Neuroscience 162 (2009) 87–95

EARLY CORTICAL PROCESSING OF LINGUISTIC PITCH PATTERNS AS REVEALED BY THE MISMATCH NEGATIVITY

G.-Q. REN, Y. YANG* AND X. LI

State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, Datun Road 4A, Chaoyang District, Beijing 100101, PR China

Abstract—Previous brain imaging studies have shown the left hemispheric dominance for processing of lexical tone in native speakers. However, the low temporal resolution related to neuroimaging techniques might not explicitly detect the brain activities that occur at a relatively small or a determined time frame. We used the mismatch negativity (MMN) and a source estimation technique (low-resolution electromagnetic tomography [LORETA]) to probe the brain activities underlying the early pre-attentive processing of Mandarin lexical tone and intonation. A passive oddball paradigm was applied to present tone and intonation contrast in a speech and nonspeech context. The results showed that no difference of the MMN amplitudes existed between speech and nonspeech conditions, although a larger MMN was found for tone than intonation condition. Source localization of the MMNs for all of the conditions showed the right hemispheric dominance, regardless of their linguistic functions (tone vs. intonation) or speech context (speech vs. nonspeech). Interestingly, the MMN generator for normal tone and hummed tone originated from the same cortical area (right parietal lobe, BA 19). These findings suggest that the pre-attentive cortical processing can be modulated not only by speech stimuli, but also by their nonspeech hums. Our data are compatible with the acoustic hypothesis of speech processing. Crown Copyright © 2009 Published by Elsevier Ltd on behalf of IBRO. All rights reserved.

Key words: auditory, pitch, mandarin Chinese, lexical tone, intonation, mismatch negativity.

Linguistic pitch patterns are used to signal different aspects of spoken language such as emphatic stress and word meaning. An important issue that is particularly relevant to the processing of linguistic pitch patterns concerns language lateralization in the human brain. Two competing hypotheses were proposed to state the neural mechanisms underlying human pitch perception. The functional hypothesis (task-dependent hypothesis) claims that the hemispheric dominance of pitch patterns perception is determined by their psychological functions (Van Lancker, 1980). Those pitch patterns that carry a greater linguistic load (e.g. lexical tone) are preferentially processed in the left hemisphere, while those that carry a less linguistic load (e.g. intonation) are preferentially processed in the right hemisphere (Van Lancker, 1980; Wong, 2002). Alternatively, the acoustic hypothesis (cue-dependent hypothesis) claims that, regardless of psychological functions, all pitch patterns are lateralized to the same hemisphere, the right hemisphere in particular (Klouda et al., 1988; Zatorre and Belin, 2001; Zatorre et al., 2002). Empiric evidence exists either for the functional hypothesis (Hsieh et al., 2001; Gandour et al., 2000, 2002; Wong et al., 2004) or for the acoustic hypothesis (Zatorre and Belin, 2001; Luo et al., 2006; Warrier and Zatorre, 2004). Up to the present, the nature of these neural mechanisms underlying hemispheric lateralization for the perception of linguistic pitch patterns still remains a matter of debate.

The evidence for the functional hypothesis usually came from dichotic listening (Wang et al., 2001) or imaging studies (Hsieh et al., 2001; Klein et al., 2001; Gandour et al., 2000, 2002; Wong et al., 2004). These studies revealed the left hemispheric superiority in lexical tone perception for native speakers of tone languages, and suggested that hemispheric lateralization seems more sensitive to language-specific factors irrespective of the low-level acoustic processing. For example, when Thai and Chinese subjects were required to perform discrimination judgments of Thai tone, only Thai subjects displayed an increased activation in the left inferior prefrontal cortex (Gandour et al., 2002). Similar lateralization was obtained in Chinese speakers when Chinese and English speakers were required to discriminate the pitch patterns in Chinese words (Klein et al., 2001). In an functional magnetic resonance imaging (fMRI) study, Gandour et al. (2003) demonstrated that pitch contours associated with Mandarin lexical tones are processed in the left hemisphere, whereas pitch contours associated with intonation are processed mainly in the right hemisphere by Chinese speakers. These findings were against the view that hemispheric lateralization is sensitive to low-level auditory processing in the perception of linguistic pitch patterns.

However, because of the low temporal resolution of dichotic or imaging measures, the results mentioned above may reflect temporally aggregated neural events. In a passive oddball paradigm, Luo et al. (2006) demonstrated the right hemispheric lateralization for early auditory processing of lexical tones. The early pre-attentive cortical processing was found to be sensitive not only to speech sound, but also to nonspeech sound which is of phonological relevance in a particular language (Tervaniemi et al., 2006). In fact, these dichotomous views need not be mutually exclusive (Zatorre and Gandour, 2008). Both linguistic and acoustic factors are all necessary for developing a neural model of speech perception, and this

---

*Corresponding author. Tel: +86-10-64888629; fax: +86-10-64872070.
E-mail address: yangyf@psych.ac.cn (Y. Yang).

Abbreviations: ANOVA, analysis of variance; ERP, event-related potential; fMRI, functional magnetic resonance imaging; F0, fundamental frequency; LORETA, low-resolution electromagnetic tomography; MMN, mismatch negativity.

0306-4522/09 $ - see front matter. Crown Copyright © 2009 Published by Elsevier Ltd on behalf of IBRO. All rights reserved.
doi:10.1016/j.neuroscience.2009.04.021

model relies on dynamic interactions between the two hemispheres (Gandour et al., 2004). Moreover, subsequent stages of linguistic processing might have developed from these early low-level acoustic processing (Zatorre et al., 2002; Luo et al., 2006).

This study focused on the early auditory processing of the two types of linguistic pitch patterns, Mandarin lexical tone and intonation. Mandarin lexical tone and intonation can be taken as stimuli for a robust test of these two hypotheses. In tonal languages such as Chinese, lexical tone is signaled by pitch variations and associated with spectral processing. Moreover, lexical tone is lexically contrastive and can distinguish lexical meaning, just as phonemes are. Mandarin Chinese has four lexical tones and tones 1–4 can be described phonetically as high level, high rising, low rising, and high falling pitch patterns respectively. For example, the syllable/ma/in Mandarin Chinese pronounced in a high level pitch means "mother," but the same syllable means "horse" when pronounced in a low rising pitch. Intonation, which is also signaled by pitch variation, does not distinguish lexical meanings. Intonation may convey several types of meanings such as attitudinal meanings, discursive meanings, or grammatical meanings which suggest that there are typical intonations associated with syntactic structures like declaratives, interrogatives, and imperatives (Cruttenden, 1997). Mandarin tone and intonation thus have equal acoustic features and different linguistic functions. The functional hypothesis predicts that the perception of lexical tone would be lateralized to the left hemisphere and the perception of intonation to the right hemisphere. On the contrary, the acoustic hypothesis predicts that both the processing of lexical tone and intonation would be lateralized to the right hemisphere regardless of their linguistic functions.

In the present study, we used event-related potentials and a source estimation technique low-resolution electromagnetic tomography (LORETA) to test which hypothesis prevails at an early pre-attentive cortical stage of linguistic pitch patterns processing. The event-related potential (ERP) component of interest is the mismatch negativity (MMN), which is an event-related response elicited by infrequent auditory stimulus (deviant) occurring among frequently repeated sounds (standard). The MMN has been proved to be an excellent tool for investigating the automatic detection of auditory changes (Näätänen and Escera, 2000; Näätänen, 2001; Näätänen et al., 2001, 2007; Pulvermüller and Shtyrov, 2006). Furthermore, LORETA was used to estimate the sources of the MMN. The LORETA approach has recently been successfully used in the studies on auditory processing (Meyer et al., 2006; Gottselig et al., 2004; Laufer and Pratt, 2005; Marco-Pallarés et al., 2005), and proved to be valuable in assessing the neural mechanisms underlying auditory processing.

A passive MMN paradigm was applied to present normal lexical tone contrast, normal intonation contrast, and their corresponding hummed versions. The aim of using hummed versions was to eliminate consonant and vowel information while preserving suprasegmental information of the normal speech stimuli. That is, compared to the normal versions, the hummed versions only preserved the purely acoustic pattern of speech melody. Our focus here was to compare the brain responses to lexical tone contrast and intonation contrast in speech and nonspeech context. By comparing the normal version and their hummed counterparts, we are able to determine whether the acoustic features of linguistic pitch patterns play an important role, and whether the modulation of ERPs to linguistic pitch patterns is speech-specific or not in the early pre-attentive processing of speech. If the functional hypothesis prevails, we would observe the left hemisphere lateralization for the lexical tone perception in speech context, and the right hemisphere lateralization for the lexical tone perception in nonspeech context and the intonation perception in both speech and nonspeech contexts. Otherwise, if the acoustic hypothesis prevails, the right hemisphere dominance for the perception of all the stimuli would be obtained.

EXPERIMENTAL PROCEDURES

Participants

Twelve graduate students (age range 21–25; six male, six female) participated in this study as paid volunteers. All the participants were native Mandarin Chinese speakers and right-handed, with no history of neurological or psychiatric impairment. Informed consent was obtained from all the participants.

Stimuli

Four experimental conditions were defined by linguistic function and context. Stimuli were presented in four oddball conditions (see Fig. 1). *The normal tone condition* consisted of two Mandarin syllables that have the same vowel and consonant /gai/ but different lexical tone (tone 3 and tone 4). Both were pronounced in a declarative intonation, syllable /gai3/ was frequently presented as the standard stimulus and syllable /gai4/ was infrequently presented as the deviant stimulus. The *hummed tone condition* was the hummed version of normal tone condition. The *normal intonation condition* consisted of the syllable /gai4/ pronounced in a declarative intonation and an interrogative intonation respectively. The declarative one was presented as the standard stimulus and the interrogative one was presented as the deviant stimulus. The *hummed intonation condition* was the hummed version of normal intonation condition.

Table 1 presents the acoustic characteristics of the experimental stimuli. The syllables used in speech conditions were pronounced by a well-trained male speaker and digitized at a sampling rate of 22,050 Hz. The hummed stimuli were created by resynthesizing the speech stimuli with Praat software (Boersma and Weenink, 2004, from http://www.praat.org) to eliminate the segmental information (vowel and consonant). The hums only conveyed the pitch changes and the suprasegmental information such as the fundamental frequency (F0), duration, and intensity of the normal speech stimuli. All the stimuli were normalized to 70 dB in intensity and 450 ms in duration, including 5 ms rising and falling times.

Concerning the acoustic manifestation of declarative and interrogative intonation in non-tonal languages, F0 contours of interrogative intonations are typically associated with final rises compared to declarative intonations. However, in tonal languages such as Chinese, the interrogative intonation involves not only local F0 variations but also more global patterns (Xu, 2005). For example, in the case of the Mandarin falling tone (as showed in Fig. 1), F0 drops even in a question but the overall height of the F0 is rising. The reasons are likely that in a tone language the local pitch targets are not easily changed, for they encode lexical information (Liu and Xu, 2005).

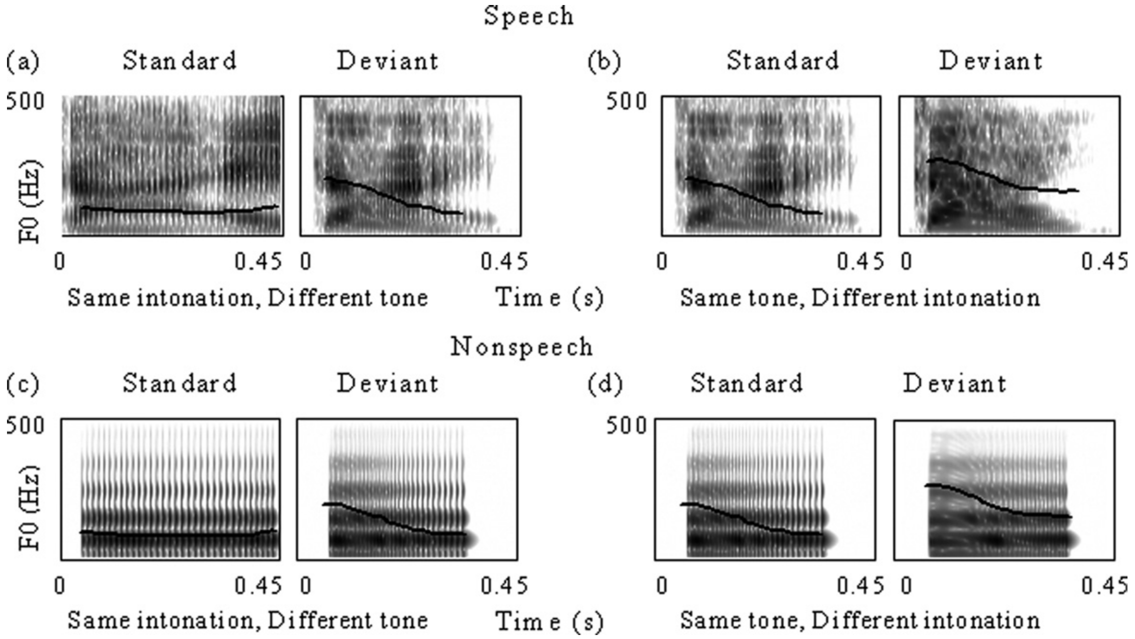

Fig. 1. Acoustic features of stimuli for Chinese speech (top) and hums (bottom). The data set consists of spectrograms with voice F0 contours (F0: 0–500 Hz) superimposed as a black line. For each stimulus pair, the left one illustrates the standard, and the right one illustrates the deviant. The four stimulus pairs consist of (a) same intonation/different lexical tone/gai3/vs./gai4/, (b) same lexical tone/different intonation /gai4/ vs. /gai4?/, (c) the hums of the stimuli in (a), and (d) the hums of the stimuli in (b).

Procedure

A passive oddball paradigm was used to record the responses to the deviant (probability of 10%) and the standard (probability of 90%) stimuli in four blocks. Each block of trials only included one type of deviant so that the entire experiment comprised four blocks of trials. Each block comprised 1015 stimuli and the first 15 stimuli were all standard. Within the blocks, the order of presentation of stimuli was pseudorandom with the restriction that each deviant be separated by at least two standard stimuli. Subjects were instructed to watch a silent movie and ignore the sounds from the headphone during the course of experiment. Stimuli were presented binaurally through headphones in a soundproof room with a stimulus onset-to-onset interval of 650 ms. The blocks were presented in a randomized order, counterbalanced across the subjects.

The electroencephalogram (EEG) was recorded using the 64 electrodes secured in an elastic cap (Neuroscan Inc.) with a sampling rate of 500 Hz, and a band-pass from 0.05 to 40 Hz. The bilateral mastoids serve as the reference and the GND electrode on the cap serve as the ground. The vertical and horizontal electrooculograms were monitored by electrodes placed at the outer canthus of each eye and the electrodes above and below the left eye respectively. All impedances were kept below 5 kΩ.

ERP analysis

Data were further filtered off-line with a 30 Hz low pass filter. Epochs were 600 ms including a 100 ms pre-stimulus baseline. Trials with artifacts exceeding ±50 μV in any channel were excluded from the averaging. The MMN in each condition was obtained by subtracting the responses to the standard from the responses to the deviant. The MMN window was defined between 120 and 240 ms after onset of stimulus in the difference waveforms. The MMN mean amplitudes were calculated by averaging the responses within a 40 ms time-window centered at the peak latency from the electrode Fz, since the largest response was observed at Fz in the grand average waveforms.

Presence of the MMN was determined using a four-way analysis of variance for repeated measures (ANOVA) on the mean

Table 1. Acoustic characteristics of experimental stimuli

Stimuli	Formant (mean values)					F0 parameters			
	F1	F2	F3	F4	F5	Onset	Offset	Range	Average
gai3	732	1871	2616	3531	4220	95	99	75–99	83
gai4	718	1893	2594	3877	4562	203	83	76–203	131
gai4?	741	2017	2819	3855	4132	262	155	154–266	201
hgai3	596	1382	2371	3464	3762				
hgai4	613	1485	2537	3627	3776				
hgai4?	604	1419	2479	3571	3892				

The values are expressed in Hz. The stimuli /gai3/, /gai4/, and /gai4?/ represent speech stimuli. The first two stimuli are Mandarin syllables (gai3, gai4) that were pronounced in a declarative intonation and the third one is a Mandarin syllable (gai4) pronounced in an interrogative intonation. The stimuli /hgai3/, /hgai4/, and /hgai4?/ represent non-speech hums of the speech stimuli. The F0 values of the speech stimuli are same as that of their non-speech counterparts.

amplitudes of the original potentials with linguistic function (tone, intonation), context (speech, nonspeech), type (standard, deviant), and electrode sites (F3, F4, Fz, C3, C4, Cz, P3, P4, Pz) as within-subject factors. To compare the MMN elicited in each condition, a four-way ANOVA on the mean amplitudes and peak latencies of difference potentials were conducted with linguistic function (tone, intonation), context (speech, nonspeech), lobe (frontal, central, parietal), and hemisphere (left, right) as within-subject factors. The Greenhouse-Geisser adjustment was applied when the variance sphericity assumption was not satisfied.

Source analysis

LORETA was used to estimate the sources of the MMNs in response to the four types of stimuli. LORETA is a tomographic technique that can help find the best possible solution of all possible solutions consistent with the scalp distribution (Pascual-Marqui et al., 1994). The solution space of LORETA consists of 2394 pixels with 7 mm resolution. The LORETA-KEY (Pascual-Marqui et al., 2002, http://www.unizh.ch/keyinst/NewLORETA/LORETA01.htm) was used in the analysis and the spatial reference used for this procedure is the Talairach brain atlas (Talairach and Tournoux, 1988). Although this is a coarse analysis compared with fMRI, it is informative when used together with the temporal information provided by ERP.

The LORETA solutions were computed for each condition and on each time point covered the MMN component. The input for LORETA was the grand averaged ERP, sampled over the MMN time points from 120 to 240 ms. The outputs were 3D maps of activity value for each of 2394 cortex pixels, based on the scalp distribution of each time point, with a subtraction of the averaged scalp distribution during the 100 ms prior to stimulus onset which corresponding to the baseline. Those pixels among the top 5% in activation value of each 3D map were treated as "active" pixels to allow focusing on a reduced set of highly activated brain regions and shown in red in LORETA maps (Fig. 4).

RESULTS

Fig. 2 shows the grand average ERP waveforms elicited by the standards and the deviants for each condition. A four-way repeated measures ANOVA conducted on the mean amplitudes revealed main effects of linguistic function [$F(1,11)=8.303$, $P=0.015$], context [$F(1,11)=6.590$, $P=0.026$], electrode site [$F(8,88)=19.577$, $P=0.000$], type [$F(1,11)=53.794$, $P=0.000$], and interaction between linguistic function and type [$F(1,11)=10.303$, $P=0.008$]. The averaged mean amplitudes showed that the deviant evoked a larger negative deflection than the standard in both tone [$F(1,11)=43.87$, $P=0.000$] and intonation conditions [$F(1,11)=27.87$, $P=0.000$]. No significant interaction was found between linguistic function and context [$F(1,11)=1.060$, $P=0.325$]. In addition, interaction between type and electrode site [$F(8,88)=21.295$, $P=0.000$] reached significance. Further tests indicated that the mean amplitudes were more negative for the deviant (-1.194 μV) than the standard (0.031 μV) stimuli at F3, Fz, F4, C3, Cz, and C4 electrode sites. No significant effects were observed on latency.

Fig. 3 shows the grand average waveforms of the mismatch negativities (MMNs) for each condition. A four-way repeated measures ANOVA conducted on the MMN mean amplitudes revealed a main effect of linguistic function [$F(1,11)=21.919$, $P=0.015$], indicating that the MMN was larger for tone than intonation condition. In addition, a significant main effect of lobe was observed [$F(2,22)=34.207$, $P=0.000$]. Bonferroni multiple comparisons ($P=0.05$) for the main effect of lobe showed that the differences of the MMN amplitudes at frontal and central were not significant [M=0.016, SE=0.127, $P=1.000$], but a larger MMN existed at frontal and central than parietal sites [M=-1.535, SE=0.295, $P=0.001$; M=-1.550, SE=0.190, $P=0.000$, respectively]. Interestingly, neither the main effect of context [$F(1,11)=0.303$, $P=0.587$] nor the interaction between linguistic function and context [$F(1,11)=1.398$, $P=0.262$] reached significance. There were no significant three- or four-way interaction effects between function, context, lobe, and hemisphere. No significant effect was observed for the MMN latency.

Fig. 4 shows the highest activation brain areas of the MMNs across experimental conditions. The local maximum of the normal tone condition was located in the right parietal lobe (precuneus, BA 19, Talairach coordinates of the maximum: x=32; y=-74; z=43). Interestingly, the maximum of the hummed tone condition was located in the same area as in the normal tone condition. The local maximum of the normal intonation condition was located in the right frontal lobe (BA 8, x=25; y=38; z=43), whereas in the hummed intonation condition, the maximum was located in the frontal lobe (BA 8, x=25; y=38; z=43) and the parietal lobe (BA 19, x=32; y=-74; z=43).

DISCUSSION

The present study investigated the neural mechanisms underlying the early pre-attentive processing of linguistic pitch patterns by combining the data of ERP and LORETA. The results showed that no difference of the MMN amplitudes existed between speech and nonspeech conditions, although a larger MMN was found for tone than intonation condition. Source estimation of the MMNs indicated that the brain areas underlying lexical tone and intonation processing were lateralized to the right hemisphere, regardless of whether the segmental information of the speech stimuli was preserved or deleted. Interestingly, the source of the MMN elicited by the normal tone and by the hummed tone was located in the same cortical area.

One major finding in this study showed the right hemispheric dominance for the early pre-attentive processing of lexical tone and intonation. Given the equal acoustic features and different linguistic functions of Mandarin tone and intonation, our finding suggests that the acoustic features of linguistic pitch patterns are crucial in determining the neural mechanisms underlying the early pre-attentive speech processing. This observation is compatible with the findings that the right hemisphere plays an important role in complex sound analysis (Tong et al., 2005; Hyde et al., 2008; Warrier and Zatorre, 2004; Zatorre et al., 1994; Zatorre and Belin, 2001), and consistent with the acoustic hypothesis, which predicts that both the processing of lexical tone and intonation would be lateralized to the right hemisphere regardless of their linguistic functions.

By recording MMN under a passive odd-ball paradigm, Luo et al. (2006) also observed the right hemispheric dom-

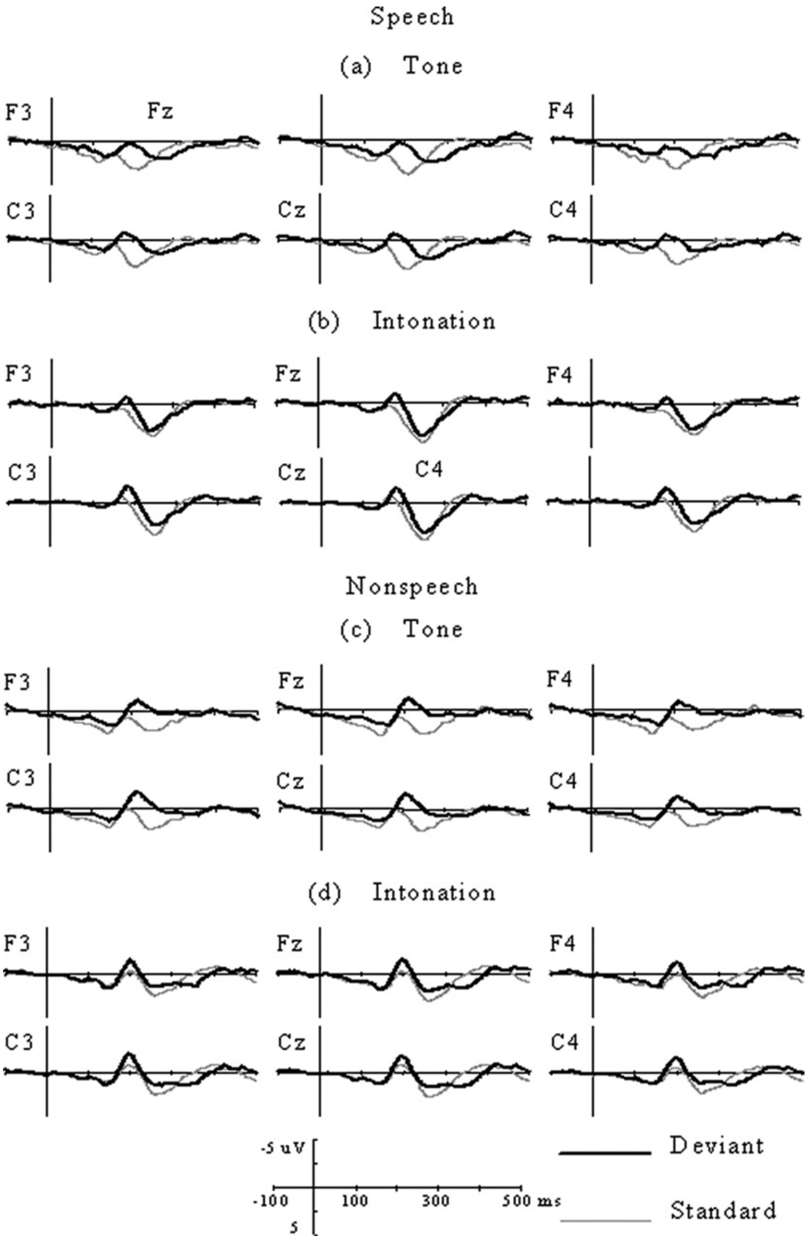

Fig. 2. Grand average waveforms elicited by the deviant (P=10%) and the standard (P=90%) for tone and intonation in speech (a, b) and nonspeech context (c, d).

inance for the perception of lexical tone at an early stage of processing. On the other hand, the data from imaging studies on speech perception revealed the left hemispheric dominance for lexical tone processing in native speakers, and suggest that hemispheric lateralization is sensitive to linguistic functions of pitch patterns and language experience (Hsieh et al., 2001; Klein et al., 2001; Gandour et al., 2002, 2003). Taken together, these findings may reflect a dynamic pattern of brain responses to linguistic pitch patterns. As proposed by Gandour et al. (2004), both linguistics and acoustics are all necessary ingredients for developing a neurobiological model of speech prosody and this model relies on dynamic interactions between the two hemispheres. High-level linguistic processing might initially have developed from low-level acoustic processing (Zatorre et al., 2002; Luo et al., 2006).

LORETA analysis in this study revealed that the source of the MMN to the normal and hummed tone was located in the same area, the right parietal lobe (BA 19). The source of the MMN to normal intonation was located in the frontal lobe (BA 8), whereas the hummed intonation was located in the frontal lobe (BA 8) and parietal lobe (BA 19). Our results are generally compatible with those seen in previous studies (Molholm et al., 2005; Escera et al., 1998; Opitz et al., 2002; Levänen et al., 1996). The frontal lobe contributes to the MMN generators (Rinne et al., 2000; Escera et al., 1998; Opitz et al., 2002; Doeller et al., 2003; Molholm et al., 2005). The frontal MMN has been inter-

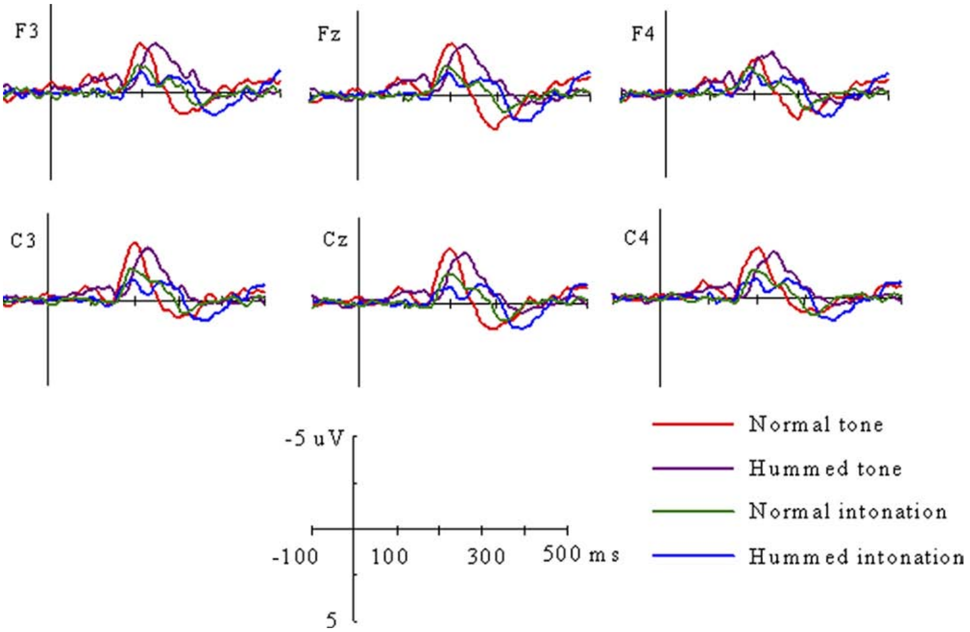

Fig. 3. MMNs displayed for normal tone (red), hummed tone (purple), normal intonation (green), and hummed intonation conditions (blue). For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.

preted as the initiation by the auditory cortex change-detection mechanism and the involvement of involuntary switch of attention towards sound change (Rinne et al., 2000; Escera et al., 1998). There is also some evidence that the parietal lobe contributes to the MMN generators (Kasai et al., 1999; Levänen et al., 1996; Marco-Pallarés et al., 2005). The parietal MMN might reflect the more global auditory change detection (Levänen et al., 1996). It is likely that the activation of frontal and parietal lobe in this study is in part related to an involuntary switch of attention mechanism and may reflect change-detection processes (Molholm et al., 2005).

However, our LORETA data failed to confirm the highest activation of temporal lobe, one major source of the MMN (Luo et al., 2006; Gottselig et al., 2004; Näätänen et al., 2001). The possible reasons may be related to the source estimation method and the time point of the current sources. Firstly, it appears that different source estimation techniques attain different aspects of brain function of the MMN (Rinne et al., 2000). In the present study, the cortical

Fig. 4. Maximum intensity projection images of the grand mean current density (LORETA) averages for the MMNs across conditions (arrowed for easy identification). Each map consists of axial, sagittal, and coronal planes showing the same activation area.

loci were the highest level of activation regions covering the time window of the MMN component, and only one region was reported as the MMN generator in each condition. Secondly, the MMN can be separated into temporally and functionally distinct subcomponents (Maess et al., 2007; Doeller et al., 2003; Opitz et al., 2002). Combining independent component analysis (ICA) and LORETA techniques, Marco-Pallarés et al. (2005) obtained six main independent components in the MMN range and showed that the sources of these components were located in the temporal, frontal, and parietal areas. The frontal MMN generator is activated later than the temporal MMN generator (Rinne et al., 2000). It appears that the change of MMN generators is associated with the time points of the current sources (Youn et al., 2004).

The loci of the MMN to tone and intonation are compatible with the notion that the MMN generator varies depending on the characteristics of the eliciting stimuli (Molholm et al., 2005; Levänen et al., 1996). The fact that the same brain region underlies the processing of normal tone and hummed tone may be due to their equal suprasegmental information. This finding may reflect the important role of acoustic cues on the early pre-attention perception of linguistic pitch patterns. However, the source location of the MMN to normal intonation and hummed intonation was not identical (frontal lobe for intonation, whereas both frontal and parietal lobe for hummed intonation), despite the equal suprasegmental information. The likely reasons for the differences may be that tone and intonation are not completely mutually exclusive in tone language (Cruttenden, 1997). There exist mutual influences between Mandarin tone and intonation (Chao, 1968). Under some circumstances, a tone that was superimposed by an interrogative intonation may sound like the tone itself, especially for its hummed version. And yet the neural mechanism of the interaction between tone and intonation is still a question to be resolved.

Another major finding showed that no difference in the MMN amplitudes existed between speech and nonspeech conditions. This finding is consistent with previous study by Sussman et al. (2004), which demonstrated that the amplitude of MMN did not differentiate between speech and their nonspeech counterparts, regardless of the stimuli were attended or ignored. Our finding suggests that speech perception would be mainly modulated by their acoustic cues at the early pre-attentive stage of processing. In the later stage of processing, however, speech perception would depend on the linguistic functions of stimuli and speech may be processed differently from nonspeech involved with controlled auditory processing (Luo et al., 2006; Sussman et al., 2004). The perception of speech appears to be a dynamic pattern (Gandour et al., 2004), and subsequent stages of speech processing might have developed from the early low-level acoustic processing (Zatorre et al., 2002; Luo et al., 2006).

Compatible with previous studies on early cortical processing of speech (Krishnan et al., 2009; Chandrasekaran et al., 2007b, 2008; Tervaniemi et al., 2006), this finding also indicates that the modification of the early cortical responses to linguistic pitch patterns is not speech-specific. By comparing the MMN of English musicians, English non-musicians, and Mandarin speakers in the perception of pitch contours, Chandrasekaran et al. (2008) found that Mandarin speakers showed larger MMN responses than the other two groups in a between-category contrast (Mandarin tone: high level vs. high rising, T1/T2) and a within-category contrast (acoustic correlates stimuli: a linear rising ramp of high rising vs. high rising, T2L/T2). Tervaniemi et al. (2006) demonstrated larger MMN responses to duration change of nonspeech sound for Finnish (for whom segmental duration can be used to distinguish word meanings) than German speakers (for whom duration cues are not phonemic). However, they found no group differences on the frequency change of nonspeech sound. Interestingly, the modulation of the MMN to frequency changes in nonspeech hums was observed in this study, which may result from the fact that frequency changes are relevant for native speakers of Mandarin.

A larger MMN was observed in this study for tone than intonation condition. Two possible interpretations would account for these differences. One possibility is that the MMN differences resulted from the acoustic differences between the two types of stimuli contrast, since the acoustic characteristics of stimuli may influence the early cortical processing of linguistic pitch contour (Chandrasekaran et al., 2007a; Kraus and Cheour, 2000; Näätänen et al., 2007). Another possibility is that there exists a distinct neural mechanism for the perception of lexical tone and intonation. When subjects were instructed to perform a discrimination task, Mandarin tone and intonation were differentially processed in left and right hemisphere respectively (Gandour et al., 2003). Liang and van Heuven (2004) found that aphasics with damage mainly in the left hemisphere had serious impairment in tone perception while their intonation perception was intact.

CONCLUSION

In conclusion, the current study provides a novel insight into the early pre-attentive cortical processing of linguistic pitch patterns. By using ERP recordings and LORETA analyses, we compared the cortical activities underlying the processing of lexical tone and intonation presented in a speech and nonspeech context. The results showed that the early pre-attentive processing for pitch patterns, regardless of their linguistic functions (tone vs. intonation), is lateralized to the right hemisphere, and can be modulated not only by speech prosody, but also by their corresponding nonspeech hums. These results are compatible with the acoustic hypothesis of speech processing.

Acknowledgments—We thank two anonymous reviewers for valuable comments on earlier versions of this manuscript. This research was supported by China Postdoctoral Science Foundation to the first author, by the National Natural Science Foundation of China (30800296), and Project for Young Scientist Fund, Institute of Psychology, Chinese Academy of Sciences (07CX122012) to the third author.

REFERENCES

Chandrasekaran B, Krishnan A, Gandour JT (2007a) Mismatch negativity to pitch contours is influenced by language experience. Brain Res 1128:148–156.

Chandrasekaran B, Krishnan A, Gandour JT (2007b) Experience-dependent neural plasticity is sensitive to shape of pitch contours. Neuroreport 18(3):1963–1967.

Chandrasekaran B, Krishnan A, Gandour JT (2008) Relative influence of musical and linguistic experience on early cortical processing of pitch contours. Brain Lang doi: 10.1016/j.bandl.2008.02.001.

Chao YR (1968) A grammar of spoken Chinese. Berkeley, CA: University of California Press.

Cruttenden A (1997) Intonation, 2nd ed. Cambridge, UK: Cambridge University Press.

Doeller CF, Opitz B, Mecklinger A, Krick C, Reith W, Schroger E (2003) Prefrontal cortex involvement in preattentive auditory deviance detection: neuroimaging and electrophysiological evidence. Neuroimage 20:1270–1282.

Escera C, Alho K, Winkler I, Näätänen R (1998) Neural mechanisms of involuntary attention to acoustic novelty and change. J Cogn Neurosci 10:590–604.

Gandour J, Dzemidzic M, Wong D, Lowe M, Tong Y, Hsieh L, Satthamnuwong N, Lurito J (2003) Temporal integration of speech prosody is shaped by language experience: an fMRI study. Brain Lang 84:318–336.

Gandour J, Tong Y, Wong D, Talavage T, Dzemidzic M, Xu Y, Li X, Lowe M (2004) Hemispheric roles in the perception of speech prosody. Neuroimage 23:344–357.

Gandour J, Wong D, Hsieh L, Weinzapfel B, Van Lancker D, Hutchins GD (2000) A crosslinguistic PET study of tone perception. J Cogn Neurosci 12(1):207–222.

Gandour J, Wong D, Lowe M, Dzemidzic M, Satthamnuwong N, Tong Y, Li X (2002) A cross-linguistic fMRI study of spectral and temporal cues underlying phonological processing. J Cogn Neurosci 14(7):1076–1087.

Gottselig JM, Brandeis D, Hofer-Tinguely G, Borbély AA, Achermann P (2004) Human central auditory plasticity associated with tone sequence learning. Learn Mem 22:162–171.

Hsieh L, Gandour J, Wong D, Hutchins GD (2001) Functional heterogeneity of inferior frontal gyrus is shaped by linguistic experience. Brain Lang 76:227–252.

Hyde KL, Peretz I, Zatorre RJ (2008) Evidence for the role of the right auditory cortex in fine pitch resolution. Neuropsychologia 46: 632–639.

Kasai K, Nakagome K, Itoh K, Koshida I, Hata A, Iwanami A, Fukuda M, Hiramatsu KI, Kato N (1999) Multiple generators in the auditory discrimination process in humans. Neuroreport 10:2267–2271.

Klein D, Zatorre R, Milner B, Zhao V (2001) A cross-linguistic PET study of tone perception in mandarin Chinese and English speakers. Neuroimage 13(4):646–653.

Klouda GV, Robin DA, Graff-Radford NR, Cooper WE (1988) The role of callosal connections in speech prosody. Brain Lang 35:154–171.

Kraus N, Cheour M (2000) Speech sound representation in the brain. Audiol Neurootol 5:140–150.

Krishnan A, Swaminathan J, Gandour JT (2009) Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to speech context. J Cogn Neurosci 21:1092–1105.

Laufer I, Pratt H (2005) The "F-complex" and MMN tap different aspects of deviance. Clin Neurophysiol 116:336–352.

Levänen S, Ahonen A, Hari R, McEvoy L, Sams M (1996) Deviant auditory stimuli activate human left and right auditory cortex differently. Cereb Cortex 6:288–296.

Liang J, Heuven VJ (2004) Evidence for separate lexical tone and sentence intonation: a perception study of Chinese aphasic patients. Brain Lang 91:60–61.

Liu F, Xu Y (2005) Parallel encoding of focus and interrogative meaning in Mandarin intonation. Phonetica 62:70–87.

Luo H, Ni JT, Li ZH, Li XO, Zhang DR, Zeng FG, Chen L (2006) Opposite patterns of hemisphere dominance for early auditory processing of lexical tones and consonants. Proc Natl Acad Sci U S A 103:19558–19563.

Maess B, Jacobsen T, Schröger E, Friederici AD (2007) Localizing pre-attentive auditory memory-based comparison: magnetic mismatch negativity to pitch change. Neuroimage 37:561–571.

Marco-Pallarés L, Grau C, Ruffini G (2005) Combined ICA-LORETA analysis of mismatch negativity. Neuroimage 25:471–477.

Meyer M, Baumann S, Jancke L (2006) Electrical brain imaging reveals spatiotemporal dynamics of timbre perception in humans. Neuroimage 32:1510–1523.

Molholm S, Martinez A, Ritter W, Javitt DC, Foxe JJ (2005) The neural circuitry of pre-attentive auditory change-detection: an fMRI study of pitch and duration mismatch negativity generators. Cereb Cortex 15:545–551.

Näätänen R (2001) The perception of Speech sounds by the human brain as reflected by the mismatch negativity. MMN and its magnetic equivalent (MMNm). Psychophysiology 38:1–21.

Näätänen R, Escera C (2000) Mismatch negativity: clinical and other applications. Audiol Neurootol 5:105–110.

Näätänen R, Paavilainen P, Rinne T, Alho K (2007) The mismatch negativity (MMN) in basic research of central auditory processing: a review. Clin Neurophysiol 118:2544–2590.

Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I (2001) "Primitive intelligence" in the auditory cortex. Trends Neurosci 24(5):283–288.

Opitz B, Rinne T, Mecklinger A, von Cramon DY, Schroger E (2002) Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. Neuroimage 15:167–174.

Pascual-Marqui RD, Esslen M, Kochi K, Lehmann D (2002) Functional imaging with low resolution brain electromagnetic tomography (LORETA): review, new comparisons, and new validation. Jpn J Clin Neurophysiol 30:81–94.

Pascual-Marqui RD, Michel CM, Lehmann D (1994) Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int J Psychophysiol 18(1):49–65.

Pulvermüller F, Shtyrov Y (2006) Language outside the focus of attention: the mismatch negativity as a tool for studying higher cognitive processes. Prog Neurobiol 79:49–71.

Rinne T, Alho K, Ilmoniemi RJ, Virtanen J, Näätänen R (2000) Separate time behaviors of the temporal and frontal mismatch negativity sources. Neuroimage 12:14–19.

Sussman E, Kujala T, Halmetoja J, Lyytinen H, Alku P, Näätänen R (2004) Automatic and controlled processing of acoustic and phonetic contrasts. Hear Res 190:128–140.

Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain: three-dimensional proportional system. In: (Rayport M, ed). New York: Thieme Medical Publishers.

Tervaniemi M, Jacobsen T, Röttger S, Kujala T, Widmann A, Vainio M, Näätänen R, Schröger E (2006) Selective tuning of cortical sound-feature processing by language experience. Eur J Neurosci 23:2538–2541.

Tong Y, Gandour J, Talavage T, Wong D, Dzemidzic M, Xu Y, Li X, Lowe M (2005) Neural circuitry underlying sentence-level linguistic prosody. Neuroimage 28:417–428.

Van Lancker D (1980) Cerebral lateralization of pitch cues in the linguistic signal. Papers in linguistics. Int J Hum Commun 13(2):201–277.

Wang Y, Jongman A, Sereno JA (2001) Dichotic perception of mandarin tones by Chinese and American listeners. Brain Lang 78:332–348.

Warrier CM, Zatorre RJ (2004) Right temporal cortex is critical for utilization of melodic contextual cues in a pitch constancy task. Brain 127:1616–1625.

Wong PCM (2002) Hemispheric specialization of linguistic pitch patterns. Brain Res Bull 59(2):83–95.

Wong PCM, Parsons LM, Martinez M, Diehl RL (2004) The role of the insular cortex in pitch pattern perception: the effect of linguistic contexts. J Neurosci 24(41):9153–9160.

Xu Y (2005) Speech melody as articulatorily implemented communicative functions. Speech Commun 46:220-151.

Youn T, Park H-J, Kwon JS (2004) Response to Rosburg: a voxel-based statistical parametric mapping of MMN current densities. Hum Brain Mapp 21:46–48.

Zatorre RJ, Belin P (2001) Spectral and temporal processing in human auditory cortex. Cereb Cortex 11:946–953.

Zatorre RJ, Belin P, Penhune VB (2002) Structure and function of auditory cortex: music and speech. Trends Cogn Sci 6(1):37–46.

Zatorre RJ, Evans AC, Meyer E (1994) Neural mechanisms underlying melodic perception and memory for pitch. J Neurosci 14(4):1908–1919.

Zatorre RJ, Gandour JT (2008) Neural specializations for speech and pitch: moving beyond the dichotomies. Philos Trans R Soc B 363:1087–1104.

(Accepted 9 April 2009)
(Available online 15 April 2009)