

RESEARCH ARTICLE

Neural correlates of intonation and lexical tone in tonal and non-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, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1a, 04103 Leipzig, Germany.
Email: chien@cbs.mpg.de

Funding information

Max Planck Society

Abstract

Intonation, the modulation of pitch in speech, is a crucial aspect of language that is processed in right-hemispheric regions, beyond the classical left-hemispheric language system. Whether or not this notion generalises across languages remains, however, unclear. Particularly, tonal languages are an interesting test case because of the dual linguistic function of pitch that conveys lexical meaning in form of tone, in addition to intonation. To date, only few studies have explored how intonation is processed in tonal languages, how this compares to tone and between tonal and non-tonal language speakers. The present fMRI study addressed these questions by testing Mandarin and German speakers with Mandarin material. Both groups categorised mono-syllabic Mandarin words in terms of intonation, tone, and voice gender. Systematic comparisons of brain activity of the two groups between the three tasks showed large cross-linguistic commonalities in the neural processing of intonation in left fronto-parietal, right frontal, and bilateral cingulo-opercular regions. These areas are associated with general phonological, specific prosodic, and controlled categorical decision-making processes, respectively. Tone processing overlapped with intonation processing in left fronto-parietal areas, in both groups, but evoked additional activity in bilateral temporo-parietal semantic regions and subcortical areas in Mandarin speakers only. Together, these findings confirm cross-linguistic commonalities in the neural implementation of intonation processing but dissociations for semantic processing of tone only in tonal language speakers.

KEYWORDS

phonology, pitch, prosody, semantics, voice

1 | INTRODUCTION

Speech intonation—the melodic contour of the voice—is an important linguistic feature for language comprehension and communication. Acoustically, it is defined by the pitch, that is, the fundamental

frequency (f_0) of the voice and its changes over time (Wagner & Watson, 2010). Functionally, intonation is used by speakers to mark phrase boundaries and stress relevant words in sentences (Cole, 2015; Cutler, Dahan, & van Donselaar, 1997; Dichter, Breshears, Leonard, & Chang, 2018), and also to convey important communicative messages in single words beyond their literal meanings (Ma, Ciocca, & Whitehill, 2011; Srinivasan & Massaro, 2003). For example,

Gesa Hartwigsen and Daniela Sammler are shared senior authors.

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, Inc.

a speaker may raise her vocal pitch at the end of the word 'coffee' to signal that she asks for confirmation instead of only referring to a tasty hot drink. Interestingly, regardless of their language backgrounds, listeners commonly rely on high rising pitch contours towards the end of an utterance to identify a question, even in unfamiliar languages (Gussenhoven & Chen, 2000). The present study investigated the neural bases of this cross-linguistic ability to understand intonation.

Previous neuroimaging studies demonstrated that perception of linguistic intonation involves fronto-temporal and parietal brain regions in both hemispheres. These regions have been associated with different steps in intonation processing, from basic auditory analysis of pitch via abstraction of pitch contours to subvocal rehearsal and categorical labelling of intonation, for example, as statement or question (for reviews, see Baum & Pell, 1999; Belyk & Brown, 2014; Paulmann, 2016; Wittman, van Ijzendoorn, van de Velde, van Heuven, & Schiller, 2011). One point of constant debate is the lateralisation of linguistic intonation perception. Some studies lend evidence for a right-hemispheric predominance (Meyer, Alter, Friederici, Lohmann, & von Cramon, 2002; Sammler, Grosbras, Anwender, Bestelmeyer, & Belin, 2015). These findings are in line with cue-dependent models of auditory speech perception that argue for a relative processing benefit of right auditory regions for pitch and spectral information (e.g., Zatorre, Belin, & Penhune, 2002) that unfolds over extended time scales (e.g., Poeppel, 2003). In turn, others have proposed that the lateralisation depends on the linguistic function of prosody: The stronger its linguistic function, for example, when processing syntactic structure based on intonation cues, the larger the left-lateralisation (Friederici & Alter, 2004; van der Burght, Goucha, Friederici, Kreitewolf, & Hartwigsen, 2019; van Lancker, 1980). Finally, others have argued for a general involvement of both hemispheres, with the relative contribution of either hemisphere depending on the *control* task with which intonation perception is compared (e.g., Kreitewolf, Friederici, & von Kriegstein, 2014). For instance, task-related activity during categorisation of intonation contours in single words seemed to be lateralised to the right hemisphere when compared with a linguistic task (i.e., phoneme categorisation), but to the left hemisphere when compared with a non-linguistic task (i.e., gender categorisation) in the latter study. However, most of the previous work has mainly focused on non-tonal languages (e.g., English or German). Consequently, it remains unclear how intonation guides language comprehension in tonal languages such as Mandarin Chinese (hereafter Mandarin) or Cantonese.

A better understanding of intonation perception in Mandarin is of great interest, not only because about 60–70% of the world's languages are tonal languages (Yip, 2002), but since they employ pitch to contrast semantic meanings at the syllable level by using lexical tone (hereafter tone), in addition to intonation (Chao, 1968). In Mandarin, for example, four tones are described that differ in pitch height and contour: 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) (Ladefoged & Johnson, 2011). Previous studies have suggested that lexical tonal information is functionally similar to phonemic information in language comprehension, given that tonal language speakers access tonal information early for word recognition (Malins & Joanisse, 2012) and

process tone quasi-categorically (Feng, Gan, Wang, Wong, & Chandrasekaran, 2018; Gandour & Krishnan, 2016; Peng et al., 2010; Xi, Zhang, Shu, Zhang, & Li, 2010). Neuroimaging work has shown that tone processing in native Mandarin speakers involves brain regions for the analysis and abstraction of acoustic signals, and the processing of phonological and/or semantic information. Depending on the experimental tasks, these processes are reflected by increased activity in bilateral fronto-parietal regions (Gandour, Dziedzic, et al., 2003; Gandour, Wong, et al., 2003; Gandour et al., 2004; Li, Gandour, Talavage, & Wong, 2010) or fronto-temporal areas (Kwok, Dan, Yakpo, Matthews, & Tan, 2016; for reviews, see Kwok et al., 2017; Liang & Du, 2018). Task-related activity for tone processing was found to be more left-lateralised when compared to intonation processing, emphasising the linguistic relevance of tonal pitch for semantic processing during language comprehension (Gandour, Dziedzic, et al., 2003; Gandour, Wong, et al., 2003; Gandour et al., 2004). Together, these findings illustrate the specific tuning of tonal language speakers to process tonal pitch as semantic and phonological information.

As a matter of fact, pitch has dual linguistic functions in tonal languages—intonation and tone—that are closely intertwined in their acoustic features (Ho, 1976; Ma et al., 2011) and cognitive processes (e.g., Kung, Chwilla, & Schriefers, 2014; Liu, Chen, & Schiller, 2016). In Mandarin, speakers raise pitch contour when signalling echo questions (like non-tonal language speakers do; Gussenhoven & Chen, 2000), but the realisation of the pitch rise differs depending on the concurrent tonal contour (Yuan, 2004). For example, in T2 (i.e., a high rising tone), a question intonation shows a higher rising pitch contour than in statement intonation; however, in T4 (i.e., a high falling tone), a question intonation shows a falling pitch contour—even though the pitch contour is less falling than that in a statement intonation, indicating that the falling tone underwent a pitch rise. Such overlap of pitch cues between intonation and tone influences intonation identification in Mandarin, where question identification is, for instance, more challenging in a T2 than in a T4 (Liu et al., 2016; Yuan, 2004). Moreover, task-related neural activity strongly overlaps during intonation and tone processing in fronto-parietal brain areas, albeit intonation processing induces bilateral activity and tone processing is more left-lateralised (Gandour, Dziedzic, et al., 2003; Gandour, Wong, et al., 2003; Gandour et al., 2004). The present study will further investigate how intonation processing in tonal languages overlaps with and dissociates from tone processing.

Notably, the observed bilateral fronto-parietal activity in Mandarin speakers processing intonation appears similar to activity patterns reported in non-tonal language speakers processing intonation in their native language (e.g., Kreitewolf et al., 2014). This similarity across tonal and non-tonal languages might indicate common phonological processes regardless of the language-specific realisation of intonation. However, only few studies have investigated how non-tonal language speakers process intonation in tonal languages both at the behavioural (Liang & Heuven, 2007) and neural level (Fournier, Gussenhoven, Jensen, & Hagoort, 2010; Gandour, Dziedzic, et al., 2003; Gandour, Wong, et al., 2003; Gandour et al., 2004). None of these studies included both linguistic and non-linguistic control tasks.

The present fMRI study was designed to investigate how intonation processing in a tonal language overlaps with and dissociates from

tone processing. To address cross-linguistic similarities and differences in intonation and tone processing, we compared Mandarin speakers with German speakers who had no previous exposure to Mandarin, both listening to Mandarin stimuli. We used audio-morphing, which gradually transforms one stimulus into another, to generate stimuli varying in intonation from statement to question and in tone from T2 to T4, both modulating pitch contours while keeping other acoustic features (e.g., duration and intensity) controlled. Furthermore, we included control continua that varied in voice gender between male and female (hereafter gender), that is, mainly modulating overall pitch height (Charest, Pernet, Latinus, Crabbe, & Belin, 2013). We used monosyllabic stimuli to avoid potential syntactic and compositional semantic processing that may interact with intonation at the sentence level (Sammler et al., 2018; Sammler, Kotz, Eckstein, Ott, & Friederici, 2010; van der Burgh et al., 2019). Three tasks were employed: Intonation categorisation (statement or question) and tone categorisation (T2 or T4) as experimental tasks, and gender categorisation (female or male) as a control task. The inclusion of three categorisation tasks on acoustically well-controlled stimuli enabled us to isolate task-specific top-down processing of intonation from non-linguistic pitch perception in the gender task and linguistic contour perception in the tone task (for a similar approach, see Li et al., 2003; von Kriegstein, Eger, Kleinschmidt, & Giraud, 2003). Importantly, the inclusion of the three tasks allowed us to provide a comprehensive characterisation of the hemispheric lateralisation of intonation (and tone) processing.

Our hypotheses were as follows: In both Mandarin and German speakers, we expected to observe strong overlap in intonation and tone processing in large-scale fronto-temporo-parietal networks (Gandour, Dzemidzic, et al., 2003; Gandour, Wong, et al., 2003; Gandour et al., 2004). In Mandarin speakers only, we expected to see dissociations of intonation and tone. For intonation, we expected increased activation in regions associated with prosodic contour evaluation, such as right inferior frontal gyrus (IFG) (Sammler et al., 2015). For tone, we hypothesised increased activation in semantic areas, such as left anterior IFG (Friederici, 2012; Kwok et al., 2016, 2017), angular gyrus (AG) (Hartwigsen et al., 2016), and/or the left posterior portion of superior and middle temporal gyrus (pSTG/pMTG) (Kwok et al., 2016). We also expected to see differences in hemispheric lateralisation, with a stronger contribution of the right hemisphere to intonation processing and a left-lateralised activity pattern for tone processing (Gandour, Dzemidzic, et al., 2003; Gandour, Wong, et al., 2003; Gandour et al., 2004; van Lancker, 1980). In German speakers, we did not expect any semantic processing of tone, and we expected a similar right-lateralisation of both intonation and tone.

2 | METHODS

2.1 | Participants

Twenty-four healthy native Mandarin speakers (8 males, mean age 25.4 years, age range 21–31) and 24 healthy native German speakers

(8 males, mean age 25.3 years, age range 18–32) participated in the fMRI experiment. 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. None of the participants reported a history of neurological or psychiatric disorders or contraindications against MRI. Both groups were matched with respect to musical training (Mandarin = mean 5.4 years, German = mean 5 years). None of the German speakers had prior experience with tonal languages. Before the experiment, all participants gave written informed consent in accordance with the procedures approved by the Ethics Committee of the University of Leipzig (126/18-ek).

2.2 | Tasks and stimuli

Participants listened to monosyllabic Mandarin words that were audio-morphed and varied in intonation, tone and the gender of the speakers. They were asked to categorise these stimuli in three different tasks via button press. In the intonation task, participants indicated whether the stimulus was spoken as a statement or question. Note that Mandarin interrogative intonation is typically most prominent at the utterance-final syllable (Liu et al., 2016; Yuan & Jurafsky, 2005). Hence, single words should be perceivable as question and statement. This was confirmed in a pilot test (see below). In the tone task, they had to decide whether the word was spoken with T2 or T4. In the control task (gender), participants judged whether they heard a male or female voice.

Stimuli were derived from the same speech recordings of the monosyllable 'bi' (International Phonetic Alphabet in English: [bi:]) spoken with two lexical tones (T2 'nose' and T4 'arm') and two intonation types (statement and question) from 4 native Mandarin speakers (2 males). Recordings were done in a soundproof chamber at the Max Planck Institute for Human Cognitive and Brain Sciences Leipzig using a Rode NT55 microphone with pop-filter. Speech recordings were digitised at a 44.1 kHz sampling rate in a 16-bit mono format using Audacity 2.1.1 (<http://audacityteam.org>).

To elicit naturally intoned speech, speakers were asked to picture themselves teaching imaginary students Mandarin tones, either by saying the words bi2 'nose' and bi4 'arm' while pointing to the corresponding character on a blackboard, or by asking the students whether the character they point to is bi2 'nose' or bi4 'arm'. This way, Mandarin questions were plausibly formed without using sentence-final particles (Liu & Xu, 2005). Speakers were allowed to practice production with a carrier sentence (e.g., *Is this the character for nose?*), but final recordings were done for isolated words. The resulting 16 recordings (2 words/lexical tones \times 2 intonation types \times 4 speakers) were prepared for audio-morphing by down-sampling to 16 kHz, normalisation to 85% and clipping of the release burst of the word-initial 'b' using Adobe Audition (version 3.0, Adobe Systems Inc., Mountain View, CA).

Audio-morphing was conducted with STRAIGHT (Kawahara, 2006) to create three types of continua with pitch information

gradually changing from one category to the other: Intonation (statement to question), tone (T2 to T4), and gender (female to male voice). Accordingly, 24 continua were generated, with two continua from statement to question (spoken with T2 and T4) for each speaker, two continua from T2 to T4 (spoken as statement and question) for each speaker, and two continua for gender (statement spoken with T2 and T4) for each female–male speaker pair. Temporal anchors were set at the onset and the offset of phonation. Spectro-temporal anchors were set at the onsets and offsets of the first to the fourth formants, the formant transitions and the characteristic pitch rise or fall of statements, questions, T2 and T4. Speech recordings were then re-synthesised to five-step continua with 25% steps by interpolating the anchor templates and the spectrogram (i.e., logarithmic interpolation of fundamental frequency, formant frequencies, spectro-temporal density, and aperiodicity, and linear interpolation of duration). Finally, the release burst of the 'b' was spliced back into all morphed stimuli (for details of the acoustic properties of the morphed stimuli, see

Table S1 in Supporting Information 6.1). Figure 1a–c illustrates the five-step continua used in the fMRI experiment.

Prior to the fMRI experiment, a behavioural pilot experiment was conducted with 24 Mandarin and 24 German speakers to ensure that both groups were able to perform the pitch categorisation tasks. A psychometric curve could be properly fitted to participants' responses in all three tasks, suggesting that the material is suitable for investigating the neural correlates of intonation, tone and gender perception. For details of the behavioural pilot, see Supporting Information 6.2.

2.3 | fMRI procedure

Before scanning, participants were first provided with a scenario instructing them to imagine that four instructors are teaching the Mandarin words bi2 'nose' and bi4 'arm' to students by saying the word while pointing to the corresponding character on a blackboard, or by asking the students whether the character they point to is bi2

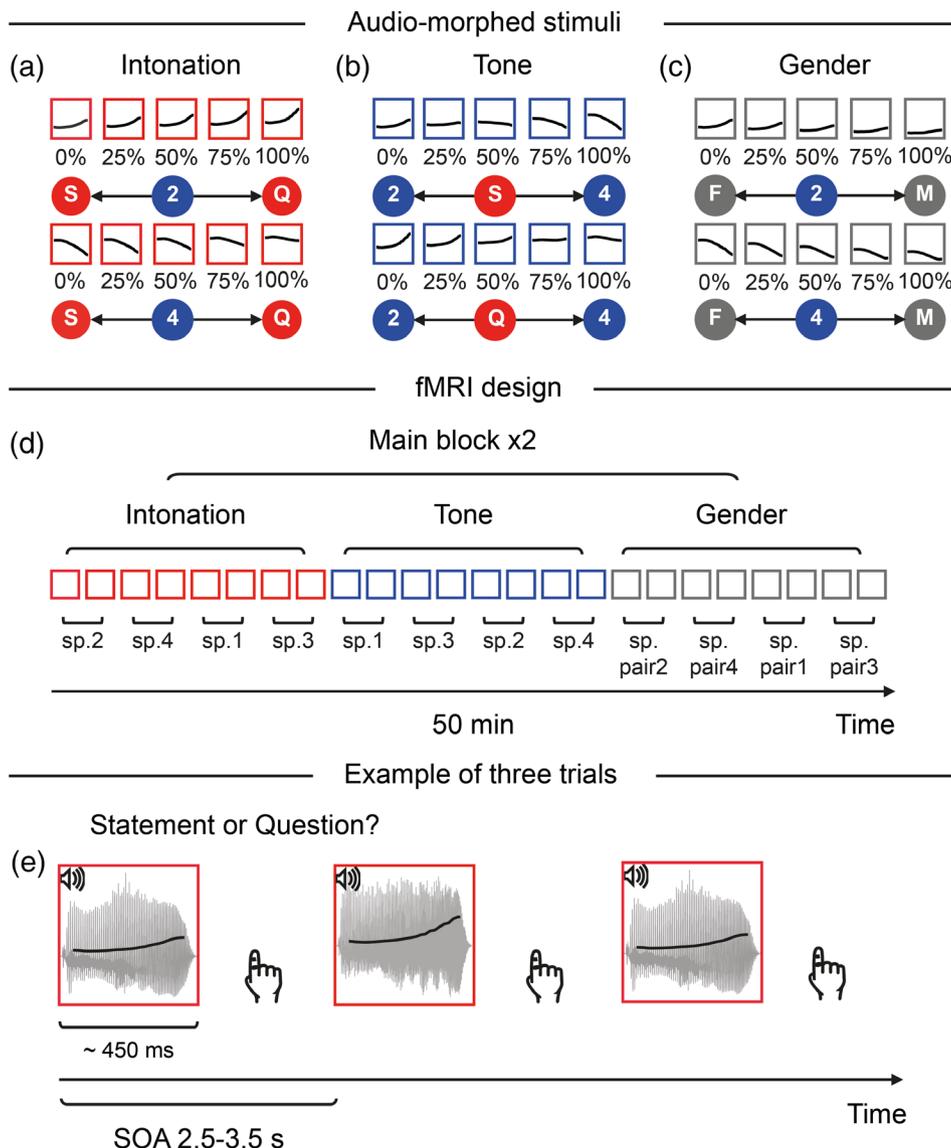


FIGURE 1 Illustration of the five-step continua and experimental design. Mandarin words bi2 'nose' and bi4 'arm' uttered as statement or question by male and female Mandarin speakers were used for audio-morphing to create (a) intonation, (b) tone, and (c) gender continua. Each box in the continuum illustrates one morph step. Participants were asked to categorise intonation (statement or question), tone (T2 or T4) and gender (female or male voice) of the respective continua. (d) Timeline of the fMRI experiment. Intonation, tone and gender categorisation tasks were carried out in separate task blocks, consisting of eight mini-blocks each (i.e., two mini-blocks per speaker; each box corresponds to one mini-block). Each mini-block consisted of 15 trials. The scanning session lasted about 50 min. sp., speaker. (e) Example of three trials in the intonation task. Stimuli were presented with a mean jittered SOA of 3 ± 0.5 s. Participants categorised the intonation of the stimulus via button press. This trial structure was analogous in the tone and gender task

'nose' or bi4 'arm' (see Supporting Information 6.3 for details). Thereafter, they practiced each task for 10 min.

During fMRI, morphed stimuli from the five-step continua were presented at comfortable volume via MR-compatible headphones (MR confon GmbH, Magdeburg, Germany). Participants judged intonation (statement or question), tone (T2 or T4), or gender (male or female) in six separate blocks (two per task) via button press with their right index and middle finger. In each block, we presented stimuli only from one of the two continua in the corresponding task (see Figure 1a–c). During each block, participants were visually presented with the two relevant categories on the screen (e.g., statement, question), separated by a central fixation point, with its left–right alignment corresponding to the button assignment. Each block had 360 trials presented in 8 mini-blocks with 15 trials each. In each task, the first mini-block began with a 5 s visual task instruction. Stimuli were presented with a jittered SOA ranging from 2.5 to 3.5 s (mean SOA = 3 s) in pseudo-random order such that each morph step followed each morph step with similar probability. Mini-blocks were separated by 15 s rest breaks with a central fixation point on the screen (Figure 1d–e). Stimulus presentation and response registration was controlled with Presentation software (Version 19.0, Neurobehavioural Systems, Inc., Berkeley, CA). Button assignment and task order were counterbalanced across participants. The fMRI session lasted ~50 min.

2.4 | Data acquisition

Functional images were acquired with a 3T Siemens Magnetom Prisma scanner (Siemens AG, Erlangen, Germany) using a multi-band echo-planar imaging sequence (EPI; 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 × 2 × 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 either taken from the brain database of the Max Planck Institute for Human Cognitive and Brain Sciences Leipzig or 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 = 2300 ms, TE = 2.98 ms, flip angle = 9°).

2.5 | Behavioural data analysis

Customised scripts in MatlabR2019 (The MathWorks, Inc., Natick, MA) and SPSS (PASW) Statistics 21.0 (SPSS Inc., Chicago, IL) were used for analysing the behavioural data. We calculated the proportion of participants' 'question', 'T4', and 'male voice' responses for each of the five morph-steps in the intonation, tone and gender task. A psychometric curve was fitted to participants' categorisation data for each task using a cumulative logistic function with the following formula:

$$y = a + \frac{b}{1 + \exp\left(\frac{c-x}{s}\right)}$$

Note that a and b correspond to the values of the left and right asymptotes, c to the centre of symmetry of the fitting curve, and s to the slope of the fitting curve at c . The slope of the psychometric curve was analysed as performance measure for intonation, tone and gender categorisation. Better performance is reflected in a steeper slope (e.g., Hallé, Chang, & Best, 2004). In addition, we analysed participants' response times (RTs) in the three categorisation tasks. Better performance is reflected in shorter RTs for clear stimuli (i.e., morph steps 1 and 5), and longer RTs for ambiguous stimuli (i.e., morph steps 2, 3, and 4), resulting in larger RT differences between clear and ambiguous stimuli (e.g., Hallé et al., 2004). Two-way analyses of variance with factors Group (Mandarin speakers/German speakers) and Task (intonation/tone/gender) were calculated for each of these measures. Post-hoc pair-wise comparisons were conducted and Bonferroni corrected on significant F -tests. p -values were adjusted with Greenhouse–Geisser correction (Greenhouse & Geisser, 1959), if necessary.

2.6 | fMRI data analysis

fMRI data were analysed with SPM 12 (Wellcome Trust Centre for Neuroimaging, London, United Kingdom). Data preprocessing 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.

For statistical analyses, we estimated a general linear model for each participant as implemented in SPM 12, including one regressor for each task (i.e., intonation, tone, and gender) and convolving the onset and duration (set as 0) of stimulus presentation with a canonical hemodynamic response function. RTs were included as parametric modulators to account for differences in task difficulty. Onsets and durations of visual task instructions and six motion parameters were modelled as nuisance regressors. T-contrasts for comparisons of interest (i.e., intonation > tone, tone > intonation, intonation > gender, and tone > gender) were calculated for each participant and subjected to one-sample t -tests at the second level for each of the two groups. To test for overlap between intonation and tone processing in each group, conjunction analyses (using the conjunction null conjunction, Nichols, Brett, Andersson, Wager, & Poline, 2005) of intonation > gender and tone > gender were performed separately for Mandarin and German speakers.

Between-group comparisons were performed by including the first-level contrasts of intonation or tone against the implicit baseline from each group into two-sample t -tests at the second level. Finally, overlap of intonation processing (relative to gender) or tone processing (relative to gender), respectively, between Mandarin and German speakers were tested with conjunction analyses.

In addition, we performed a lateralisation analysis for intonation processing. Participants' raw fMRI data were preprocessed with the same preprocessing pipeline but segmented and normalised using a symmetric MNI template. For statistical analyses, we estimated a general linear model (cf. whole-brain analysis above), and the resulting contrast image of interest (i.e., intonation > tone) was then left-right flipped. The contrast image of interest and its flipped equivalent was compared with paired *t*-tests at the second level (Bozic, Tyler, Ives, Randall, & Marslen-Wilson, 2010; Liégeois et al., 2002; van der Burght et al., 2019). All comparisons were thresholded using a cluster-forming threshold of $p < .001$ at voxel level (uncorrected) and a family-wise error (FWE) correction of $p < .05$ at the cluster level. Anatomical locations were identified using the SPM Anatomy Toolbox 2.2b (Eickhoff et al., 2005) and Jülich probabilistic cytoarchitectonic maps. fMRI results were visualised by 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).

3 | RESULTS

3.1 | Behavioural data

Figure 2 illustrates group-averaged psychometric curves and RTs across the morph steps in all tasks for both groups. Overall, both groups were able to perform the three pitch categorisation tasks accurately. Statistical analyses of the mean slope (indicated in degrees) showed a significant interaction between Group and Task ($F[2,92] = 20.486, p < .001$), and main effects of Group ($F[1,46] = 13.449, p = .001$) and Task ($F[2,92] = 50.985, p < .001$). Further pair-wise comparisons showed that Mandarin speakers were less successful in categorising intonation than tone and gender, but did not reveal significant differences between tone and gender categorisation. In contrast, German speakers categorised intonation and tone equally well, both of which were poorer than gender. Between-group comparisons revealed similar performance of

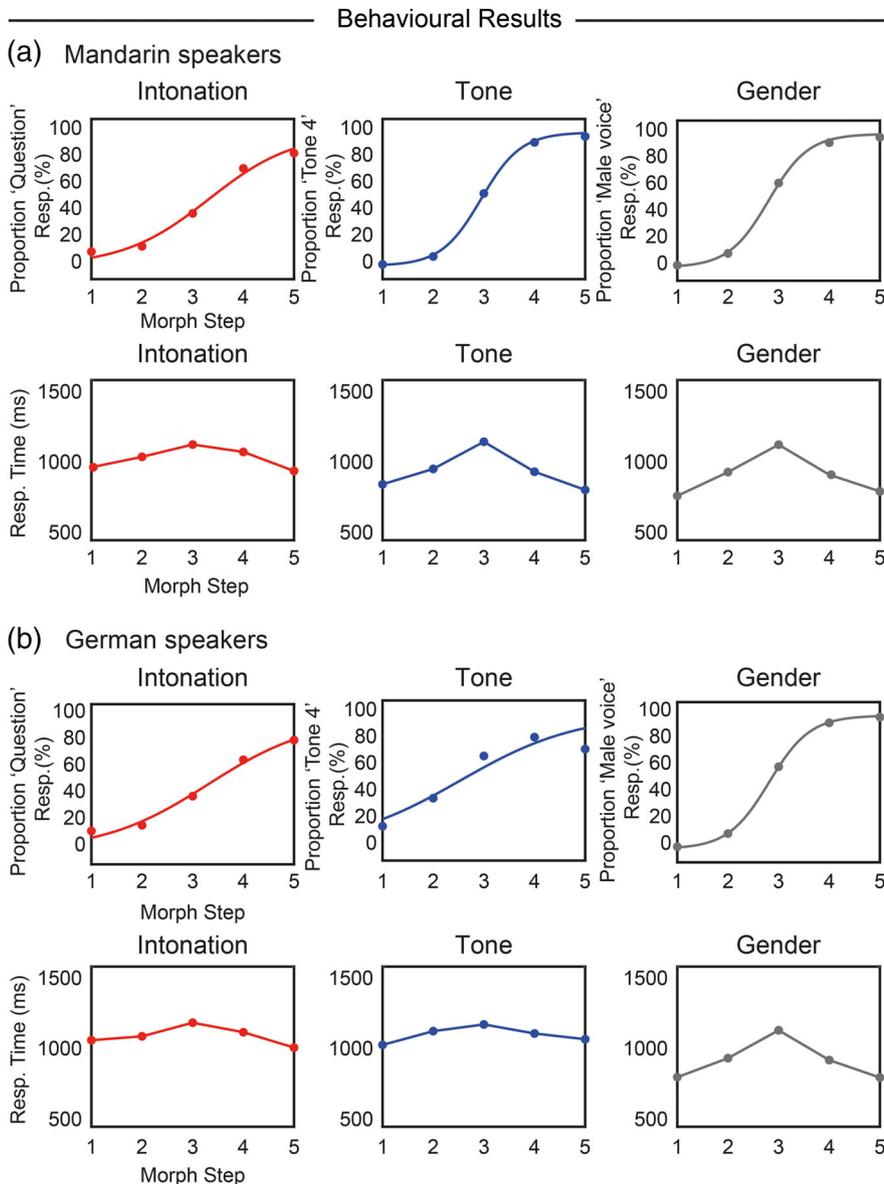


FIGURE 2 Group-averaged psychometric curves and RTs across the morph steps in intonation, tone, and gender task in the fMRI experiment. (a) Mandarin speakers. (b) German speakers. Resp., Response. Steeper slopes, faster responses in clear stimuli (morph steps 1 and 5), slower responses in ambiguous stimuli (steps 2–4), and larger RT differences between ambiguous and clear stimuli indicate better performance. RT, response time

TABLE 1 Post-hoc pair-wise comparisons of the behavioural results in the fMRI experiment

Contrast	Mean slope (in degrees)		Mean RTs of clear stimuli		Mean RT difference between clear and ambiguous stimuli	
	t-value ^a	p-value	t-value ^a	p-value	t-value ^a	p-value
Mandarin speakers						
Intonation vs. Tone	10.405	<.001	-8.190	<.001	4.583	<.001
Intonation vs. Gender	-6.230	<.001	6.491	<.001	-3.728	.001
Tone vs. Gender	1.549	.135	2.206	.038	0.823	.419
German speakers						
Intonation vs. Tone	1.150	.262	0.576	.570	0.412	.684
Intonation vs. Gender	-9.771	<.001	6.838	<.001	-4.697	<.001
Tone vs. Gender	-5.975	<.001	8.595	<.001	-3.855	.001
Mandarin speakers vs. German speakers						
Intonation	1.697	.096	-1.822	.075	1.468	.149
Tone	5.431	<.001	-4.778	<.001	3.898	<.001
Gender	-0.200	.842	-0.553	.583	0.425	.673

^at(23) for paired-samples t-tests in Mandarin and German speakers, t(46) for between-group comparison of Mandarin versus German speakers. Bold font indicates significant results after Bonferroni correction.

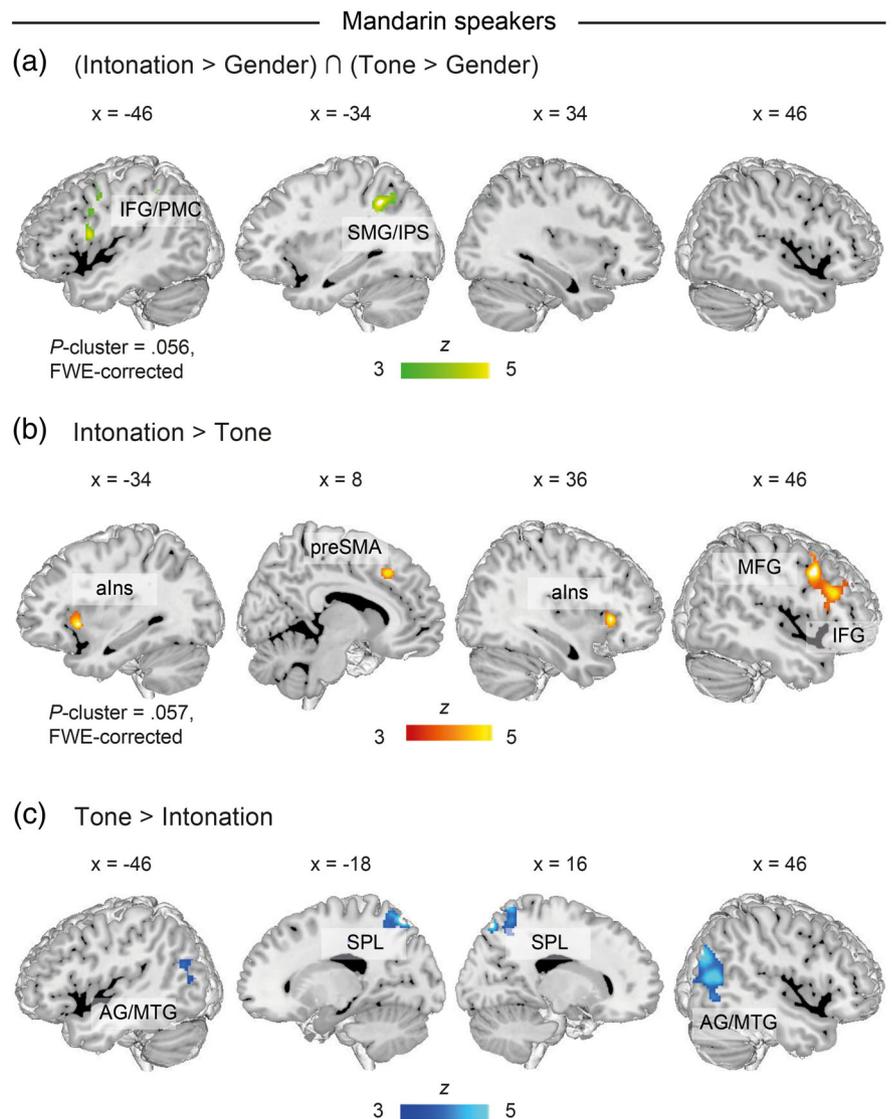


FIGURE 3 Comparisons of tasks in Mandarin speakers. (a) Conjunction of Intonation > Gender and Tone > Gender, (b) Intonation > Tone, and (c) Tone > Intonation. All clusters are thresholded at p-cluster < .05 (FWE-corrected) unless otherwise indicated. FWE, family-wise error

TABLE 2 Comparisons of tasks in Mandarin speakers

Region	BA	k	z-value	p-value	MNI coordinates		
					x	y	z
(Intonation > Gender) \cap (Tone > Gender)							
L SMG	40	540	4.60	.009	-34	-48	40
	40		4.00		-54	-34	44
L IFG (p. op.)	44	322	4.36	.056	-46	8	14
L PMC	6		3.80		-42	-2	42
	6		3.59		-40	2	32
Intonation > Tone							
R IFG (p. tri.)	45	1,120	4.88	.000	54	24	28
R IFG (p. op.)	44		4.50		48	12	38
R alns	13		4.17		36	24	2
R MFG	6		3.68		46	12	52
L preSMA	32	418	4.23	.025	-4	18	46
R preSMA	8		4.02		8	26	42
L alns	13	324	4.17	.057	-34	18	-2
Tone > Intonation							
R SPL	7	2,197	4.45	.000	16	-74	58
	7		4.23		22	-58	70
L PCC	31		4.35		-8	-34	36
R PCC	31		3.70		8	-20	40
L SPL	7	694	4.44	.003	-18	-68	60
	7		4.20		-30	-56	58
R AG	39	1,169	4.29	.000	54	-64	14
	39		4.16		46	-74	32
R MTG	19		4.03		44	-68	12
L MTG	37	454	3.90	.019	-58	-66	6
	39		3.40		-46	-64	22
L AG	19		3.57		-36	-78	36

Notes: Peak voxels in clusters are in bold. *p*-values refer to *p*-cluster < .05, FWE-corrected.

Abbreviations: BA, Brodmann area; L, left hemisphere; R, right hemisphere; k, cluster size (number of voxels); p. op., pars opercularis; p. tri., pars triangularis.

Mandarin and German speakers both in intonation and gender, while Mandarin speakers categorised tone better than German speakers. Statistical results of the post-hoc pair-wise comparisons (Bonferroni corrected) are provided in Table 1. Similar patterns were found for mean RTs for clear stimuli and RT differences between clear and ambiguous stimuli (see also Table S3 in Supporting Information 6.4).

3.2 | fMRI data

3.2.1 | Mandarin speakers

To identify brain regions in which intonation processing overlaps with and dissociates from tone processing, we first calculated a conjunction analysis of the contrasts intonation > gender and tone > gender,

followed by direct contrasts between intonation and tone. The conjunction analysis revealed strong overlap in the left intraparietal sulcus (IPS)/supramarginal gyrus (SMG). With a more liberal threshold (*p*-cluster = .056, FWE-corrected), left IFG and premotor cortex (PMC) were also identified in the overlap (Figure 3a and Table 2; see Supporting Information 6.5 for details of the intonation > gender and tone > gender contrasts).

Direct comparisons between intonation and tone showed their dissociation in frontal and temporo-parietal regions. The contrast of intonation > tone revealed a stronger involvement of right frontal regions, including right IFG, right middle frontal gyrus (MFG), right anterior insula (alns), and pre-supplementary motor area (preSMA). Left alns was also identified with a more liberal threshold (*p*-cluster = .057, FWE-corrected) (Figure 3b and Table 2). The lateralisation analysis (see Section 2.6) showed that activity in

FIGURE 4 Comparisons of tasks in German speakers (p -cluster < .05, FWE-corrected). FWE, family-wise error

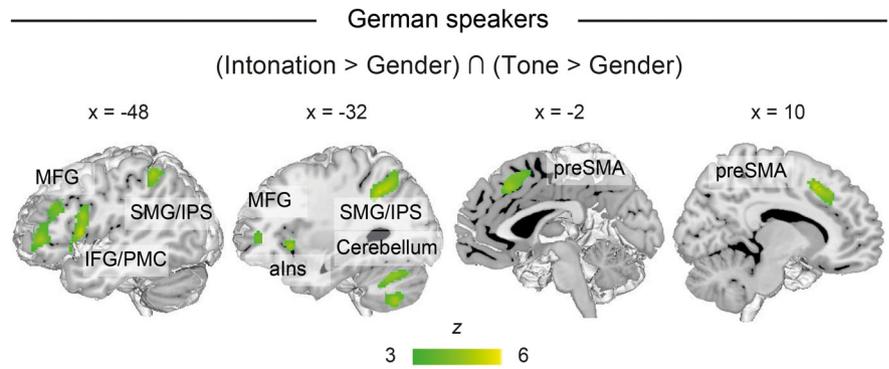


TABLE 3 Comparisons of tasks in German speakers

Region	BA	k	z-value	p-value	MNI coordinates		
					x	y	z
(Intonation > Gender) ∩ (Tone > Gender)							
L MFG	45	1,182	5.14	.000	-42	34	28
L orbito-frontal cortex	46		4.71		-44	44	-2
L IPS	40	986	5.04	.000	-36	-50	42
	40		4.48		-50	-42	52
L IFG (p. op.)	44	603	4.83	.002	-48	12	6
	44		4.46		-50	10	18
L IFG (p. orb.)	13		3.87		-42	20	-4
L alns	13		4.33		-32	22	-4
L PMC	6		3.33		-40	2	36
R preSMA	6	950	4.62	.000	10	18	48
	32		4.32		8	28	42
L preSMA	8		4.33		-2	26	44
L cerebellum (VIIb)	-	834	4.48	.000	-32	-62	-50
L cerebellum (VI)	-		4.40		-28	-60	-28
L cerebellum (VIIa)	-		4.06		-8	-78	-28
R cerebellum (VIIa)	-	271	4.42	.063	32	-68	-48

Notes: Peak voxels in clusters are in bold. p -values refer to p -cluster < .05, FWE-corrected. Abbreviations: BA, Brodmann area; L, left hemisphere; R, right hemisphere; k, cluster size (number of voxels); p. op., pars opercularis; p. orb., pars orbitalis.

FIGURE 5 Comparison of intonation processing between Mandarin and German speakers. p -cluster < .05, FWE-corrected, if not otherwise indicated. FWE, family-wise error

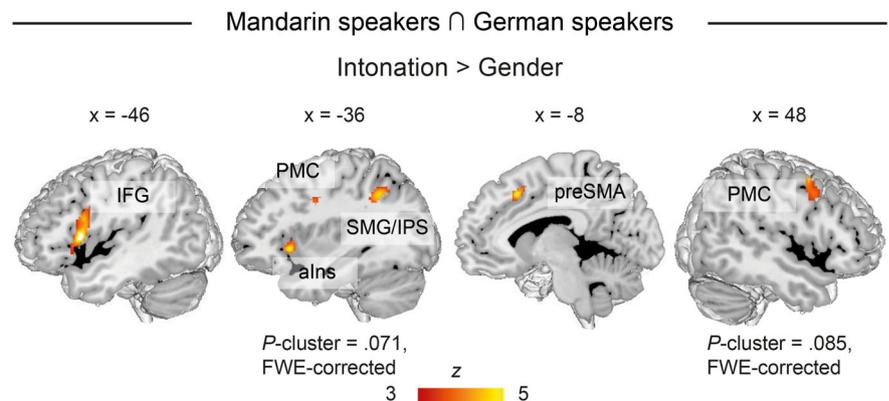


TABLE 4 Comparison of intonation processing between Mandarin and German speakers

Region	BA	k	z-value	p-value	MNI coordinates		
					x	y	z
Mandarin speakers \cap German speakers: Intonation > Gender							
L IFG (p. op.)	44	651	4.80	.004	-46	12	8
	9		4.20		-48	10	20
L alns	13		4.38		-32	20	-4
L PMC	6		3.56		-36	0	36
L preSMA	32	456	4.44	.018	-8	20	46
	8		3.98		8	26	42
L IPS	40	298	4.21	.071	-36	-50	40
L SMG	40		3.18		-54	-36	44
R PMC	44	279	4.19	.085	48	6	52

Notes: Peak voxels in clusters are in bold. *p*-values refer to *p*-cluster < .05, FWE-corrected.

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

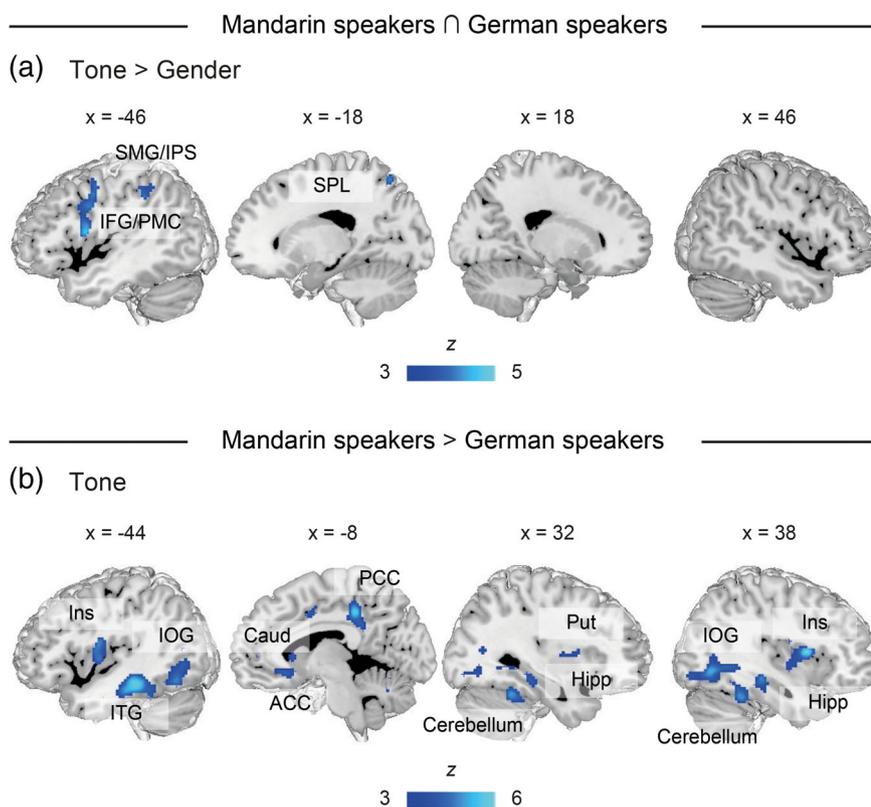


FIGURE 6 Comparisons of tone processing between Mandarin and German speakers. (a) Conjunction of Tone > Gender in Mandarin and German speakers, (b) direct comparison of Mandarin speakers > German speakers in the Tone task. All clusters are thresholded at *p*-cluster < .05, FWE-corrected. FWE, family-wise error

IFG/MFG ($x, y, z = 50, 24, 24$; $z = 3.87$) was significantly lateralised to the right hemisphere. The comparison of tone > intonation showed stronger activity in bilateral AG/MTG and superior parietal lobule (SPL) as well as posterior cingulate cortex (PCC) (Figure 3c and Table 2).

3.2.2 | German speakers

The conjunction analysis of intonation > gender and tone > gender showed strong overlap in left fronto-parietal cortex and the

cerebellum. Cortical areas included the left MFG/orbito-frontal cortex, IFG/PMC, alns, preSMA, and IPS/SMG (Figure 4 and Table 3; see Supporting Information 6.6 for details of the simple contrasts of intonation > gender and tone > gender). The direct comparisons between intonation and tone did not reveal any significant clusters.

3.2.3 | Mandarin versus German speakers

The conjunction analysis of intonation > gender in Mandarin and German speakers revealed shared activity in the left IFG, left alns and

TABLE 5 Comparisons of tone processing between Mandarin and German speakers

Region	BA	k	z-value	p-value	MNI coordinates		
					x	y	z
Mandarin speakers \cap German speakers: Tone > Gender							
L IPS	40	1,122	4.92	.000	-34	-46	42
	7		4.11		-32	-60	54
L SMG	40		4.32		-52	-34	44
L SPL	7		3.77		-18	-66	58
L PMC	6	638	4.26	.001	-36	0	28
	6		3.83		-50	0	50
L IFG (p. op.)	44		4.11		-46	6	12
Mandarin speakers > German speakers: Tone							
R PCC	31	5,333	5.35	.000	6	-28	42
	31		5.32		10	-34	38
L PCC	31		4.94		-8	-36	42
R IOG	19		4.52		38	-68	-6
R hippocampus	-		4.41		36	-28	-10
R cerebellum	-		4.41		32	-44	-24
R insula	13	522	5.19	.004	38	8	10
R putamen	-		3.37		24	-6	8
	-		3.49		32	-4	8
L ITG	20	890	5.02	.000	-44	-32	-16
	36		4.86		-42	-24	-20
L parahippocampal gyrus	-		4.29		-30	-26	-18
L hippocampus	-		3.58		-30	-36	-8
L MTG	21		3.37		-54	-22	-18
L Rolandic operculum	13	724	4.28	.001	-44	-2	16
L Heschl's gyrus	13		4.28		-52	-12	10
L insula	13		4.14		-42	0	10
L IOG	37	557	4.21	.003	-44	-62	-12
L MOG	37		3.94		-40	-68	0
R cerebellum (V)	-	682	4.21	.001	8	-58	-20
R cerebellum (VI)	-		4.06		18	-58	-14
L cerebellum (VI)	-		3.64		-16	-54	-22
R ACC	24	507	3.93	.004	6	40	-6
L caudate nucleus	-		3.77		-10	20	4

Notes: Peak voxels in clusters are in bold. *p*-values refer to *p*-cluster < .05, FWE-corrected.

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

preSMA. With a more liberal threshold, the left IPS/SMG (*p*-cluster = .071, FWE-corrected) and the right PMC (*p*-cluster = .085, FWE-corrected) were also identified (Figure 5 and Table 4). Direct comparisons of intonation (relative to the implicit baseline) between groups did not show any significant differences.

The conjunction analysis of tone > gender in Mandarin and German speakers revealed overlap in left fronto-parietal regions, specifically in the left IFG/PMC, IPS/SMG, and SPL (Figure 6a and Table 5). In the direct comparison of tone processing (tone relative to implicit baseline)

between groups, Mandarin speakers showed significantly stronger activity than German speakers in fronto-temporo-occipital and subcortical regions. This activation pattern included anterior and posterior cingulate cortex (ACC, PCC), bilateral insula, left Rolandic operculum/Heschl's gyrus, left inferior temporal gyrus (ITG), left inferior occipital gyrus (IOG)/middle occipital gyrus (MOG), bilateral cerebellum, and basal ganglia (right putamen, left caudate). Left parahippocampal gyrus was also active (Figure 6b and Table 5). The opposite contrast of tone processing in German > Mandarin speakers did not show significant effects.

4 | DISCUSSION

The goal of the present fMRI study was to elucidate cross-linguistic commonalities and differences of intonation and tone processing in tonal and non-tonal language speakers. To this end, we used three different pitch categorisation tasks on a range of intonational and tonal pitch contours. As a first main finding, we observed a strong overlap of intonation and tone processing in left fronto-parietal regions in both Mandarin and German speakers. Additionally, intonation (but not tone) showed further overlap between groups in right frontal regions. These combined bilateral IFG/PMC and left SMG/IPS activity patterns argue for a general bilateral network for the processing of intonation in both tonal and non-tonal language speakers. However, while left-hemispheric regions are involved in phonological processing and evaluation of linguistic pitch contours irrespective of whether they express intonation or tone, right frontal areas seem to be specific for the categorical evaluation of intonation. This functional specificity was further confirmed by the direct comparison of intonation relative to tone processing in Mandarin speakers that revealed a significantly right-lateralised contribution of IFG/MFG to intonation evaluation. As a second main finding, Mandarin speakers showed stronger involvement of bilateral temporo-parietal and subcortical regions when processing tonal pitch, both compared to intonation and compared to German speakers, likely reflecting semantic processing. Together, our results demonstrate cross-linguistic commonalities in the neural processing of intonation that overlaps with the phonological (but not semantic) processing of tone across Mandarin and German speakers. In contrast, semantic processing of tone was only observed in Mandarin speakers.

4.1 | Left fronto-parietal regions contribute to linguistic pitch contours in Mandarin and German speakers

As the first key finding of our study, intonation processing overlapped with tone processing in left posterior IFG (i.e., pars opercularis) and adjacent PMC as well as in left SMG/IPS, both in Mandarin (Figure 3a) and German speakers (Figure 4) and in the between-group conjunctions (Figures 5 and 6a), suggesting a general and cross-linguistic role of these regions in the categorisation of linguistic pitch contours. These areas have been previously found in Mandarin speakers for tone or intonation processing (Gandour, Dziedzic, et al., 2003; Gandour, Wong, et al., 2003; Gandour et al., 2004), as well as in non-tonal language speakers processing intonation in their native language (Kreitewolf et al., 2014; Merrill et al., 2012; for a review, see Belyk & Brown, 2014). The present study is the first to reveal these areas in a direct comparison between tonal and non-tonal language speakers, both processing Mandarin material. Notably, all these regions were observed compared to the non-linguistic control task, that is, gender categorisation. This strengthens our assumption that these fronto-parietal activity patterns reflect linguistic processes, more specifically, the phonological processing of pitch contours.

The functional relevance of the observed fronto-parietal regions for phonological processing is well established, both as part of the articulatory language network for audio-motor mapping (e.g., Hickok & Poeppel, 2004, 2007) and the phonological working memory loop (Baddeley, 1992, 2003b). Accordingly, perturbation of left posterior IFG or SMG with neurostimulation has been shown to disrupt phonological judgments and short-term retention of words (Hartwigsen et al., 2010, 2016; Romero, Walsh, & Papagno, 2006). More specifically, it has been proposed that left IFG/PMC is associated with articulatory-based representations (for reviews, see Hickok & Poeppel, 2004, 2007) and covert rehearsal (Baddeley, 2003a, 2003b), while left SMG is relevant for storing and manipulating auditory phonological information (Baddeley, 2003a).

The present data extend the notion of fronto-parietal phonological processing to linguistic pitch contours in intonation and tone (e.g., Gandour, Dziedzic, et al., 2003; Gandour, Wong, et al., 2003). Accordingly, the observed left IFG/PMC activity may indicate that participants subvocally rehearsed the pitch contours to better recognise the pitch categories in both the intonation (Hickok & Poeppel, 2004; Kreitewolf et al., 2014) and tone task (Liang & Du, 2018). Left SMG activity may indicate that participants temporarily stored and compared pitch contours of consecutive stimuli to ease categorisation. The SMG cluster further extended into the left IPS, in line with previous reports on intonation (Gandour, Dziedzic, et al., 2003; Gandour, Wong, et al., 2003; Gandour et al., 2004; Kristensen, Wang, Petersson, & Hagoort, 2013; Sammler et al., 2015) and tone (Gandour, Dziedzic, et al., 2003; Gandour, Wong, et al., 2003; Gandour et al., 2004; Tong et al., 2005). IPS has been associated with numerous functions (Grefkes & Fink, 2005), including sensorimotor (Hamilton & Grafton, 2007) or attentional and executive control (Corbetta & Shulman, 2002; Duncan, 2010) that may both more generally support the categorisation of pitch contours and speech sounds (Husain et al., 2006).

Together, the present results allow a generalisation of previous findings by demonstrating that left fronto-parietal areas contribute to both intonation and tone processing, independent of speakers' language background. Both tonal and non-tonal language speakers appear to commonly rely on phonological processes such as storage and covert rehearsal of phonological pitch information and executive control when categorising linguistic pitch contours. The shared activity patterns across intonation and tone may highlight genuine similarities in the categorisation of sublexical, but linguistically meaningful pitch contours in the left hemisphere.

4.2 | Right frontal regions are recruited for intonation processing in Mandarin and German speakers

Compared to the general involvement of fronto-parietal regions in the left hemisphere in both intonation and tone, IFG and PMC in the right hemisphere showed a preference for intonation over tone. Specifically in Mandarin speakers, the direct comparison between intonation and

tone revealed a significantly right-lateralised contribution of IFG/MFG to intonation processing (Figure 3b). Furthermore, the conjunction of intonation relative to gender (but not tone compared to gender) across Mandarin and German speakers revealed the shared contribution of right PMC, at the level of the larynx representation (Brown et al., 2009; Brown, Ngan, & Liotti, 2008; Dichter et al., 2018). Increased activity in right IFG has been previously associated with the explicit evaluation of prosodic categories (Belyk & Brown, 2014; Liang & Du, 2018; Witteman et al., 2011), including question and statement intonation (Kreitewolf et al., 2014; Sammler et al., 2015). Likewise, right PMC activity has been found during the perception of intonation contours (Dichter et al., 2018), and its causal role in the categorisation of intonation, presumably via subvocal simulation of laryngeal gestures, has been shown in a recent neurostimulation study (Sammler et al., 2015).

Together with the observed left fronto-parietal activity during intonation processing (see above), our findings are in agreement with studies that argue for a bilateral processing of intonation (Kreitewolf et al., 2014). At the same time, our data also support the notion that right IFG/PMC is specific to intonation (not tone), across languages. More generally, our data resonate with the idea that, at least at the higher cognitive level, the *function* of pitch information rather than pitch as stimulus *feature* determines the lateralisation of intonation and tone perception (Gandour, Dziedzic, et al., 2003; Gandour, Wong, et al., 2003; Gandour et al., 2004; van der Burght et al., 2019; van Lancker, 1980), irrespective of speakers' language background. While the processing of pitch as (lexical) tone is located in the left-hemispheric phonological system, pitch perceived as intonation involves bilateral regions that include right-hemispheric prosodic areas together with left-hemispheric phonological areas (Friederici, 2011; Friederici & Alter, 2004; Kreitewolf et al., 2014). Future studies should clarify how these results generalise beyond the word level to sentence-level intonation.

More generally, our findings overlap with the dorsal stream in the dual-stream model of intonation and tone processing proposed by a recent meta-analysis (Liang & Du, 2018). In both groups, we identified the left IFG/PMC shared by intonation and tone processing and the right IFG/PMC additionally recruited by intonation processing (which is also in line with the dorsal stream for prosody by Sammler et al., 2015). However, we did not observe significant temporal activity, that is, in the ventral stream, for intonation and tone processing. The most likely explanation is that the acoustic properties of our stimuli were well-matched such that the analysis and abstraction of incoming auditory signals was subtracted out in the direct contrasts.

4.3 | Contribution of cingulo-opercular regions to categorical decisions in Mandarin and German speakers

Apart from left fronto-parietal and right frontal activity, our data also showed activity in bilateral anterior insula and preSMA in intonation processing, compared with tone in Mandarin speakers (Figure 3b), and

also compared with the gender task in Mandarin and German speakers (Figures 4 and 5). Similar activity patterns have been previously reported in studies that required prosodic decisions (for review, see Belyk & Brown, 2014), particularly when these decisions were difficult, for example, when pitch contours were ambiguous (Hellbernd & Sammler, 2018; Sammler et al., 2015). More generally, the observed bilateral cingulo-opercular activity most likely reflects explicit categorical decision making under high cognitive control. Indeed, increased activity in these regions has been previously linked to decision-making processes under challenging conditions (Camilleri et al., 2018; Duncan, 2010; Geranmayeh, Brownsett, & Wise, 2014), and categorical decision making can be accompanied by increased cognitive control (Duncan, 2010). This fits with our observation that the intonation task was the behaviourally most challenging condition compared to the gender task in both groups and compared to tone in Mandarin speakers. Together, it is likely that the observed increase in task-related activity reflects both categorical decision making and increased task demands that are closely intertwined (Duncan, 2010). The comparable recruitment of these regions in both groups underlines that categorising intonation in a tonal language is challenging not only for non-tonal language speakers but also for tonal language speakers, even if the latter have long-term experience with tonal intonation patterns from everyday conversation.

4.4 | Tone processing engages semantic areas in Mandarin speakers

Finally, Mandarin (but not German) speakers showed bilateral temporo-parietal and subcortical activation that was specific to tone processing. Tone compared to intonation processing involved bilateral AG/pMTG, SPL and PCC (Figure 3c), likely reflecting semantic processing. AG and pMTG have been consistently identified when participants are engaged in task-related semantic decisions and lexical semantic processing (Binder, Desai, Graves, & Conant, 2009; Hartwigsen et al., 2017; Noonan, Jefferies, Visser, & Lambon Ralph, 2013). The specific involvement of AG/pMTG for tone processing fits well with this account since tonal pitch in Mandarin is employed to contrast lexical semantic meanings (Chao, 1968). Although no explicit semantic decisions were required in our tone categorisation task (i.e., participants were asked to identify T2 or T4, not to indicate whether the speakers said 'nose' or 'arm'), automatic semantic processes may have been triggered in our Mandarin speakers (Copland et al., 2003).

Notably, there is an on-going debate whether AG is engaged in semantic processing per se or is part of the default mode network (DMN) (Hartwigsen, 2018; Humphreys, Hoffman, Visser, Binney, & Lambon Ralph, 2015; Lambon Ralph, Jefferies, Patterson, & Rogers, 2016; Seghier, 2013; Seghier, Fagan, & Price, 2010). Indeed, the present effects in Mandarin speakers reflect less *deactivation* during tone compared to intonation processing, in line with previously observed DMN activity patterns (Binder et al., 2009; Binder & Desai, 2011). It may also be argued that the lower difficulty of tone than intonation

categorisation in our Mandarin participants (see Figure 2) supports a DMN interpretation, even though we controlled for task difficulty in our statistical design (see above). On a more abstract level, current views have argued for a close relationship between the DMN and the semantic system (Binder et al., 2009), for example, by showing that mind wandering at rest is closely linked to semantic memory processes (Binder & Desai, 2011). Accordingly, we may propose that the observed AG/pMTG effects are associated with the intrinsically easier categorisation of tone than intonation in addition to automatic semantic processing.

Similar conclusions can be drawn from the direct comparison of tone processing between Mandarin and German speakers. Here, Mandarin speakers also showed stronger activity in a set of regions that have been frequently found during semantic decisions, semantic memory retrieval and articulatory planning in language comprehension, including the bilateral insula, left ITG, bilateral IOG, PCC, ACC, cerebellum, hippocampus, and the basal ganglia (Binder et al., 2009; Brown et al., 2009; Levy, Bayley, & Squire, 2004; Price, 2010). It should be noted that stimuli in all tasks contained semantic information, even if semantic meaning was not task-relevant. Therefore, it is likely that activity in other semantic regions (e.g., left anterior IFG) was contrasted out in the direct between-task/-group comparisons.

Together, the present results may demonstrate the inevitable semantic processing of tonal pitch in Mandarin (but not German) speakers, corroborating the notion that the processing of lexical tone is tuned differently depending on speakers' language background (e.g., Gandour, Dziedzic, et al., 2003; Gandour, Wong, et al., 2003; Gandour et al., 2004; Hallé et al., 2004).

5 | CONCLUSION

The present study identified cross-linguistic commonalities and dissociations between intonation and tone processing in speakers of a tonal or a non-tonal language. Our results specify three core neural systems for the processing of intonation that are shared by Mandarin and German speakers: (a) left fronto-parietal regions for general phonological processing of linguistic pitch contours, (b) right frontal regions for prosodic category evaluation specific to intonation, and (c) bilateral cingulo-opercular areas for controlled categorical decision making. Although tone processing overlapped with intonation in left fronto-parietal regions in both groups, it showed an additional contribution of bilateral temporo-parietal semantic areas and subcortical regions in Mandarin speakers only. Our findings demonstrate that the bilateral processing of intonation with specific right-frontal contribution generalises across languages, while the semantic processing of tonal pitch is limited to tonal language speakers.

ACKNOWLEDGMENTS

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 INTERESTS

The authors declare that there are no conflicts of interest.

DATA AVAILABILITY STATEMENT

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

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

- Baddeley, A. (1992). Working memory. *Science*, 255, 556–559. <https://doi.org/10.1126/science.1736359>
- Baddeley, A. (2003a). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4(10), 829–839. <https://doi.org/10.1038/nrn1201>
- Baddeley, A. (2003b). Working memory and language: An overview. *Journal of Communication Disorders*, 36, 189–208. [https://doi.org/10.1016/S0021-9924\(03\)00019-4](https://doi.org/10.1016/S0021-9924(03)00019-4)
- 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>
- Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, 15(11), 527–536. <https://doi.org/10.1016/j.tics.2011.10.001>
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19, 2767–2796. <https://doi.org/10.1093/cercor/bhp055>
- Bozic, M., Tyler, L. K., Ives, D. T., Randall, B., & Marslen-Wilson, W. D. (2010). Bihemispheric foundations for human speech comprehension. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 17439–17444. <https://doi.org/10.1073/pnas.1000531107>
- Brown, S., Laird, A. R., Pfordresher, P. Q., Thelen, S. M., Turkeltaub, P., & Liotti, M. (2009). The somatotopy of speech: Phonation and articulation in the human motor cortex. *Brain and Cognition*, 70(1), 31–41. <https://doi.org/10.1016/j.bandc.2008.12.006>
- Brown, S., Ngan, E., & Liotti, M. (2008). A larynx area in the human motor cortex. *Cerebral Cortex*, 18, 837–845. <https://doi.org/10.1093/cercor/bhm131>
- Camilleri, J. A., Müller, V. I., Fox, P., Laird, A. R., Hoffstaedter, F., Kalenscher, T., & Eickhoff, S. B. (2018). Definition and characterization of an extended multiple-demand network. *NeuroImage*, 165, 138–147. <https://doi.org/10.1016/j.neuroimage.2017.10.020>
- Chao, Y. R. (1968). *A grammar of spoken Chinese*. Berkeley, CA: University of California Press.
- Charest, I., Pernet, C., Latinus, M., Crabbe, F., & Belin, P. (2013). Cerebral processing of voice gender studied using a continuous carryover fMRI design. *Cerebral Cortex*, 23(4), 958–966. <https://doi.org/10.1093/cercor/bhs090>
- Cole, J. (2015). Prosody in context: A review. *Language, Cognition and Neuroscience*, 30, 1–31. <https://doi.org/10.1080/23273798.2014.963130>

- Copland, D. A., de Zubicaray, G. I., McMahon, K., Wilson, S. J., Eastburn, M., & Chenery, H. J. (2003). Brain activity during automatic semantic priming revealed by event-related functional magnetic resonance imaging. *NeuroImage*, 20(1), 302–310. [https://doi.org/10.1016/S1053-8119\(03\)00279-9](https://doi.org/10.1016/S1053-8119(03)00279-9)
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3(3), 201–215. <https://doi.org/10.1038/nrn755>
- 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>
- Dichter, B. K., Breshears, J. D., Leonard, M. K., & Chang, E. F. (2018). The control of vocal pitch in human laryngeal motor cortex. *Cell*, 174, 21–31. <https://doi.org/10.1016/j.cell.2018.05.016>
- Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, 14, 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>
- 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>
- Fournier, R., Gussenhoven, C., Jensen, O., & Hagoort, P. (2010). Lateralization of tonal and intonational pitch processing: An MEG study. *Brain Research*, 1328, 79–88. <https://doi.org/10.1016/j.brainres.2010.02.053>
- 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. (2012). The cortical language circuit: From auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, 16(5), 262–268.
- Friederici, A. D., & Alter, K. (2004). Lateralization of auditory language functions: A dynamic dual pathway model. *Brain and Language*, 89(2), 267–276.
- Gandour, J., Wong, D., Dziedzic, M., Lowe, M., Tong, Y., & Li, X. (2003). A cross-linguistic fMRI study of perception of intonation and emotion in Chinese. *Human Brain Mapping*, 18, 149–157. <https://doi.org/10.1002/hbm.10088>
- Gandour, J., & Krishnan, A. (2016). Processing tone languages. In G. Hickok & S. L. Small (Eds.), *Neurobiology of language* (pp. 1095–1107). Cambridge: Academic Press. <https://doi.org/10.1016/B978-0-12-407794-2.00087-0>
- 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.
- Geranmayeh, F., Brownsett, S. L. E., & Wise, R. J. S. (2014). Task-induced brain activity in aphasic stroke patients: What is driving recovery? *Brain*, 137, 2632–2648. <https://doi.org/10.1093/brain/awu163>
- Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24, 95–112. <https://doi.org/10.1007/BF02289823>
- Grefkes, C., & Fink, G. R. (2005). The functional organization of the intraparietal sulcus in humans and monkeys. *Journal of Anatomy*, 207, 3–17. <https://doi.org/10.1111/j.1469-7580.2005.00426.x>
- Gussenhoven, C., & Chen, A. (2000). Universal and language-specific effects in the perception of question intonation. In B. Yuan, T. Huang, & X. Tang (Eds.), *Proceedings of the Sixth International Conference on Spoken Language Processing* (pp. 91–94). Beijing: China Military Friendship Publish.
- Hallé, P. A., Chang, Y. C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, 32(3), 395–421. [https://doi.org/10.1016/S0095-4470\(03\)00016-0](https://doi.org/10.1016/S0095-4470(03)00016-0)
- Hamilton, A. F. d. C., & Grafton, S. T. (2007). The motor hierarchy: From kinematics to goals and intentions. In P. Haggard, Y. Rossetti, & M. Kawato (Eds.), *Sensorimotor foundations of higher cognition* (pp. 381–408). Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199231447.003.0018>
- Hartwigsen, G. (2018). Flexible redistribution in cognitive networks. *Trends in Cognitive Sciences*, 22, 687–698. <https://doi.org/10.1016/j.tics.2018.05.008>
- Hartwigsen, G., Baumgaertner, A., Price, C. J., Koehnke, M., Ulmer, S., & Siebner, H. R. (2010). Phonological decisions require both the left and right supramarginal gyri. *Proceedings of the National Academy of Sciences of the United States of America*, 107(38), 16494–16499. <https://doi.org/10.1073/pnas.1008121107>
- 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. (2018). Neural bases of social communicative intentions in speech. *Social Cognitive and Affective Neuroscience*, 13, 604–615. <https://doi.org/10.1093/scan/nsy034>
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, 92, 67–99. <https://doi.org/10.1016/j.cognition.2003.10.011>
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8, 393–402. <https://doi.org/10.1038/nrn2113>
- Ho, A. T. (1976). Project on linguistic analysis Mandarin tones in relation to sentence intonation and grammatical structure. *Journal of Chinese Linguistics*, 4(1), 1–13.
- Humphreys, G. F., Hoffman, P., Visser, M., Binney, R. J., & Lambon Ralph, M. A. (2015). Establishing task- and modality-dependent dissociations between the semantic and default mode networks. *Proceedings of the National Academy of Sciences of the United States of America*, 112, 7857–7862. <https://doi.org/10.1073/pnas.1422760112>
- Husain, F. T., Fromm, S. J., Pursley, R. H., Hosey, L. A., Braun, A. R., & Horwitz, B. (2006). Neural bases of categorization of simple speech and nonspeech sounds. *Human Brain Mapping*, 27(8), 636–651. <https://doi.org/10.1002/hbm.20207>
- 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.
- Kung, C., Chwilla, D. J., & Schriefers, H. (2014). The interaction of lexical tone, intonation and semantic context in on-line spoken word

- recognition: An ERP study on Cantonese Chinese. *Neuropsychologia*, 53(1), 293–309.
- 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>
- Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2016). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, 18(1), 42–55. <https://doi.org/10.1038/nrn.2016.150>
- Levy, D. A., Bayley, P. J., & Squire, L. R. (2004). The anatomy of semantic knowledge: Medial vs. lateral temporal lobe. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 6710–6715. <https://doi.org/10.1073/pnas.0401679101>
- Li, X., Gandour, J., Talavage, T., & Wong, D. (2010). Hemispheric asymmetries in phonological processing of tones vs. segmental units. *NeuroReport*, 21(10), 690–694.
- Li, X., Gandour, J., Talavage, T., Wong, D., Dziedzic, M., Lowe, M., & Tong, Y. (2003). Selective attention to lexical tones recruits left dorsal frontoparietal network. *Neuroreport*, 14(17), 2263–2266. <https://doi.org/10.1097/00001756-200312020-00025>
- 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>
- Liang, J., & van Heuven, V. J. (2007). Chinese tone and intonation perceived by L1 and L2 listeners. In C. Gussenhoven & T. Riad (Eds.), *Phonology and phonetics* (Vol. 2, pp. 27–61). Berlin/NewYork: Mouton de Gruyter.
- Liégeois, F., Connelly, A., Salmond, C. H., Gadian, D. G., Vargha-Khadem, F., & Baldeweg, T. (2002). A direct test for lateralization of language activation using fMRI: Comparison with invasive assessments in children with epilepsy. *NeuroImage*, 17, 1861–1867. <https://doi.org/10.1006/nimg.2002.1327>
- Liu, F., & Xu, Y. (2005). Parallel encoding of focus and interrogative meaning in Mandarin intonation. *Phonetica*, 62, 70–87. <https://doi.org/10.1159/000090090>
- Liu, M., Chen, Y., & Schiller, N. O. (2016). Online processing of tone and intonation in Mandarin: Evidence from ERPs. *Neuropsychologia*, 91, 307–317.
- Ma, J. K.-Y., Ciocca, V., & Whitehill, T. L. (2011). The perception of intonation questions and statements in Cantonese. *The Journal of the Acoustical Society of America*, 129(2), 1012–1023. <https://doi.org/10.1121/1.3531840>
- Malins, J. G., & Joanisse, M. F. (2012). Setting the tone: An ERP investigation of the influences of phonological similarity on spoken word recognition in Mandarin Chinese. *Neuropsychologia*, 50(8), 2032–2043.
- Merrill, J., Sammler, D., Bangert, M., Goldhahn, D., Lohmann, G., Turner, R., & Friederici, A. D. (2012). Perception of words and pitch patterns in song and speech. *Frontiers in Psychology*, 3, 76. <https://doi.org/10.3389/fpsyg.2012.00076>
- 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., Olan, C. A., Auerbach, E., Strupp, J., Harel, N., & Ugurbil, 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>
- Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J. B. (2005). Valid conjunction inference with the minimum statistic. *NeuroImage*, 25, 653–660. <https://doi.org/10.1016/j.neuroimage.2004.12.005>
- Noonan, K. A., Jefferies, E., Visser, M., & Lambon Ralph, M. A. (2013). Going beyond inferior prefrontal involvement in semantic control: Evidence for the additional contribution of dorsal angular gyrus and posterior middle temporal cortex. *Journal of Cognitive Neuroscience*, 25(11), 1824–1850. https://doi.org/10.1162/jocn_a_00442
- 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)
- Paulmann, S. (2016). The Neurocognition of prosody. In G. Hickok & S. L. Small (Eds.), *Neurobiology of language* (pp. 1109–1120). Amsterdam: Elsevier. <https://doi.org/10.1016/b978-0-12-407794-2.00088-2>
- Peng, G., Zheng, H. Y., Gong, T., Yang, R. X., Kong, J. P., & Wang, W. S. Y. (2010). The influence of language experience on categorical perception of pitch contours. *Journal of Phonetics*, 38, 616–624. <https://doi.org/10.1016/j.wocn.2010.09.003>
- Poeppl, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as “asymmetric sampling in time.”. *Speech Communication*, 41(1), 245–255. [https://doi.org/10.1016/S0167-6393\(02\)00107-3](https://doi.org/10.1016/S0167-6393(02)00107-3)
- Price, C. J. (2010). The anatomy of language: A review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences*, 1191, 62–88. <https://doi.org/10.1111/j.1749-6632.2010.05444.x>
- Romero, L., Walsh, V., & Papagno, C. (2006). The neural correlates of phonological short-term memory: A repetitive transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience*, 18(7), 1147–1155. <https://doi.org/10.1162/jocn.2006.18.7.1147>
- 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>
- Seghier, M. L. (2013). The angular gyrus: Multiple functions and multiple subdivisions. *The Neuroscientist*, 19, 43–61. <https://doi.org/10.1177/1073858412440596>
- Seghier, M. L., Fagan, E., & Price, C. J. (2010). Functional subdivisions in the left angular Gyrus where the semantic system meets and diverges from the default network. *Journal of Neuroscience*, 30(50), 16809–16817. <https://doi.org/10.1523/jneurosci.3377-10.2010>
- Srinivasan, R. J., & Massaro, D. W. (2003). Perceiving prosody from the face and voice: Distinguishing statements from echoic questions in English. *Language and Speech*, 46(1), 1–22.
- 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>
- 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>

- von Kriegstein, K., Eger, E., Kleinschmidt, A., & Giraud, A. L. (2003). Modulation of neural responses to speech by directing attention to voices or verbal content. *Cognitive Brain Research*, 17(1), 48–55. [https://doi.org/10.1016/S0926-6410\(03\)00079-X](https://doi.org/10.1016/S0926-6410(03)00079-X)
- 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>
- Witteman, J., van Ijzendoorn, M. H., van de Velde, D., van Heuven, V. J. J. P., & Schiller, N. O. (2011). The nature of hemispheric specialization for linguistic and emotional prosodic perception: A meta-analysis of the lesion literature. *Neuropsychologia*, 49(13), 3722–3738. <https://doi.org/10.1016/j.neuropsychologia.2011.09.028>
- Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience*, 170(1), 223–231.
- Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.
- Yuan, J. (2004). Perception of Mandarin intonation. In *International symposium on Chinese spoken language processing* (pp. 45–48). Piscataway: IEEE.
- Yuan, J., & Jurafsky, D. (2005). Detection of questions in Chinese conversational speech. In *Proceedings of ASRU 2005: 2005 IEEE Automatic Speech Recognition and Understanding Workshop*. Piscataway: IEEE. <https://doi.org/10.1109/ASRU.2005.1566536>
- Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, 6(1), 37–46. [https://doi.org/10.1016/S1364-6613\(00\)01816-7](https://doi.org/10.1016/S1364-6613(00)01816-7)

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Chien P-J, Friederici AD, Hartwigsen G, Sammler D. Neural correlates of intonation and lexical tone in tonal and non-tonal language speakers. *Hum Brain Mapp.* 2020;1–17. <https://doi.org/10.1002/hbm.24916>