UC Irvine
UC Irvine Previously Published Works

Title
Neural predictors of auditory word learning

Permalink

Journal

ISSN

Authors
Mei, Leilei
Chen, Chuansheng
Xue, Gui
et al.

Publication Date

DOI

Copyright Information

Peer reviewed

Neural predictors of auditory word learning

Leilei Mei[a], Chuansheng Chen[b], Gui Xue[c], Qinghua He[a], Tian Li[a], Feng Xue[a], Qin Yang[a] and Qi Dong[a]

[a]State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, [b]Department of Psychology and Social Behavior, University of California, Irvine and [c]Department of Psychology, University of Southern California, Los Angeles, USA

Correspondence to Dr Qi Dong, State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China
Tel: +86 10 58807615; fax: +86 10 58807615; e-mail: dongqi@bnu.edu.cn

The present fMRI study aimed to identify neurofunctional predictors of auditory word learning. Twenty-four native Chinese speakers were trained to learn a logographic artificial language (LAL) for 2 weeks and their behavioral performance was recorded. Participants were also scanned before and after the training while performing a passive listening task. Results showed that, compared to 'poor' learners (those whose performance was below average during the training), 'good' (i.e. above-average) learners showed more activation in the left MTG/STS and less activation in the right IFG during the pretraining scan. These results confirmed the hypothesis that preexisting individual differences in neural activities can predict the efficiency in learning words in a new language. *NeuroReport* 19:215–219 © 2008 Wolters Kluwer Health | Lippincott Williams & Wilkins.

Keywords: fMRI, individual differences, language learning, speech perception

Introduction

Cognitive neuroscientists have traditionally been interested in the general neural mechanisms of language processing and learning. Several recent studies, however, have found remarkable interindividual variability in neural responses to language processing [1,2], especially when processing a nonfluent or new language [3,4]. Two lines of research have further shown that such individual differences are linked to efficiency in language learning. First, several studies have revealed that neural changes due to language training were correlated with improvement in behavioral performance [5–7]. For instance, in a phonetic training study, Golestani *et al.* showed that individuals with greater behavioral improvement showed more neural changes (i.e. neuroplasticity) in the classical frontal speech area and the temporal-parietal speech area than those with less improvement [6].

Other studies have found that preexisting individual differences in neural activity might serve as neurofunctional predictors of learning [4,8–10]. In visual language learning, for example, Xue *et al.* [4] recently revealed that interindividual variability in brain asymmetry in the fusiform area predicted visual word learning (also see Ref. [8] for related findings). This line of evidence is important because it shows that interindividual variability in neural patterns is not merely a result of differential training effectiveness (i.e. neuroplasticity). Instead, preexisting individual differences in neural responses predict and possibly affect subsequent learning outcomes.

The present study was aimed at extending our existing research on visual word learning by identifying potential neurofunctional predictors of auditory word learning. We trained 24 Chinese college students to learn a logographic artificial language (LAL) for 2 weeks. The data on visual word learning and the fusiform gyrus activation have been reported elsewhere [8]. In this article, we mainly focused on auditory word learning. Many neuroimaging studies have examined the neural changes associated with speech training by focusing on the training of one or two aspects of language, such as tones [11], phonetics [6,12], or the association between sounds and meanings [13]. The present study adopted a comprehensive language training paradigm (i.e. simultaneously training the visual form, phonology, and semantics of LAL) so that the learning of LAL resembles language learning in real life. On the basis of the findings of previous studies [3,4,13], we hypothesized that (a) there would exist significant individual differences in neural responses to novel auditory words (pretraining) in the auditory language processing regions, such as the left temporal lobe and the inferior frontal gyri, and (b) these differences would be associated with the efficiency of learning auditory words.

Methods
Participants

Twenty-four native Chinese college students (11 females and 13 males), with normal or corrected-to-normal visual acuity, participated in this training study. They were all strongly right-handed as judged by Snyder and Harris's handedness inventory [14]. None of them had a previous history of neurological or psychiatric disease. All of them learned English as their second language and none had any formal knowledge of Korean language before the training. Informed written consent was obtained from the participants before the experiment. This study was approved by both the Beijing 306 Hospital and the State Key Laboratory of Cognitive Neuroscience and Learning at Beijing Normal University.

Materials and training procedure

Participants went through a 2-week training program (2 h per day and 5 days per week) to learn the visual forms, phonology, and semantics of the LAL characters. The LAL was created by borrowing the writing and sounds of 60 Korean Hangul characters and were assigned arbitrary meanings, half natural kinds (e.g. Sun) and half man-made artifacts (e.g. table). The sounds were recorded from four native Korean speakers (two males and two females), and normalized to the same length (600 ms) and loudness. Participants learned 20 characters per day during the first 3 days of training and all 60 characters for the remaining training days. For more details about the training, please refer to Ref. [8].

In this study, we used two phonological tests: dictation (selecting the correct character after hearing an LAL sound) and listening comprehension (selecting the correct meaning after hearing an LAL sound). During the tests, participants heard an LAL sound through earphones and at the same time saw a pair of characters on the screen which would stay on until participants responded. Participants needed to press one of two keys to select the visual form or meaning that corresponded to the LAL sound. If no responses were made in 4 s after stimulus presentation, the stimulus would disappear. The next stimulus would begin after an interval of 1 s. During the first 3 training days, participants were only tested on the 20 characters they learned that day. Participants were trained each day until they attained a correct ratio of 90% or higher.

Functional MRI paradigm and parameters

Participants were scanned before and after the 2-week training, using the same passive listening task. Passive tasks have often been used in studies of foreign language processing [3,4,8] because they are more likely to reveal reliable individual differences, less likely to involve neural compensation due to task difficulty, and less likely to confound task difficulty with other study variables. A rapid event-related design was used in this study. At the beginning of the scanning session, there was a 9-s fixation to allow for stability in magnetization, and these images were excluded from analyses. Stimuli were programmed with DMDX and were projected onto a translucent screen via a projector. Participants listened to the stimuli through a functional MRI-compatible headphone, and at the same time, a white fixation cross could be seen through a mirror attached to the head coil. Each sound lasted for 600 ms, and the ISI was 1200 ms. A pure tone (600 Hz) which matched LAL's sounds in length and loudness, served as the baseline. Sixty LAL sounds, 60 matched unfamiliar (i.e. not trained) Korean sounds, and the baseline were presented randomly. The total scanning session lasted 6 m 12 s. All participants followed the same sequence for stimulus presentation to avoid confounding stimulus sequence with individual differences in neural responses.

Structural and functional MRI scans were performed on a 2.0 T GE/Elscint Prestige whole-body MRI scanner (Elscint Ltd., Haifa, Israel) with a standard head coil at the MRI Center of the Beijing 306 Hospital. Single-shot T2*-weighted gradient-echo, EPI sequence was used for functional imaging acquisition with the following parameters: TR/TE/θ=3000 ms/60 ms/90°, FOV=375 × 210 mm, matrix= 128 × 72, and slice thickness=6 mm. Nineteen contiguous axial slices parallel to AC-PC line were obtained to cover the whole cerebrum and partial cerebellum. Anatomical MRI was acquired using a T1-weighted, three-dimensional, gradient-echo pulse-sequence. Parameters for this sequence were: TR/TE/θ=25 ms/6 ms/28°, FOV=220 × 220 mm, matrix=220 × 220, and slice thickness=2 mm. Eighty-nine axial slices parallel to the AC-PC line were acquired to provide a high-resolution structural image of the whole brain.

Functional MRI data analysis

Image preprocessing and statistical analyses were performed with the Statistical Parametric Mapping (SPM2, Wellcome Department of Cognitive Neurology, London, UK) implemented in Matlab (Mathworks Inc. Sherborn, Mass., USA). The first three images were excluded from analysis to avoid the initial instability of the magnet. Functional images were realigned, unwarped, normalized to the MNI template, and smoothed with an 8 mm FWHM Gaussian filter. The general linear model was used to estimate the condition effect for individual participants [15]. The canonical haemodynamic response function (HRF) was used to model the BOLD response of each stimulus type. Group effects were computed with a random-effects model. Group-averaged results were computed using one-sample *t*-tests. Clusters with more than 30 voxels activated above the threshold of $P < 0.001$ (uncorrected) were considered as significant. To examine the neural differences between 'good' and 'poor' learners, we submitted the two groups of learners to a permutation test using Statistical nonParametric Mapping (SnPM3, *http://www.sph.umich.edu/ni-stat/ SnPM/*). In this analysis, a slightly less stringent threshold was used, that is, 10 voxels activated above the threshold of $P < 0.005$, uncorrected. All activations were localized according to the MNI coordinates.

Results
Behavioral results

Figure 1a shows the mean reaction times (RTs) for the listening comprehension and dictation tasks. One-way repeated measures analysis of variance (ANOVA) indicated that RTs decreased significantly over the training period for listening comprehension [$F(9,207)=65.54$, $P < 0.001$] and dictation [$F(9,207)=73.29$, $P < 0.001$]. Owing to the strict criterion for meeting the target performance in the listening comprehension and dictation tasks during training (i.e. > 90%), correct ratios were generally high. As shown in Fig. 1b, however, training still resulted in increased accuracy in both phonological tasks [$F(9,207)=2.17$, $P < 0.05$; $F(9,207)=21.71$, $P < 0.001$]. To capture reliable individual differences in learning outcomes, we conducted a principal component analysis (PCA) to extract a core phonological component for reaction times of the listening comprehension and dictation tasks during the last five training days. One component, explaining 80.96% of the total variance, was extracted. On the basis of this component, we divided the participants (i.e. median split) into 'good' and 'poor' groups. Each group included 12 participants.

Functional MRI results
Whole-brain analysis

During the pretraining scan, the passive listening task activated a broad sensorimotor network, including bilateral
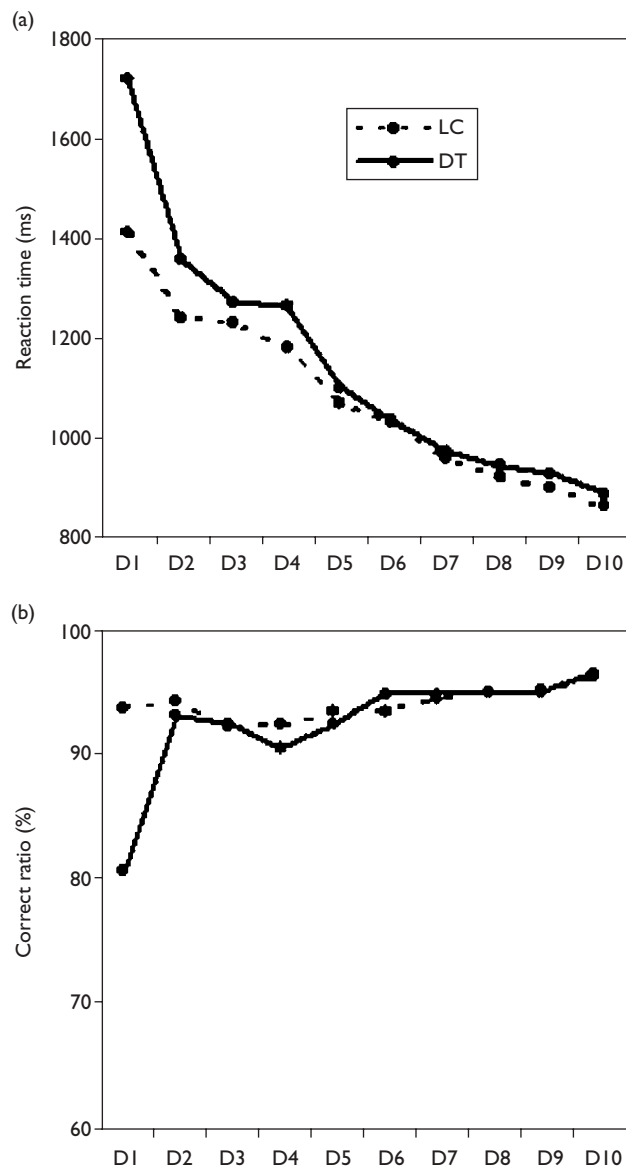
(a)

(b)

Fig. 1 Reaction times (a) and correct ratios (b) for the listening comprehension (LC) and dictation (DT) tasks across training sessions. D=day.

frontal, temporal, and motor areas. After training, activation was more extensive in the frontal lobe and subcortical areas (Fig. 2). Specifically, participants showed increased activation in the left inferior/superior parietal lobe (I/SPL), inferior frontal gyrus (IFG), middle frontal gyrus (MFG), and right supplemental motor area (SMA), cerebellum, and insula/putamen. Decreased activation was found in bilateral temporal lobe.

Comparisons between 'good' and 'poor' learners
To identify the brain regions that contributed to the success of auditory word learning, we compared the pretraining neural responses to auditory words between 'good' and 'poor' learners using a permutation test implemented in SnPM. This test enabled us to directly show the probability of getting our results if the participants were grouped as 'good' and 'poor' learners at random. The results revealed
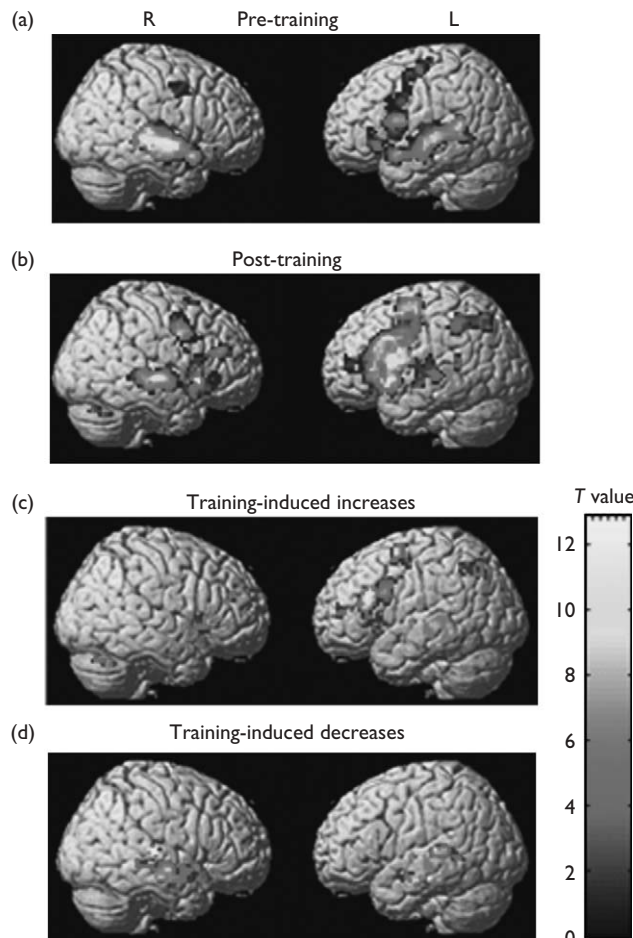
(a)

Fig. 2 Group-averaged brain activation. (a) LAL sounds > tone at pre-training stage. (b) LAL sounds > tone at post-training stage. (c) Post-training > pre-training LAL sounds. (d) Pre-training > post-training LAL sounds. L=left, R=right, LAL=logographic artificial language. The threshold of this comparison was $P < 0.001$, uncorrected.

that 'good' learners showed increased activation in the left posterior middle temporal gyrus (MTG)/the superior temporal sulcus (STS) ($x=-63$, $y=-42$, $z=+6$; $P=0.0012$, uncorrected), whereas 'poor' learners showed increased activation in the right IFG ($x=+54$, $y=+27$, $z=+9$; $P=0.0016$, uncorrected) (Fig. 3a and b).

Correlational analysis
To supplement the group comparisons, we conducted a correlational analysis. First, we quantified the activation (using effect size, that is, beta values in the general linear model) in the two regions that showed significant differences between 'good' and 'poor' learners (i.e. the left pMTG/STS and the right IFG). We then correlated the amount of activation in the two regions with the core phonological component across individuals. These analyses showed that the left pMTG/STS activation at the pretraining stage was positively correlated with the fluency of auditory word processing (i.e. negatively with RT in the core phonological component) ($r=-0.414$, $P<0.05$), whereas the right IFG activation was negatively correlated with word fluency (i.e. positively with RT, $r=0.479$, $P<0.05$) (Fig. 3c and d).
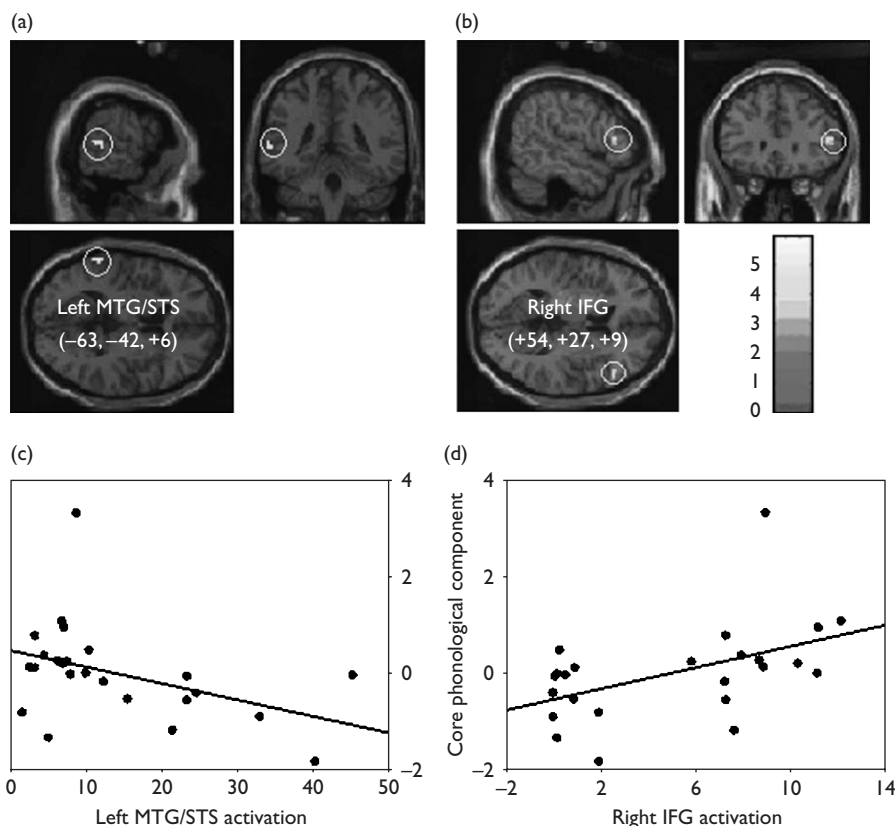
Fig. 3 The top panel shows activation maps of the direct comparison between 'good' and 'poor' learners before training. (a) Good > Poor learners. (b) Poor > Good learners. The threshold of this comparison was $P < 0.005$, uncorrected. The bottom panel shows correlations between pre-training activation (the left pMTG/STS and the right IFG) and the core phonological component.

Discussion

In this study, we aimed to discover preexisting individual differences in neural responses that could predict auditory word learning. Whole-brain analysis showed significant training-induced changes in the left frontal areas, I/SPL, right SMA, insula/putamen, cerebellum, and bilateral temporal lobe. On the basis of the findings of previous research, these neural changes might represent improvement in the following aspects of the processing of the newly acquired speech sounds: processing efficiency (i.e. decreased activation in bilateral temporal regions) [9], access to phonological and semantic representations (i.e. increased activation in the left frontal areas) [16], storage of phonological representations (i.e. increased activation in the right I/SPL) [5], and demand of articulation preparation or speech production control (i.e. increased activation in the right insula/putamen and cerebellum) [7,12].

Consistent with previous studies [3,4,8], this study observed significant individual differences in neural responses (especially in the language network such as the left pMTG/STS) to novel auditory words at the pretraining stage. More importantly, we found that pretraining individual differences in activation in the left pMTG/STS and the right IFG predicted learning outcomes. Previous studies have shown that the left pMTG/STS was involved in the analysis of complex sounds, including speech sounds [17,18], nonspeech vocalizations [19], and other familiar environmental sounds [20]. This analysis might represent an intermediate stage of speech processing in the functional

pathway linking dorsal STG and STS (which are engaged in the analysis of physical features of complex sounds) to more ventral regions in the left MTG and anterior STS (which are engaged in higher-level linguistic processes such as semantic and syntactic processes) [18]. Two other studies also found that the left posterior MTG/STS played an important role in perceiving sounds as speech [21,22]. In light of these findings and the fact that we used a passive listening task in this study, we think that the individual differences in the left pMTG/STS might primarily reflect individual differences in speech perception between 'good' and 'poor' learners.

Unlike the left pMTG/STS, the right IFG is an area known to be involved in nonlinguistic processing [13,23]. Our result of a negative association between activation in the right IFG and auditory word learning seems to suggest that, if the nonlinguistic network (the right IFG) is activated by the novel language, participants may not be as efficient in learning the language. This result parallels a finding by Xue et al. [4] that visual word learning was not as efficient when participants did not rely on the visual word fusiform area (VWFA) to process novel words.

In sum, the results of our two previous studies [4,8] and this study clearly show that preexisting individual differences in neural responses are important neurofunctional predictors of language learning. The specific mechanisms (or reasons) for this prediction are still unclear. One possible explanation is that the native language tunes relevant brain regions and such tuning affects the learning of a second language. For example, years of language experience tune the

left mid-fusiform to be very efficient in processing visual words which can be transferred to the processing of a novel writing [4,8]. Similarly, the left pMTG/STS is efficient in processing new speech because it has been tuned by native speech. This perspective differs from that of some researchers who have argued for a separate language system for second language [24], but is consistent with newer findings that second language shares the same neural representation with the native language, even when the second language is dramatically different from the native language and at a very low level of proficiency [25]. Further studies are needed to examine the origin of preexisting individual differences in neural responses and the mechanisms underlying the effects of these responses on subsequent language learning.

Conclusion

Using an artificial language training paradigm, this study found significant functional differences in the left pMTG/STS and right IFG between 'good' and 'poor' learners of auditory words. These results provided strong evidence for our hypothesis that preexisting interindividual variability in neural activities can predict the efficiency of word learning.

Acknowledgements

References

1. Tzourio-Mazoyer N, Josse G, Crivello F, Mazoyer B. Interindividual variability in the hemispheric organization for speech. *NeuroImage* 2004; **21**:422–435.
2. Xiong J, Rao S, Gao JH, Woldorff M, Fox PT. Evaluation of hemispheric dominance for language using functional MRI: a comparison with positron emission tomography. *Hum Brain Mapp* 1998; **6**:42–58.
3. Dehaene S, Dupoux E, Mehler J, Cohen L, Paulesu E, Perani D, et al. Anatomical variability in the cortical representation of first and second language. *NeuroReport* 1997; **8**:3809–3815.
4. Xue G, Chen C, Jin Z, Dong Q. Cerebral asymmetry in the fusiform areas predicted the efficiency of learning a new writing system. *J Cogn Neurosci* 2006; **18**:923–931.
5. Breitenstein C, Jansen A, Deppe M, Foerster AF, Sommer J, Wolbers T, et al. Hippocampus activity differentiates good from poor learners of a novel lexicon. *NeuroImage* 2005; **25**:958–968.
6. Golestani N, Zatorre RJ. Learning new sounds of speech: reallocation of neural substrates. *NeuroImage* 2004; **21**:494–506.
7. Raboyeau G, Marie N, Balduyck S, Gros H, De'monet J, Cardebat D. Lexical learning of the English language: a PET study in healthy French subjects. *NeuroImage* 2004; **22**:1808–1818.
8. Chen C, Xue G, Dong Q, Jin Z, Li T, Xue F, et al. Sex determines the neurofunctional predictors of visual word learning. *Neuropsychologia* 2007; **45**:741–747.
9. Kelly AMC, Garavan H. Human functional neuroimaging of brain changes associated with practice. *Cereb Cortex* 2005; **15**:1089–1102.
10. Hoeft F, Ueno T, Reiss AL, Meyler A, Whitfield-Gabrielie S, Glover GH, et al. Prediction of children's reading skills using behavioral, functional, and structural neuroimaging measures. *Beh Neurosci* 2007; **121**:602–613.
11. Wang Y, Sereno JA, Jongman A, Hirsch J. fMRI evidence for cortical modification during learning of Mandarin lexicon tone. *J Cogn Neurosci* 2003; **15**:1019–1027.
12. Callan DE, Tajima K, Callan AM, Kubo R, Masaki S, Akahane-Yamada R. Learning-induced neural plasticity associated with improved identification performance after training of a difficult second-language phonetic contrast. *NeuroImage* 2003; **19**:113–124.
13. Wong PCM, Perrachione TK, Parrish TB. Neural characteristics of successful and less successful speech and word learning in adults. *Hum Brain Mapp* 2007; **28**:995–1006.
14. Snyder PJ, Harris LJ. Handedness, sex, and familial sinistrality effects on spatial tasks. *Cortex* 1993; **29**:115–134.
15. Friston KJ, Frith CD, Frackowiak RS, Turner R. Characterizing dynamic brain responses with fMRI: a multivariate approach. *NeuroImage* 1995; **2**:166–172.
16. Poldrack RA, Wagner AD, Prull MW, Desmond JE, Glover GH, Gabrieli JD. Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *NeuroImage* 1999; **10**:15–35.
17. Binder JR, Frost JA, Hammeke TA, Bellgowan PSF, Springer JA, Kaufman JN, et al. Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex* 2000; **10**:512–528.
18. Liebenthal E, Binder JR, Spitzer SM, Possing ET, Medler DA. Neural substrates of phonemic perception. *Cereb Cortex* 2005; **15**:1621–1631.
19. Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. Voice-selective areas in human auditory cortex. *Nature* 2000; **403**:309–312.
20. Giraud AL, Price CJ. The constraints functional neuroimaging places on classical models of auditory word processing. *J Cogn Neurosci* 2001; **13**:754–765.
21. Dehaene-Lambertz G, Pallier C, Semiclaes W, Sprenger-Charolles L, Jobert A, Dehaene S. Neural correlates of switching from auditory to speech perception. *NeuroImage* 2005; **24**:21–33.
22. Möttönen R, Calvert GA, Jääskeläinen IP, Matthews PM, Thesen T, Tuomainen J, et al. Perceiving identical sounds as speech or non-speech modulates activity in the left posterior superior temporal sulcus. *NeuroImage* 2006; **30**:563–569.
23. Klein D, Zatorre RJ, Milner B, Zhao V. A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. *NeuroImage* 2001; **13**:636–653.
24. Kim KH, Relkin NR, Lee KM, Hirsch J. Distinct cortical areas associated with native and second languages. *Nature* 1997; **388**:171–174.
25. Xue G, Jin Z, Dong Q, Chen C. Mapping of verbal working memory in nonfluent Chinese–English bilinguals with functional MRI. *NeuroImage* 2004; **22**:1–10.