

# Hemispheric asymmetries in phonological processing of tones versus segmental units

Xiaojian Li<sup>a</sup>, Jackson T. Gandour<sup>b</sup>, Thomas Talavage<sup>c,d</sup>, Donald Wong<sup>e</sup>, Angela Hoffa<sup>c</sup>, Mark Lowe<sup>f</sup> and Mario Dzemidzic<sup>g</sup>

**The aim of this functional magnetic resonance imaging study is to identify neuroanatomical substrates underlying phonological processing of segmental (consonant, rhyme) and suprasegmental (tone) units. An auditory verbal recognition paradigm was used in which native speakers of Mandarin Chinese were required to match a phonological unit that occurs in a list of three syllables to the corresponding unit of a following probe. The results show that hemispheric asymmetries arise depending on the type of phonological unit. In direct contrasts between phonological units, tones, relative to consonants and rhymes, yield increased activation in frontoparietal areas of the right hemisphere. This finding indicates that the cortical circuitry subserving lexical tones differs from that of consonants or rhymes. *NeuroReport* 21:690–694 © 2010 Wolters Kluwer Health | Lippincott Williams & Wilkins.**

Tone languages give us a unique window on the neurobiology of suprasegmental versus segmental processing because of the phonemic status of pitch variations at the level of the syllable or morpheme. For instance, in Mandarin Chinese, consonants and rhymes differ in duration and the order in which their information unfolds over the course of a syllable. Both are the units of segmental information. Rhymes and tones, in contrast, are coterminous in duration and order in the syllable, but rhymes are segmental and tones are suprasegmental [1].

Functional magnetic resonance imaging (fMRI) suggests that speech prosody perception, including lexical tone, is mediated primarily by the right hemisphere (RH), but is lateralized to the left hemisphere for postperceptual processing depending on its linguistic status in a particular language [2–5]. Differential patterns of cortical activation, however, are not driven by language experience alone. They may also be driven by differences in acoustic features associated with specific types of phonological units. For instance, it has been shown that hemispheric specialization of consonants is dissociable from vowels during phonetic discrimination [6]. In the production of Mandarin tones, tones elicit more activity in the RH than vowels [7]. This is especially remarkable as tones are primarily realized on vowels or rhymes with

*NeuroReport* 2010, 21:690–694

**Keywords:** auditory, functional magnetic resonance imaging, language, lexical tone, Mandarin, prosody, selective attention, verbal working memory

<sup>a</sup>Research Center for Psychological Application, South China Normal University, Guangzhou, PR China, <sup>b</sup>Department of Speech Language Hearing Sciences, <sup>c</sup>School of Electrical and Computer Engineering, <sup>d</sup>Weldon School of Biomedical Engineering, Purdue University, West Lafayette, <sup>e</sup>Departments of Anatomy and Cell Biology, <sup>f</sup>Imaging Institute, Cleveland Clinic Foundation, Cleveland, Ohio and <sup>g</sup>Neurology and Radiology, Indiana University School of Medicine, Indianapolis, Indiana, USA

Correspondence to Jackson T. Gandour, PhD, Department of Speech Language Hearing Sciences, Purdue University, 1353 Heavilon Hall, West Lafayette, IN 47907-2038, USA

Tel: +1 765 494 3821; fax: +1 765 494 0771; e-mail: gandour@purdue.edu

Received 22 March 2010 accepted 15 April 2010

which they are associated. In Mandarin, early event-related brain potentials show RH dominance regardless of linguistic function (tone, intonation) [8], but opposite patterns of hemispheric dominance for tones (RH) versus consonants (left hemisphere) [9]. In Japanese, near infrared spectroscopy shows stronger left-dominant and right-dominant responses for vowels and prosodic contrasts in sentence type, respectively [10]. Using fMRI, it has been shown in Mandarin that selective attention to a target tone of a syllable relative to the whole syllable recruits a left dorsal frontoparietal network [11], and that distracters between the target tone and its probe may induce articulatory encoding with engagement of a frontocerebellar network including a left dorsal frontal region [12]. Whether hemispheric specialization of tones is dissociable from that of rhymes is an empirical question.

The aim of this fMRI study is to identify the neuroanatomical substrates subserving phonological processing by providing pairwise contrasts between suprasegmental (tone, T) and segmental (consonant, C; rhyme, R) phonological units concurrently. An auditory recognition paradigm is used in which individuals are asked to match a phonological unit within a three-syllable list to the corresponding unit of a following probe syllable. Two types of matching sequences are distinguished by either fixing or randomly varying the position of syllables containing the target units in a three-syllable list.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Website ([www.neuroreport.com](http://www.neuroreport.com)).

Matching judgments for random sequences, compared with fixed, are expected to increase the neural activity for encoding of phonological units and their associated working memory. We further expect to elicit differential patterns of hemispheric asymmetry as a function of the type of phonological unit – segmental versus suprasegmental.

## Methods

### Participants

Twelve adult native speakers of Mandarin (six male; six female) from mainland China, ranging in age from 23–32 years, participated in this study. All the participants were strongly right-handed (laterality quotient:  $M = 94\%$ ,  $SD = 9$ ) [13]; and exhibited normal hearing sensitivity. All the participants gave informed consent in compliance with a protocol approved by the Institutional Review Board of Indiana University-Purdue University Indianapolis and Clarian Health.

### Stimuli

A stimulus list of three Mandarin monosyllables followed by a probe monosyllable made up a sequence for each trial. All syllables (maximum duration = 450 ms) were produced by a native Chinese male speaker. A matching sequence contained a target syllable in the stimulus list sharing a phonological unit (consonant; rhyme; or tone) in common with the probe. In a nonmatching sequence, none of the three syllables in the stimulus list contained a phonological unit in common with the following probe.

Two types of matching sequences were constructed based on the positions of the target syllables (Table 1). A fixed matching sequence located the target syllable in the last position of the stimulus list. A corresponding random matching sequence was derived from a fixed matching sequence by varying the target syllable in a random position (first, second, or last) of the stimulus list. Two types of nonmatching sequences were also similarly constructed (see text, Supplemental digital content 1, <http://links.lww.com/WNR/A57> for listeners' guide to samples; see Audio, Supplemental digital content 2, <http://links.lww.com/WNR/A58>, <http://links.lww.com/WNR/A59>, <http://links.lww.com/WNR/A60>, <http://links.lww.com/WNR/A61>, <http://links.lww.com/WNR/A62>, <http://links.lww.com/WNR/A63>, <http://links.lww.com/WNR/A64>, <http://links.lww.com/WNR/A65>, to listen to samples). There were a total of 16 fixed matching sequences and corresponding random matching sequences, sharing consonant, rhyme, and tone, respectively (see Table, Supplemental digital content 3, <http://links.lww.com/WNR/A66> for complete list of sequences). Within each sequence, no adjacent syllables formed a disyllabic word. Occurrences of different consonants, rhymes, and tones were balanced across the sequences.

### Task procedure

There were three paired experimental tasks designed to contrast random (r) versus fixed (f) target positions on the three phonological units: Cr versus Cf, Rr versus Rf,

**Table 1** Sample auditory stimuli for consonant, rhyme, and tone matching tasks

Unit	Position	Syllable list			Probe	Matching
		P1	P2	P3		
Consonant	Random	bian <sup>4</sup>	wang <sup>1</sup>	she <sup>2</sup>	<b>ba</b> <sup>3</sup>	Yes
		cen <sup>2</sup>	<b>zhai</b> <sup>4</sup>	feng <sup>3</sup>	<b>zhe</b> <sup>1</sup>	Yes
		liu <sup>1</sup>	zuo <sup>2</sup>	song <sup>4</sup>	<b>ku</b> <sup>3</sup>	No
Consonant	Fixed	she <sup>2</sup>	wang <sup>1</sup>	<b>bian</b> <sup>4</sup>	<b>ba</b> <sup>3</sup>	Yes
		cen <sup>2</sup>	feng <sup>3</sup>	<b>zhai</b> <sup>4</sup>	<b>zhe</b> <sup>1</sup>	Yes
		liu <sup>1</sup>	song <sup>4</sup>	zuo <sup>2</sup>	<b>kui</b> <sup>3</sup>	No
Rhyme	Random	xi <sup>1</sup>	tuan <sup>2</sup>	lou <sup>3</sup>	pi <sup>4</sup>	Yes
		zhuo <sup>1</sup>	pin <sup>4</sup>	<b>da</b> <sup>3</sup>	<b>ma</b> <sup>2</sup>	Yes
		zou <sup>1</sup>	se <sup>4</sup>	nian <sup>2</sup>	<b>kuang</b> <sup>3</sup>	No
Rhyme	Fixed	tuan <sup>2</sup>	lou <sup>3</sup>	xi <sup>1</sup>	pi <sup>4</sup>	Yes
		zhuo <sup>1</sup>	pin <sup>4</sup>	<b>da</b> <sup>3</sup>	<b>ma</b> <sup>2</sup>	Yes
		se <sup>4</sup>	nian <sup>2</sup>	zou <sup>1</sup>	<b>kuang</b> <sup>3</sup>	No
Tone	Random	bao <sup>1</sup>	hun <sup>4</sup>	mu <sup>2</sup>	zhi <sup>1</sup>	Yes
		kai <sup>3</sup>	tu <sup>2</sup>	huang <sup>4</sup>	mei <sup>2</sup>	Yes
		hei <sup>1</sup>	liu <sup>2</sup>	gong <sup>3</sup>	can <sup>4</sup>	No
Tone	Fixed	hun <sup>4</sup>	mu <sup>2</sup>	bao <sup>1</sup>	zhi <sup>1</sup>	Yes
		kai <sup>3</sup>	huang <sup>4</sup>	tu <sup>2</sup>	mei <sup>2</sup>	Yes
		hei <sup>1</sup>	gong <sup>3</sup>	liu <sup>2</sup>	can <sup>4</sup>	No

Chinese syllables are written in Pinyin transcription. Phonological units (consonant, rhyme, tone) in the probe are marked in bold; a target unit from the syllable list is also marked in bold when it matches the corresponding unit in the probe. P1, P2, and P3 represent the first, second, and third positions, respectively, in the syllable list. Superscript numbers (1-4) refer to the four Mandarin lexical tones: high level (1), high rising (2), low falling rising (3), and high falling (4).

and Tr versus Tf. Tasks with fixed target positions served as the control (baseline) for tasks with random target positions. In the Cr task, for example, the participants were instructed to judge whether any of the three syllables in the list had a consonant matching to the consonant of the probe. In the Cf task, the participants were instructed to judge whether the consonant of the last syllable matched that of the probe, ignoring the first and second syllables of the list. They responded by pressing the left mouse button. Instructions were delivered in Mandarin through the headphones immediately preceding each task block: for example, Cr, “consonant-random position”; Rf, “rhyme-fixed position”. Before imaging, participants were trained to a high level of accuracy ( $\geq 85\%$ ) on all tasks using different stimuli from those presented during fMRI.

Three fMRI scans were conducted, each focusing on a single phonological unit (consonant, rhyme, and tone). In each 7.5 min scan, a pair of tasks (e.g. Cr and Cf) were presented in blocked format (36 s) in an alternating boxcar design with 18-s rest periods separating the task blocks. A block design paradigm was chosen to enhance statistical power of detection. There were eight task blocks in a scan, four per task (e.g. four Cr and four Cf blocks). Each block contained eight 4.5 s trials, four matching and four nonmatching sequences, presented in random order. For each of the three scans, there were 32 trials per fixed and random matching sequences, respectively, for a sum total of 64 trials. The cumulative total of trials across the three scans was 192: (matching + nonmatching = 8 trials)  $\times$  (4 blocks)  $\times$  (2 positions)  $\times$  (3 units). The order

of imaging scans for phonological units (consonant, rhyme, and tone) and task blocks (random, fixed) within each scan was counterbalanced across the participants.

The timing architecture of a trial consisted of the syllable list + probe (2300 ms, on average) and the response interval (2200 ms). Stimulus onset asynchrony within the syllable list was 500 ms. A silent interval of 350 ms was inserted between the syllable list and its probe. Each sequence of four syllables (list + probe) fell within the span for short-term memory and attention [14,15].

### Image acquisition

Imaging was performed on a 1.5T Signa GE LX Horizon scanner (Waukesha, Wisconsin, USA) equipped with a birdcage transmit-receive radiofrequency head coil. Blood oxygenation level dependent contrast sensitive functional volumes were acquired with an echo-planar imaging pulse sequence (gradient echo; 2.25s repetition time; 50 ms echo time; 90° flip angle; 64 × 64 acquisition matrix; 24 × 24 cm field of view, sixteen 7.5 mm thick contiguous axial slices). Before fMRI, whole-brain high-resolution anatomic images were acquired in 124 contiguous axial slices using a 3-D spoiled gradient-recalled acquisition in the steady state sequence for purposes of anatomic localization and transformation to a standard stereotactic system.

### Image analysis

Image analysis was conducted using the SPM5 software package (Wellcome Department of Imaging Neuroscience, University College, London, UK). For each participant, functional image volumes were corrected for slice acquisition timing differences and rigid-body realigned to the initial volume of the first functional scan. Each participant's high-resolution anatomical images were coregistered to the mean image of all the three functional scans and segmented into tissue components. Spatial parameters generated during the segmentation were applied to transform functional volumes into the Montreal Neurological Institute space, and then resampled to 2 mm (isotropic) voxels and smoothed by a 6 mm full-width at half-maximum Gaussian kernel.

Participants' responses to various stimuli were modeled using SPM's canonical hemodynamic response function and its time and dispersion derivatives to account for variations in response onsets and durations. The model also included six-movement parameter regressors obtained during realignment, allowing for residual movement-induced effects. The effects of serial correlations in fMRI time series were taken into account using a first order autoregressive model, whereas a high-pass filter with a cutoff of 1/128 Hz was applied to each voxel's time series to remove the low frequency noise.

For consonant, rhyme, and tone scans, respectively, summary contrast images representing average activation differences between random and fixed matching se-

quences were calculated across blocks (Cr > Cf, Rr > Rf, and Tr > Tf). Comparisons of Cr, Rr, or Tr with rest were done for calibration purposes. In addition, contrast with the rest blocks also facilitated better capture of the hemodynamic response for each task block of interest. For direct comparisons of phonological units (Consonant vs. Rhyme; Consonant vs. Tone; Rhyme vs. Tone), three summary contrast images representing average activation differences between phonological units [(Cr > Cf)–(Rr > Rf); (Cr > Cf)–(Tr > Tf); (Rr > Rf)–(Tr > Tf)] were similarly calculated across the blocks within and between related functional scans.

Statistical inferences for each phonological unit relative to rest were made using a Gaussian field theory derived cluster level significance ( $P_{\text{cluster}} < 0.05$ ), corrected for multiple comparisons in a search volume comprising all voxels within SPM's gray matter template after smoothing with a 6 mm full-width at half-maximum Gaussian kernel. The voxelwise height threshold for comparisons between random and fixed conditions within phonological units was set at  $P_{\text{voxel}} < 0.001$  (uncorrected), whereas comparisons between phonological units were conducted using  $P_{\text{voxel}} < 0.005$  (uncorrected).

### Results

A comparison of random versus fixed matching positions yielded numerous common areas of increased activity (Table 2), regardless of phonological unit, in frontal and parietal areas bilaterally, and in the anterior insula, frontal operculum, and anterior cingulate gyrus (see Figure, Supplemental digital content 4, <http://links.lww.com/WNR/A67> for activation maps). A direct comparison of phonological units showed significant frontoparietal activations predominantly in the RH for tone relative to consonant or rhyme (Fig. 1; see Table, Supplemental digital content 5, <http://links.lww.com/WNR/A68> for summary of significant clusters of activation). In the case of tone versus consonant, activity was centered in dorsal aspects of the inferior frontal gyrus in the RH near the junction of the inferior frontal/precentral sulci. Activations were observed more extensively in the right inferior parietal lobule. In the case of tone versus rhyme, there were two activation foci in the right frontal lobe, one localized predominantly in the pars opercularis, with the other centered more anteriorly in the inferior frontal sulcus. The peak focus of activation in the right inferior parietal lobule was centered dorsally near the intraparietal sulcus.

Two-way (position × unit) mixed model analyses of variance of reaction time and response accuracy showed that regardless of phonological unit, reaction time and correct percentage were larger in the random-matching than in the fixed-matching task (see Figure, Supplemental digital content 6, <http://links.lww.com/WNR/A69> for display of reaction time and correct percentage by unit and task). Post hoc multiple comparisons ( $\alpha_{\text{Bonferroni}} = 0.05$ ) further

**Table 2 Summary of significant clusters of activation for single subtraction comparisons between random and fixed conditions per phonological unit**

Side	Brain region	BA	MNI coordinates (mm)				Z
			$k_E$	x	y	z	
<b>Cr &gt; Cf</b>							
L	Inferior frontal gyrus	44	1071	-36	8	24	5.05
L	Inferior parietal lobule	40	411	-32	-56	34	4.93
R	Anterior cingulate gyrus	32	275	4	20	44	3.72
R	Insula, frontal operculum	13/47	267	34	30	6	4.26
R	Middle frontal gyrus	6	199	32	-2	44	4.59
R	Inferior parietal lobule	40	133	30	-56	40	4.07
<b>Rr &gt; Rf</b>							
L	Inferior frontal gyrus	44	1828	-36	22	26	4.77
L	Inferior parietal lobule	40	823	-26	-66	44	5.39
L	Anterior cingulate gyrus	32	660	-4	22	42	4.69
R	Insula, frontal operculum	13/47	289	32	26	-4	4.56
M	Cerebellar vermis	201	10	-78	-28	4.14	
R	Inferior parietal lobule	40	152	36	-54	40	4.16
R	Middle frontal gyrus	6	98	28	-4	38	3.79
<b>Tr &gt; Tf</b>							
L	Inferior/middle frontal gyrus	44/9	2702	-44	8	30	5.50
L	Inferior parietal lobule	40	1561	-28	-64	38	5.02
M	Cerebellar vermis	236	-2	-78	-30	4.98	
R	Insula, frontal operculum	13/47	2233	30	30	-4	5.48
R	Inferior parietal lobule	40	1533	40	-60	48	5.45
R	Middle frontal gyrus	6	238	36	-4	48	4.34
M	Anterior cingulate gyrus	32	887	0	18	50	5.17

$k_E$  refers to cluster extent. Coordinates (x, y, z) of peak activation are expressed in millimeters in the MNI space. Z refers to peak Z-score value within a cluster. Statistical significance was inferred at the cluster level,  $P_{cluster} < 0.05$ , after correcting for multiple comparisons within gray matter voxels in the whole brain. Voxelwise height threshold,  $P = 0.001$ .

BA, Brodmann area; Cf, consonant fixed; Cr, consonant random; L, left; M, medial; MNI, Montreal Neurological Institute; Rf, rhyme fixed; R, right; Rr, rhyme random; Tf, tone fixed; Tr, tone random.

showed that response accuracy for both rhyme and tone was higher than consonant across tasks, whereas no difference was observed between rhyme and tone.

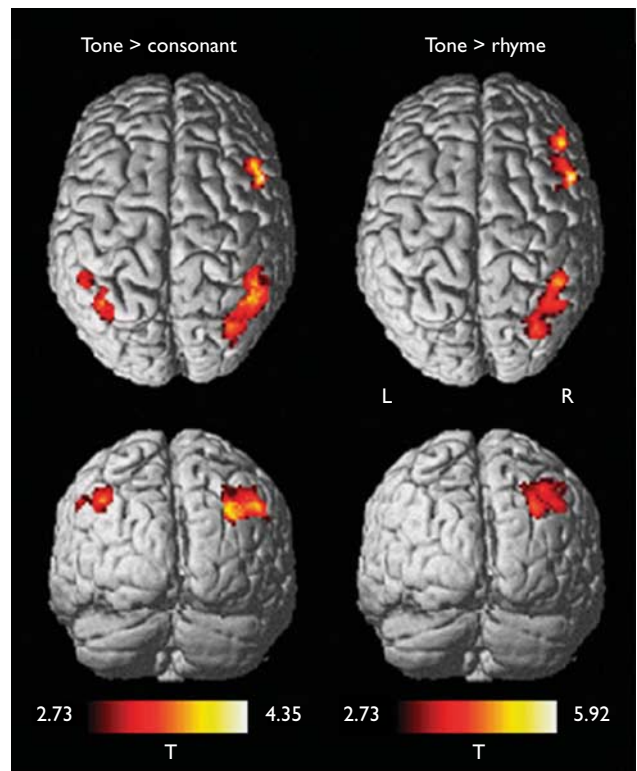
**Discussion**

Using an auditory immediate recognition paradigm [16], it is shown that hemispheric asymmetries arise as a function of the type of phonological unit. In direct contrasts of phonological units, tone, as compared with consonant or rhyme, shows increased activation in frontoparietal areas of the RH. This rightward asymmetry of a suprasegmental unit (tone) whose primary acoustic correlate is voice fundamental frequency, as compared with segmental units (consonant, rhyme), is congruent with the well-established role of the RH in mediating speech prosody [3,17].

**Neural substrates of phonological processing**

Direct contrasts between segmental and suprasegmental units [(Tr-Tf) > (Cr-Cf); (Tr-Tf) > (Rr-Rf)] show rightward asymmetry of the frontoparietal network for tones, as compared with consonants and rhymes. As the task paradigm is identical across the units for random-matching and fixed-matching conditions, all three units (consonant, rhyme, and tone) recruit key structures of a frontoparietal network consistent with the extant literature on verbal short-term memory. Using direct contrasts,

**Fig. 1**



Rendered statistical maps showing significant differences in blood oxygenation level-dependent response of tone over consonant [(Tr-Tf) > (Cr-Cf); left column] and tone over rhyme [(Tr-Tf) > (Rr-Rf); right column]. Only voxels found within significant clusters ( $P < 0.05$ , corrected for multiple comparisons) are shown. The color scale depicts the range of t-statistic values. Voxelwise display threshold,  $P = 0.005$ , uncorrected;  $Z > 3.72$ . L, left; R, right; T, tone.

we are able to observe neural activity specific to tonal encoding and its separate memory processes, as compared with segmental.

However, the limited temporal resolution of fMRI does not permit us to tease apart specific processes associated with verbal working memory, that is, encoding, storage, retrieval, comparison, matching, decision making, etc. All of these processes, perceptual and postperceptual, are processed rapidly within hundreds of milliseconds [18]. For instance, the rightward asymmetry of the frontoparietal network for tones, relative to consonants and rhymes, may be attributable to the well-established role of the RH in mediating pitch. This would be consistent with the view that hemispheric asymmetries arise from low-level features of sounds [19,20]. In contrast, the observed rightward asymmetry for tones in the frontal lobe is also consistent with the view that dorsolateral prefrontal cortex carries out temporal integration of information when making stimulus comparisons in short-term memory, and that it actively organizes sequences of responses based on explicit retrieval of information from posterior cortical association systems [21].

Although unable to fractionate temporal stages of phonological processing in this study, these data point to a fruitful line of research using magnetoencephalography to show differences in spatiotemporal dynamics associated with suprasegmental and segmental information.

### Effects of task performance on brain activation patterns

It is unlikely that differences in task performance can account for the differential patterns of activation in tone versus consonant or rhyme in right frontoparietal cortex. Chinese participants' reaction times are homogeneous irrespective of phonological unit. Reaction time is presumed to reflect decision-making processes, and seems to be positively correlated with increased activity in inferior frontal regions [22]. Yet we find no differences in reaction time among consonant, rhyme, and tone. However, response accuracy is observed to be higher for rhymes and tones than for consonants. This disparity is likely because of the relative degree-of-change over time in acoustic properties of rhymes and tones (slowly changing voice fundamental frequency and higher harmonics) versus onset consonants (rapidly changing bursts and formant transitions). Indeed, the perceptual trace of rapidly changing cues has been shown to decay faster in working memory [23]. Thus, the RH advantage for tone over rhyme cannot be accounted for by differences in decay rates. Instead, our findings argue for a view of working memory that emerges from the integrated action of neural processes subserving acoustic/auditory features associated with specific types of phonological units, that is, suprasegmental versus segmental.

### Conclusion

This study shows that neural circuitry subserving phonological processing is differentially engaged depending on whether the unit is segmental or suprasegmental. The rightward asymmetry in frontoparietal regions for tones, relative to consonants and rhymes, is consistent with the idea of differential hemispheric specialization on the basis of both attentional demands and perceptual cues.

### Acknowledgements

Sources of support: NIH research grants R01 DC04584 for J.T.G.; R01 EB03990 for T.M.T.; and NIH post-doctoral traineeship for (X.L.)

### References

- 1 Yip M. *Tone*. New York: Cambridge University Press; 2003.
- 2 Gandour JT, Tong Y, Wong D, Talavage T, Dzemidzic M, Xu Y, et al. Hemispheric roles in the perception of speech prosody. *Neuroimage* 2004; **23**:344–357.
- 3 Wildgruber D, Ackermann H, Kreifelts B, Ethofer T. Cerebral processing of linguistic and emotional prosody: fMRI studies. *Prog Brain Res* 2006; **156**:249–268.
- 4 Wong PC. Hemispheric specialization of linguistic pitch patterns. *Brain Res Bull* 2002; **59**:83–95.
- 5 Zatorre RJ, Gandour JT. Neural specializations for speech and pitch: moving beyond the dichotomies. *Philos Trans R Soc Lond B Biol Sci* 2008; **363**:1087–1104.
- 6 Jancke L, Wustenberg T, Scheich H, Heinze HJ. Phonetic perception and the temporal cortex. *Neuroimage* 2002; **15**:733–746.
- 7 Liu L, Peng D, Ding G, Jin Z, Zhang L, Li K, et al. Dissociation in the neural basis underlying Chinese tone and vowel production. *Neuroimage* 2006; **29**:515–523.
- 8 Ren GQ, Yang Y, Li X. Early cortical processing of linguistic pitch patterns as revealed by the mismatch negativity. *Neuroscience* 2009; **162**:87–95.
- 9 Luo H, Ni JT, Li ZH, Li XO, Zhang DR, Zeng FG, et al. Opposite patterns of hemisphere dominance for early auditory processing of lexical tones and consonants. *Proc Natl Acad Sci U S A* 2006; **103**:19558–19563.
- 10 Minagawa-Kawai Y, Naoi N, Kikuchi N, Yamamoto J, Nakamura K, Kojima S. Cerebral laterality for phonemic and prosodic cue decoding in children with autism. *Neuroreport* 2009; **20**:1219–1224.
- 11 Li X, Gandour J, Talavage T, Wong D, Dzemidzic M, Lowe M, et al. Selective attention to lexical tones recruits left dorsal frontoparietal network. *Neuroreport* 2003; **14**:2263–2266.
- 12 Li X, Wong D, Gandour J, Dzemidzic M, Tong Y, Talavage T, et al. Neural network for encoding immediate memory in phonological processing. *Neuroreport* 2004; **15**:2459–2462.
- 13 Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 1971; **9**:97–113.
- 14 Cowan N. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav Brain Sci* 2000; **24**:87–185.
- 15 Zhang G, Simon HA. STM capacity for Chinese words and idioms: chunking and acoustical loop hypotheses. *Mem Cognit* 1985; **13**:193–201.
- 16 Sternberg S. Memory-scanning: mental processes revealed by reaction-time experiments. *Am Sci* 1969; **57**:421–457.
- 17 Friederici AD, Alter K. Lateralization of auditory language functions: a dynamic dual pathway model. *Brain Lang* 2004; **89**:267–276.
- 18 Halgren E, Boujon C, Clarke J, Wang C, Chauvel P. Rapid distributed fronto-parieto-occipital processing stages during working memory in humans. *Cereb Cortex* 2002; **12**:710–728.
- 19 Poeppel D. The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. *Speech Commun* 2003; **41**:245–255.
- 20 Zatorre RJ, Belin P. Spectral and temporal processing in human auditory cortex. *Cereb Cortex* 2001; **11**:946–953.
- 21 Goldman-Rakic PS. The prefrontal landscape: implications of functional architecture for understanding human mentation and the central executive. In: Roberts A, Robbins T, Weiskrantz L, editors. *The prefrontal cortex: executive and cognitive functions*. New York: Oxford University Press; 1998, pp. 88–102.
- 22 Binder JR, Liebenthal E, Possing ET, Medler DA, Ward BD. Neural correlates of sensory and decision processes in auditory object identification. *Nat Neurosci* 2004; **7**:295–301.
- 23 Mirman D, Holt LL, McClelland JL. Categorization and discrimination of nonspeech sounds: differences between steady-state and rapidly-changing acoustic cues. *J Acoust Soc Am* 2004; **116**:1198–1207.