M., Auerbach, E., Ramanna, S., Glasser, M. F., ... Yacoub, E. (2010). Multiplexed echo planar imaging for sub-second whole brain fMRI and fast diffusion imaging. *PLoS One*, 5(12), e15710. <https://doi.org/10.1371/journal.pone.0015710>
- Feng, G., Gan, Z., Wang, S., Wong, P. C. M., & Chandrasekaran, B. (2018). Task-general and acoustic-invariant neural representation of speech categories in the human brain. *Cerebral Cortex*, 28(9), 3241–3254. <https://doi.org/10.1093/cercor/bhx195>
- Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews*, 91(4), 1357–1392. <https://doi.org/10.1152/physrev.00006.2011>
- Friederici, A. D., & Alter, K. (2004). Lateralization of auditory language functions: A dynamic dual pathway model. *Brain and Language*, 89(2), 267–276.
- Friederici, A. D., von Cramon, D. Y., & Kotz, S. A. (2007). Role of the corpus callosum in speech comprehension: Interfacing syntax and prosody. *Neuron*, 53, 135–145. <https://doi.org/10.1016/j.neuron.2006.11.020>
- Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *NeuroImage*, 6(3), 218–229. <https://doi.org/10.1006/nimg.1997.0291>
- Gandour, J., Dziedzic, M., Wong, D., Lowe, M., Tong, Y., Hsieh, L., ... Lurito, J. (2003). Temporal integration of speech prosody is shaped by language experience: An fMRI study. *Brain and Language*, 84(3), 318–336.
- Gandour, J., Tong, Y., Wong, D., Talavage, T., Dziedzic, M., Xu, Y., ... Lowe, M. (2004). Hemispheric roles in the perception of speech prosody. *NeuroImage*, 23(1), 344–357.
- Ge, J., Peng, G., Lyu, B., Wang, Y., Zhuo, Y., Niu, Z., ... Gao, J. H. (2015). Cross-language differences in the brain network subserving intelligible speech. *Proceedings of the National Academy of Sciences of the United States of America*, 112, 2972–2977. <https://doi.org/10.1073/pnas.1416000112>
- Gläscher, J. (2009). Visualization of group inference data in functional neuroimaging. *Neuroinformatics*, 7, 73–82. <https://doi.org/10.1007/s12021-008-9042-x>
- Glasser, M. F., & Rilling, J. K. (2008). DTI tractography of the human brain's language pathways. *Cerebral Cortex*, 18, 2471–2482. <https://doi.org/10.1093/cercor/bhn011>
- Grice, M., Ritter, S., Niemann, H., & Roettger, T. B. (2017). Integrating the discreteness and continuity of intonational categories. *Journal of Phonetics*, 64, 90–107. <https://doi.org/10.1016/j.jwocn.2017.03.003>
- Grinband, J., Wager, T. D., Lindquist, M., Ferrera, V. P., & Hirsch, J. (2008). Detection of time-varying signals in event-related fMRI designs. *NeuroImage*, 43, 509–520. <https://doi.org/10.1016/j.neuroimage.2008.07.065>

- Hagoort, P. (2014). Nodes and networks in the neural architecture for language: Broca's region and beyond. *Current Opinion in Neurobiology*, 28, 136–141. <https://doi.org/10.1016/j.conb.2014.07.013>
- Hampshire, A., Chamberlain, S. R., Monti, M. M., Duncan, J., & Owen, A. M. (2010). The role of the right inferior frontal gyrus: Inhibition and attentional control. *NeuroImage*, 50, 1313–1319. <https://doi.org/10.1016/j.neuroimage.2009.12.109>
- Hartwigsen, G. (2018). Flexible redistribution in cognitive networks. *Trends in Cognitive Sciences*, 22(8), 687–698. <https://doi.org/10.1016/j.tics.2018.05.008>
- Hartwigsen, G., Bzdok, D., Klein, M., Wawrzyniak, M., Stockert, A., Wrede, K., ... Saur, D. (2017). Rapid short-term reorganization in the language network. *eLife*, 6, e25964. <https://doi.org/10.7554/eLife.25964>
- Hartwigsen, G., Weigel, A., Schuschan, P., Siebner, H. R., Weise, D., Classen, J., & Saur, D. (2016). Dissociating parieto-frontal networks for phonological and semantic word decisions: A condition-and-perturb TMS study. *Cerebral Cortex*, 26(6), 2590–2601.
- Hellbernd, N., & Sammler, D. (2016). Prosody conveys speaker's intentions: Acoustic cues for speech act perception. *Journal of Memory and Language*, 88, 70–86. <https://doi.org/10.1016/j.jml.2016.01.001>
- Hellbernd, N., & Sammler, D. (2018). Neural bases of social communicative intentions in speech. *Social Cognitive and Affective Neuroscience*, 13(6), 604–615. <https://doi.org/10.1093/scan/nsy034>
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8, 393–402. <https://doi.org/10.1038/nrn2113>
- Ischebeck, A. K., Friederici, A. D., & Alter, K. (2008). Processing prosodic boundaries in natural and hummed speech: An fMRI study. *Cerebral Cortex*, 18(3), 541–552. <https://doi.org/10.1093/cercor/bhm083>
- Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, 62(2), 782–790. <https://doi.org/10.1016/j.neuroimage.2011.09.015>
- Kawahara, H. (2006). STRAIGHT, exploitation of the other aspect of VOCODER: Perceptually isomorphic decomposition of speech sounds. *Acoustical Science and Technology*, 27(6), 349–353. <https://doi.org/10.1250/ast.27.349>
- Kreitewolf, J., Friederici, A. D., & von Kriegstein, K. (2014). Hemispheric lateralization of linguistic prosody recognition in comparison to speech and speaker recognition. *NeuroImage*, 102, 332–344. <https://doi.org/10.1016/j.neuroimage.2014.07.038>
- Kristensen, L. B., Wang, L., Petersson, K. M., & Hagoort, P. (2013). The interface between language and attention: Prosodic focus marking recruits a general attention network in spoken language comprehension. *Cerebral Cortex*, 23(8), 1836–1848.
- Kwok, V. P. Y., Dan, G., Yakpo, K., Matthews, S., Fox, P. T., Li, P., & Tan, L.-H. (2017). A meta-analytic study of the neural systems for auditory processing of lexical tones. *Frontiers in Human Neuroscience*, 11, 375. <https://doi.org/10.3389/fnhum.2017.00375>
- Kwok, V. P. Y., Dan, G., Yakpo, K., Matthews, S., & Tan, L. H. (2016). Neural systems for auditory perception of lexical tones. *Journal of Neurolinguistics*, 37, 34–40.
- Ladefoged, P., & Johnson, K. (2011). *A course in phonetics* (6th ed.). Belmont: Thomson Wadsworth. <https://doi.org/10.1080/07268600600885494>
- Leaver, A. M., & Rauschecker, J. P. (2010). Cortical representation of natural complex sounds: Effects of acoustic features and auditory object category. *Journal of Neuroscience*, 30(22), 7604–7612. <https://doi.org/10.1523/JNEUROSCI.0296-10.2010>
- Leitman, D. I., Wolf, D. H., Ragland, J. D., Laukka, P., Loughhead, J., Valdez, J. N., ... Gur, R. C. (2010). 'It's not what you say, but how you say it': A reciprocal temporo-frontal network for affective prosody. *Frontiers in Human Neuroscience*, 4, 19. <https://doi.org/10.3389/fnhum.2010.00019>
- Levy, D. F., & Wilson, S. M. (2020). Categorical encoding of vowels in primary auditory cortex. *Cerebral Cortex*, 30(2), 618–627. <https://doi.org/10.1093/cercor/bhz112>
- Li, W., & Yang, Y. (2009). Perception of prosodic hierarchical boundaries in Mandarin Chinese sentences. *Neuroscience*, 158, 1416–1425. <https://doi.org/10.1016/j.neuroscience.2008.10.065>
- Liang, B., & Du, Y. (2018). The functional neuroanatomy of lexical tone perception: An activation likelihood estimation meta-analysis. *Frontiers in Neuroscience*, 12, 495. <https://doi.org/10.3389/fnins.2018.00495>
- Lima, C. F., Krishnan, S., & Scott, S. K. (2016). Roles of supplementary motor areas in auditory processing and auditory imagery. *Trends in Neurosciences*, 39, 527–542. <https://doi.org/10.1016/j.tins.2016.06.003>
- Liu, M., Chen, Y., & Schiller, N. O. (2016). Online processing of tone and intonation in Mandarin: Evidence from ERPs. *Neuropsychologia*, 91, 307–317.
- Luks, T. L., Nusbaum, H. C., & Levy, J. (1998). Hemispheric involvement in the perception of syntactic prosody is dynamically dependent on task demands. *Brain and Language*, 65, 313–332. <https://doi.org/10.1006/brln.1998.1993>
- Meyer, M., Alter, K., Friederici, A. D., Lohmann, G., & von Cramon, D. Y. (2002). fMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. *Human Brain Mapping*, 17(2), 73–88.
- Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E., Strupp, J., Harel, N., & Uğurbil, K. (2010). Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magnetic Resonance in Medicine*, 63(5), 1144–1153. <https://doi.org/10.1002/mrm.22361>
- Norman-Haignere, S., Kanwisher, N. G., & McDermott, J. H. (2015). Distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition. *Neuron*, 88, 1281–1296. <https://doi.org/10.1016/j.neuron.2015.11.035>
- Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience*, 5(3), 263–281. <https://doi.org/10.3758/CABN.5.3.263>
- Obleser, J., & Eisner, F. (2009). Pre-lexical abstraction of speech in the auditory cortex. *Trends in Cognitive Sciences*, 13, 14–19. <https://doi.org/10.1016/j.tics.2008.09.005>
- 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](https://doi.org/10.1016/0028-3932(71)90067-4)
- Perkins, J. M., Baran, J. A., & Gandour, J. (1996). Hemispheric specialization in processing intonation contours. *Aphasiology*, 10(4), 343–362. <https://doi.org/10.1080/02687039608248416>
- Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage*, 62, 816–847. <https://doi.org/10.1016/j.neuroimage.2012.04.062>
- Przeźdźik, I., Haak, K. V., Beckman, C. F., & Bartsch, A. (2019). The human language connectome. In P. Hagoort (Ed.), *Human language: From genes and brains to behavior* (pp. 467–480). London: The MIT Press.
- Rorden, C., & Brett, M. (2000). Stereotaxic display of brain lesions. *Behavioural Neurology*, 12(4), 191–200. <https://doi.org/10.1155/2000/421719>
- Sammler, D., Cunitz, K., Gierhan, S. M. E., Anwender, A., Adermann, J., Meixensberger, J., & Friederici, A. D. (2018). White matter pathways for prosodic structure building: A case study. *Brain and Language*, 183, 1–10. <https://doi.org/10.1016/j.bandl.2018.05.001>
- Sammler, D., Grosbras, M. H., Anwender, A., Bestelmeyer, P. E. G., & Belin, P. (2015). Dorsal and ventral pathways for prosody. *Current Biology*, 25(23), 3079–3085.
- Sammler, D., Kotz, S. A., Eckstein, K., Ott, D. V. M., & Friederici, A. D. (2010). Prosody meets syntax: The role of the corpus callosum. *Brain*, 133(9), 2643–2655. <https://doi.org/10.1093/brain/awq231>

- Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vrya, M. S., ... Weiller, C. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 18035–18040. <https://doi.org/10.1073/pnas.0805234105>
- Saur, D., Schelter, B., Schnell, S., Kratochvil, D., Küpper, H., Kellmeyer, P., ... Weiller, C. (2010). Combining functional and anatomical connectivity reveals brain networks for auditory language comprehension. *NeuroImage*, 49, 3187–3197. <https://doi.org/10.1016/j.neuroimage.2009.11.009>
- Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences*, 10, 24–30. <https://doi.org/10.1016/j.tics.2005.11.009>
- Si, X., Zhou, W., & Hong, B. (2017). Cooperative cortical network for categorical processing of Chinese lexical tone. *Proceedings of the National Academy of Sciences of the United States of America*, 114, 12303–12308. <https://doi.org/10.1073/pnas.1710752114>
- Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language*, 48, 103–130. [https://doi.org/10.1016/S0749-596X\(02\)00519-3](https://doi.org/10.1016/S0749-596X(02)00519-3)
- Tong, Y., Gandour, J., Talavage, T., Wong, D., Dziedzic, M., Xu, Y., ... Lowe, M. (2005). Neural circuitry underlying sentence-level linguistic prosody. *NeuroImage*, 28, 417–428. <https://doi.org/10.1016/j.neuroimage.2005.06.002>
- Tyler, L. K., & Marslen-Wilson, W. (2008). Fronto-temporal brain systems supporting spoken language comprehension. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363, 1037–1054. <https://doi.org/10.1098/rstb.2007.2158>
- van der Burght, C. L., Goucha, T., Friederici, A. D., Kreitewolf, J., & Hartwigsen, G. (2019). Intonation guides sentence processing in the left inferior frontal gyrus. *Cortex*, 117, 122–134. <https://doi.org/10.1016/j.cortex.2019.02.011>
- van Lancker, D. (1980). Cerebral lateralization of pitch cues in the linguistic signal. *Paper in Linguistics*, 13(2), 201–277. <https://doi.org/10.1080/08351818009370498>
- Vigneau, M., Beaucois, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., ... Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage*, 30, 1414–1432. <https://doi.org/10.1016/j.neuroimage.2005.11.002>
- Wagner, M., & Watson, D. G. (2010). Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes*, 25 (7–9), 905–945. <https://doi.org/10.1080/01690961003589492>
- Wildgruber, D., Hertrich, I., Riecker, A., Erb, M., Anders, S., Grodd, W., & Ackermann, H. (2004). Distinct frontal regions subserve evaluation of linguistic and emotional aspects of speech intonation. *Cerebral Cortex*, 14, 1384–1389. <https://doi.org/10.1093/cercor/bhh099>
- Xiang, H. D., Fonteijn, H. M., Norris, D. G., & Hagoort, P. (2010). Topographical functional connectivity pattern in the perisylvian language networks. *Cerebral Cortex*, 20, 549–560. <https://doi.org/10.1093/cercor/bhp119>
- Yuan, J. (2011). Perception of intonation in Mandarin Chinese. *The Journal of the Acoustical Society of America*, 130(6), 4063–4069. <https://doi.org/10.1121/1.3651818>
- Zhang, L., Xi, J., Xu, G., Shu, H., Wang, X., & Li, P. (2011). Cortical dynamics of acoustic and phonological processing in speech perception. *PLoS One*, 6(6), 1–7.

How to cite this article: Chien P-J, Friederici AD, Hartwigsen G, Sammler D. Intonation processing increases task-specific fronto-temporal connectivity in tonal language speakers. *Hum Brain Mapp*. 2021;42:161–174. <https://doi.org/10.1002/hbm.25214>