
Pre-Attentive Processing of Mandarin Tone and Intonation: Evidence from Event-Related Potentials

Gui-Qin Ren, Yi-Yuan Tang, Xiao-Qing Li and
Xue Sui

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/56503>

1. Introduction

In tonal languages (e.g., Chinese and Thai), the meaning of a word cannot be defined solely by consonants and vowels without a lexical tone, which varies in pitch patterns. The pitch patterns associated with Mandarin lexical tones are used to distinguish lexical meaning. 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. The syllable /ma/ in Chinese, for example, can stand for “mother” [tone1], “hemp” [tone2], “horse” [tone3], or “scold” [tone4]. Previous studies have shown that the native speakers of tone languages are highly sensitive to changes in lexical tones regardless of whether the subjects focus their attention on the stimuli or not (Tsang et al., 2011; Ren et al., 2009). The pitch patterns associated with Mandarin intonation, however, may serve a variety of linguistic functions such as attitudinal meanings, discursal meanings, or grammatical meanings (Cruttenden, 1997; Pell, 2006). A cross-linguistic (Chinese and English) study showed that whereas pitch contours associated with intonation are processed predominantly in the right hemisphere whereas the pitch contours associated with tones are processed in the left hemisphere by Chinese listeners only (Gandour et al., 2003).

The neurophysiological study of the processing of tone and intonation can provide valuable insight into the nature of pitch patterns perception. Other than non-tone languages such as English, lexical tone and intonation in Chinese are both signaled primarily by changes in fundamental frequency (F_0) while their linguistic functions are different. Furthermore, most models on Mandarin intonation suggest that lexical tone affects intonation (Chao, 1968; Shen, 1992; Yuan, 2004). Chao (1968) has pointed that intonation is realized by changing pitch range and interacts with lexical tone by addition. In Chinese, interrogative intonation has a higher

pitch contour than that of its declarative counterpart (Yuan, 2004). Among the four Mandarin tones, tone 2 in isolation is rising in its phonological representation (Xu, 2005) and the average F0 contour of tone 2 resembles that of interrogative intonation. The questions arise as to whether the cortical processing of pitch patterns associated with lexical tone is distinguishable from that associated with intonation, and whether the interrogative intonation associated with tone 2 can be identified when native speakers' attention is withdrawn from stimuli input.

The issue of whether or not attention is needed during speech perception has provoked a large amount of researches such as the influence of attention in audiovisual speech perception (Astheimer & Sanders, 2009; Navarra et al., 2010) and role of selective attention in speech perception (Astheimer & Sanders, 2009; 2012). There exists evidences either supports for the view that the audiovisual integration of speech is an attention modulated process, or for the view that audiovisual integration of speech is an automatic process (Navarra et al., 2010; Astheimer & Sanders, 2009; Jones & Munhall, 1997). Concerning the role of attention in speech comprehension, Andersen et al. (2009) demonstrated that temporally selective attention may serve a function that allows preferential processing of highly relevant acoustic information such as word-initial segments during normal speech perception. In subsequently study, Andersen et al. (2012) examined the use of temporally selective attention in 3- to 5-year-old children and found that, like adults, preschool aged children modulate temporally selective attention to preferentially process the initial portions of words in continuous speech. By directly comparing the effects of attention on different speech stimuli, Hugdahl et al. (2003) revealed that attention to speech sounds may act to recruit stronger neuronal activation compared to when the same stimulus is processed in the absence of attention. Although previous results showed that cognitive processing of many aspects of language such as semantic, syntactic, and pitch information take place indexed by MMN regardless of whether subjects focus their attention on linguistic stimuli, the size of the MMN can be modulated by the level of attention (Pulvermüller & Shtyrov, 2003; 2006). The MMN is larger when subjects attend to the stimuli, as compared with that of subjects are involved in a distraction task.

Tone languages are advantageous for examining the nature of pitch patterns processing. In recent years, the functional asymmetry of two human cerebral hemispheres in the processing prosody information has received a considerable attention. The left hemisphere has been thought to be dominant for language-related behaviours (Gandour et al., 2002; Klein, Zatorre, Milner, & Zhao, 2001) and the right hemisphere to be dominant for pitch-related behaviours (Warrier & Zatorre, 2004; Zatorre & Belin, 2001). However, what cues are used by the brain to determine the labor division is still a matter of debate. The functional hypothesis (Pell & Baum, 1997; Wong, 2002) states that the psychological functions of sounds determine which neural mechanisms are engaged during speech processing. Those sounds 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. However, the acoustic hypothesis (Zatorre & Belin, 2001; Zatorre, Belin, & Penhune, 2002) states that all pitch patterns are lateralized to the right hemisphere regardless of psychological functions.

More recently, the dynamic models such as two-stage model and a more comprehensive model have been put forward, which integrate the acoustic hypothesis and functional hypothesis. The two-stage model (Luo et al., 2006) states that speech is initially processed as a general acoustic signal with lateralized to the right hemisphere at a pre-attentive stage, and then mapped into a semantic representation with lateralized to the left hemisphere at an attentive stage. This point of view is compatible with the notion put forward by Zatorre et al (2002) that the left hemisphere lateralization effect in linguistic functions may arise from a slight initial advantage in decoding speech sounds. According to the more comprehensive model proposed by Gandour and his colleagues (Gandour et al., 2004; Tong et al., 2005), speech prosody perception is mediated primarily by the right hemisphere for complex sound analysis while left hemisphere is dominant when language processing is required. What is more, both the left and right hemispheres were found to contribute to pitch patterns perception (Pell, 2006; Xi, Zhang, Shu, Zhang, & Li, 2010). The prosodic speech information can be processed on either hemisphere depending on whether the speech information is emotional or the linguistic prosodic cues (Pell, 2006). The acoustic and linguistic information is processed in parallel at an early stage of speech perception (Xi et al., 2010).

The left hemisphere lateralization in the perception of lexical tones is supported by evidence from a number of studies including dichotic-listening (Wang, Jongman, & Sereno, 2001) and functional imaging studies (Gandour et al., 2002; Klein, Zatorre, Milner, & Zhao, 2001). 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 hemispheric dominance was obtained in Chinese speakers when Chinese and English speakers were required to discriminate the pitch patterns in Chinese words (Klein, Zatorre, Milner, & Zhao, 2001). Nevertheless, those studies mentioned above likely reveal the temporally aggregated brain activity of auditory processing due to the coarse temporal resolution of fMRI or PET.

The specific aims of this study are to further investigate the neural mechanisms underlying the perception of linguistic pitch patterns by comparing the early pre-attentive processing of Mandarin tone and intonation, and examine whether the pitch changes of the intonation associated with Mandarin tone 2 can be detected by native speaker of Chinese at the early pre-attentive stage. Here a method of combining event-related potentials and a source estimation technique Low-resolution electromagnetic tomography (LORETA) was used. The ERP component of interest is the mismatch negativity (MMN), which peaks at about 100-250 ms after stimulus onset and is present by any discriminable changes in auditory processing irrespective of subjects' attention or task (Näätänen & Escera, 2000; Näätänen, Paavilainen, Rinne, & Alho, 2007; Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001). A new MMN paradigm was applied in this study, which allows one to obtain different MMNs in a short time (Näätänen, Pakarinen, Rinne, & Takegata, 2004). The sources of the MMNs were estimated by LORETA, an approach that has been successfully used in the studies on auditory processing to locate the sources of the neural activities (Liu & Perfetti, 2003; Marco-Pallarés, Grau, & Ruffini, 2005).

2. Materials and methods

Subjects:

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

Stimuli and procedure:

Stimuli consisted of two meaningful auditory Chinese words that have the same consonant and vowel (/lai/) but different lexical tone, pronounced in high rising tone (tone 2) and high falling tone (tone 4) respectively. The syllable /lai4/ was pronounced in a declarative intonation, and the syllable /lai2/ was pronounced in a declarative intonation or an interrogative intonation respectively. The standard stimulus was the syllable /lai2/ pronounced in a declarative intonation. Deviant stimuli differed from the standard in either intonation (/lai2/, intonation deviant) or lexical tone (/lai4/, lexical tone deviant).

A new passive auditory odd-ball paradigm (Näätänen, Pakarinen, Rinne, & Takegata, 2004) was applied to present the stimuli. In order to control the effect of physical stimulus features to obtain the relatively pure contribution of the memory network indexed by MMN (Pulvermüller & Shtyrov, 2006; Pulvermüller, Shtyrov, Ilmoniemi, & Marslen-Wilson, 2006), we created three sequences including one oddball sequence and two control sequences to calculate the identity MMN. The oddball sequence preceded by 15 standard was a pseudorandom block of 1015 stimuli which included standard ($P = 0.8$) and two deviants ($P = 0.1$ for each). The two control sequences for each deviant comprised 400 trials respectively and each deviant stimulus was presented alone ($P = 1$). The subjects were instructed to ignore the sounds from the headphone and watch a silent movie during the course of experiment. The order of the presentation of the three sequences was randomized across the subjects.

The auditory stimuli were pronounced in isolation by a trained female speaker and digitized at a sampling rate of 22,050 Hz. The stimuli were modified with Praat software (doing phonetics by computer version 4.4.13, download from www.praat.org) and normalized to 450 ms in duration, including 5 ms rise and fall times. The stimuli were presented binaurally at an intensity of 70 dB through headphones in a soundproof room with a stimulus onset asynchrony of 700 ms. The maximum fundamental frequency between the two deviants was comparable. Fig. 1 shows the acoustic features of the experimental stimuli.

Recording:

The 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 k Ω .

Data analysis:

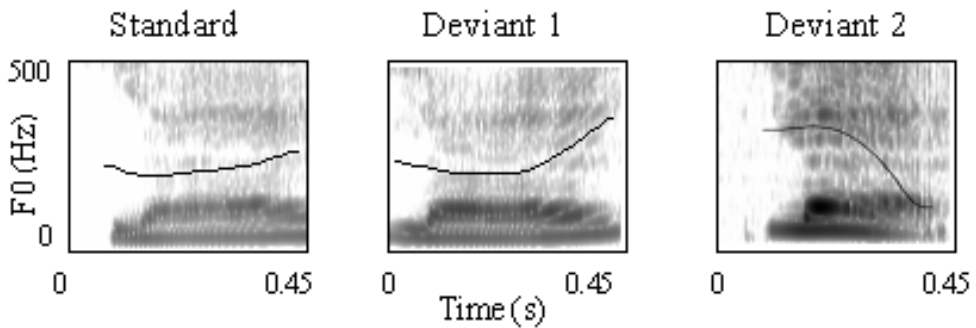


Figure 1. Acoustic features of the stimuli. The data set consists of spectrograms with voice fundamental frequency contours superimposed as a black line. The standard, the pronunciation of syllable /lai2/ in a declarative intonation; Deviant 1, same syllable as the standard but in an interrogative intonation; Deviant 2, syllable /lai4/ in a declarative intonation.

The raw EEG data were first corrected for eye-blink artifacts and filtered with a band-pass filter 0.1-30 Hz. Trials with artifacts exceeding $\pm 75\mu\text{V}$ in any channel were excluded from the averaging. Epochs were 600 ms including a 100 ms pre-stimulus baseline. ERPs elicited by deviants in the oddball sequence and the identical stimulus in the control sequences were averaged separately across subjects and electrodes.

Although the MMN is generally obtained by subtracting the responses to standard from that to deviant stimulus, it is possible that the physical differences between standard and deviant stimuli influence the responses. In order to control the physical stimulus properties in a more stringent manner (Pulvermüller & Shtyrov, 2006; Pulvermüller, Shtyrov, Ilmoniemi, & Marslen-Wilson, 2006), we calculated the MMN (identity MMN) by subtracting from the ERP to a deviant stimulus presented in the oddball sequence, the ERP to the identical stimulus in the control sequence.

The grand ERP wave shapes were first analyzed by visual inspection and the time window of MMN was defined 110-240 ms. The MMN amplitudes were measured as mean voltages using a 40 ms time-window centered at the peak latency from the electrode Cz, since the largest response was observed at Cz in the grand average waveform. Two different analyses of variance (AVOVA) were done on the mean amplitudes. A original ANOVA for the original mean amplitudes was performed to estimate the two MMNs (one for the lexical tone and the other for the intonation) with condition (lexical tone, intonation), type (deviant, the identical stimulus presented alone), and electrode (F3, F4, Fz, C3, C4, Cz, P3, P4, Pz) as independent factors. To compare the MMN elicited by lexical tone condition with the MMN elicited by the intonation condition, a difference ANOVA was conducted for the difference waveforms with condition (lexical tone, intonation), 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.

Low resolution electromagnetic tomography (LORETA) was used to estimate the sources of the MMN elicited in the experiment. LORETA is a tomographic technique that can help find the best possible solution of all possible solutions consistent with the scalp distribution (Pascual-Marqui, Michel, & Lehmann, 1994). The LORETA-KEY (<http://www.unizh.ch/keyinst/NewLORETA/LORETA01.htm>) (Pascual-Marqui, Esslen, Kochi, & Lehmann, 2002) was used and the results are illustrated in Talairach space (Talairach & Tournoux, 1988). We computed the LORETA solutions on each time point covered the MMN. The input for LORETA was the grand averaged ERP, sampled over the MMN window. The outputs were 3D maps of activity value for each of 2,394 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 (Liu & Perfetti, 2003; Ren, Liu, & Han, 2009 a; Ren, Yang, & Li, 2009 b).

3. Results

Fig. 2 shows the grand average waveforms to the deviant stimuli in the oddball sequence ($P=0.1$) and to the identical stimuli in the control sequences ($P=1$). The deviant-minus-control difference waveforms are shown in Fig. 3.

The original three-way [condition \times type \times electrode] ANOVA revealed a significant main effect of condition ($F(1, 12) = 34.132, p = 0.000$) indicating a larger negative-going deflection for lexical tone than intonation (effect magnitude: -1.328 uV), a main effect of type ($F(1, 12) = 15.263, p = 0.002$) due to a larger negative-going deflection for the deviant in the oddball sequence than the identical stimulus in the control sequence (effect magnitude: -1.065 uV), and a main effect of electrode ($F(8, 96) = 9.096, p = 0.001$). In addition, there was a significant interaction effect between condition and type ($F(1, 12) = 7.211, p = 0.02$), and between electrode and type ($F(8, 96) = 6.923, p = 0.002$). Subsequently simple effect analyses showed that the ERPs to the deviant were significantly more negative than to the identical controls only in the lexical tone condition ($F(1, 12) = 17.45, p = 0.001$), and the same differences patterns of types was observed at Fz, F4, Cz, C3, C4, and P4 electrode sites ($F(1, 12) = 7.99, p = 0.015$; $F(1, 12) = 8.09, p = 0.015$; $F(1, 12) = 12.36, p = 0.004$; $F(1, 12) = 7.29, p = 0.019$; $F(1, 12) = 29.52, p = 0.000$; $F(1, 12) = 19.42, p = 0.001$).

A difference three-way [condition \times lobe \times hemisphere] ANOVA revealed a main effect of condition ($F(1, 12) = 23.924, p = 0.000$) indicating that the lexical tone evoked a larger negative deflection than the intonation (effect magnitude: -1.879 uV), a main effect of lobe ($F(2, 24) = 7.677, p = 0.013$) due to the fact that a larger negative deflection existed at frontal and central than parietal sites (effect magnitude: -0.929 uV, -0.893 uV, respectively), and a main effect of hemisphere ($F(1, 12) = 5.691, p = 0.034$) due to the fact that a larger negative deflection existed for right than left hemisphere (effect magnitude: -0.296 uV). In addition, there was a significant interaction between condition and lobe ($F(2, 24) = 8.459, p = 0.002$). Subsequently simple effect analyses showed that a larger negative deflection existed over frontal and central than parietal sites only in the tone condition ($F(2, 24) = 11.71, p = 0.000$). However, the difference ANOVA with peak latency as dependent variable found no significant effect.

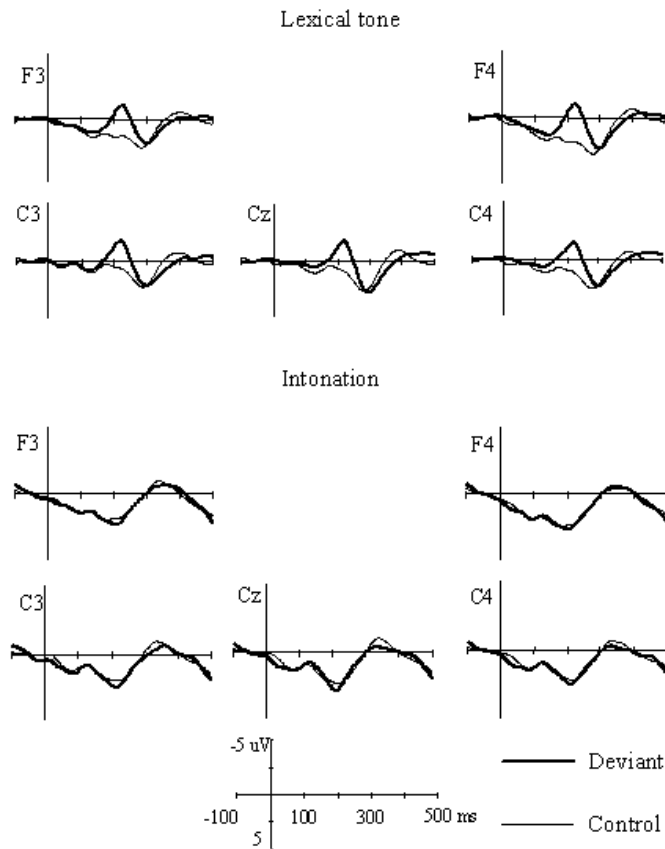


Figure 2. Grand average waveforms to the deviant stimuli in the oddball sequence (thick line) and to the identical stimuli in the control sequence (thin line).

Given the latency and topography of the difference negative deflection (see fig.1 and fig. 2), we classified it as MMN (Rinne et al., 2006). Since no MMN was elicited for the intonation, only the source of the MMN for the lexical tone was analyzed. The local maximum of the MMN was located in the right middle temporal gyrus (BA 21, Talairach coordinates of the maximum: $x = 53$; $y = 3$; $z = -13$).

4. Discussion

The present study examined the early cortical processing of linguistic pitch patterns by comparing the ERP responses to Mandarin tone and intonation. The results demonstrated that MMN was elicited only by the lexical tone contrast and no MMN was obtained to the intonation

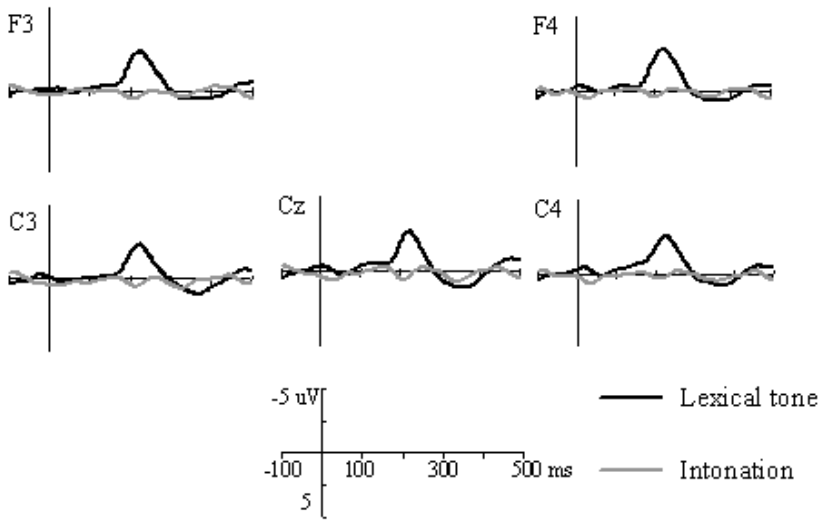


Figure 3. Grand difference waveforms (deviant — control ERPs) in the tone and intonation conditions

contrast which associated with a Mandarin tone 2. Source estimation of the MMN showed that the highest activation of brain areas underlying lexical tone processing was located into the right hemisphere, the right middle temporal gyrus (BA 21).

A clear MMN was observed for the lexical tone contrast and the highest activation of the MMN was located into the right temporal gyrus. The result of right hemisphere dominance for lexical tone in the early pre-attentive processing is converging with previous studies (Luo et al., 2006; Ren et al., 2009 b). Ren et al. (2009 b) demonstrated that both the sources of the MMNs to Mandarin lexical tone and its hummed version were located in the right hemisphere in the early pre-attentive processing. By comparing the early pre-attentive processing of Mandarin tones and consonants, Luo et al (2006) found that Mandarin tones evoked a stronger pre-attentive response in the right hemisphere than in the left hemisphere. Those results above presumably reflect the role of right hemisphere in acoustic processing and compatible with the acoustic hypothesis (Zatorre & Belin, 2001) and the dynamic models (Zatorre et al., 2002; Gandour et al., 2004; Tong et al., 2005 ; Luo et al., 2006), but cannot be explained by the functional hypothesis which predicts lexical tones are preferentially processed in the left hemisphere (Pell & Baum, 1997; Wong, 2002). However, the functional hypothesis was supported by the data from fMRI or PET studies (Hsieh et al., 2001; Klein et al., 2001; Gandour et al., 2002; 2003) which revealed the left hemisphere dominance of native speakers in the perception of lexical tones and suggest that hemispheric lateralization is sensitive to linguistic functions of pitch patterns and language experience. Taken together, these findings seem to reflect the dynamic interaction between the two hemispheres and are compatible with the dynamic models of speech perception (Gandour et al., 2004; Luo et al., 2006; Zatorre, Belin, & Penhune, 2002). Just as proposed by Gandour et al (2004) that both acoustics and linguistics

are all necessary ingredients for developing a neurobiological model of speech prosody. High-level linguistic processing might initially have developed from low-level acoustic processing (Zatorre et al., 2002; Luo et al., 2006).

LORETA analysis for the MMN to lexical tone was located into the right middle temporal gyrus (BA 21), one major source of the MMN (Luo et al., 2006; Näätänen et al., 2001). Besides the temporal lobe, there are other sources contribute to the MMN generators, such as frontal lobe (Molholm, Martinez, Ritter, Javitt, & Foxe, 2005) and parietal lobe (Levänen, Ahonen, Hari, McEvoy, & Sams, 1996; Marco-Pallarés, Grau, & Ruffini, 2005). It appears that the change of MMN generators is associated with the time points of the current sources and is feature dependent (Levänen et al., 1996; Molholm et al., 2005). In this study, the cortical locus reported as the MMN generator was the highest level of activation region covering the time window of the MMN. It can be seen that although the solution of LORETA produces a “blurred-localized” image of a point source, it conserves the location of maximal activity and allows at least the discussion of asymmetric hemispheric involvement in pitch perception (Liu & Perfetti, 2003; Mulert et al., 2007; Ren et al., 2009a; 2009b).

The result that no MMN was elicited by the intonation contrast (declarative vs. interrogative) demonstrated the perceptual difficulties when the intonation is combined with the Mandarin tone 2. The MMN, an index of change-detection of brain response to any change in auditory stimuli, can enable one to determine discrimination accuracy which usually with a good correspondence with behavioural discrimination (Näätänen et al., 2007; 2012). It suggested in the study that the listeners cannot tease apart the two types of intonation at the pre-attentive processing stage. For the four Mandarin tones, the average fundamental frequency contours produced in isolation reflect directly the canonical forms of the tones (Xu, 1997). The F0 contour of tone 2 in isolation is rising in its phonological representation (Xu, 2005; Yuan, 2004) and resembles that of interrogative intonation. When tone 2 is at the end of sentences, it is more difficult for native speakers of Chinese to identify the interrogative intonation (Ren et al., 2011; Yuan, 2004). Three mechanisms were proposed by Yuan (2004) to explain the perception of interrogative intonation, such as the phrase curve mechanism, the strength mechanism, and the tone-dependent mechanism. Among the mechanisms, the strength mechanism may conflict with the tone-dependent mechanism on the Mandarin tone 2. This conflict likely leads to the perceptual difficulties of interrogative intonation for tone 2.

The perceptual difficulties of intonation contrast showed in the present study also suggested the interaction between lexical tone and intonation. Gandour et al. (1997) demonstrated the interaction between lexical tone and intonation in Thai by analyzing intonational characteristics of the Thai sentences which produced by normal and brain-damaged speakers at a conversational speaking rate. Most models on Mandarin intonation are in terms of contour interaction (Chao, 1968; Shen, 1992; Yuan, 2004). For example, Chao (1968) likened syllabic tone and sentence intonation to small ripples riding on large waves in the ocean and stated that they interact by addition. Based on perceptual and acoustic studies, Yuan (2004) proposed a tone-dependent mechanism for intonation perception, which flattens the falling slope of the falling tone (such as Mandarin tone 4) and steepens the rising slope of the rising tone (such as Mandarin tone 2). It can be reasoned that the contrast of declarative and interrogative intona-

tion might be detected more difficult or easier for certain tone. In a prior experiment, we found a clear MMN to intonation contrast for Mandarin tone 4 (Ren et al., 2009 b).

In summary, the present study demonstrated the right hemispheric dominance of lexical tone in the early pre-attentive processing, which is compatible with the acoustic hypothesis (Zatorre & Belin, 2001) and the dynamic models (Gandour et al., 2004; Tong et al., 2005 ; Luo et al., 2006), but cannot be counted by the functional hypothesis (Pell & Baum, 1997, Wong, 2002). Moreover, the current results provide clearly evidence that listeners can not tease apart the declarative and interrogative intonation when the target was Mandarin tone 2 at the early stage of pre-attentive processing. However, how tone and intonation interact and how intonation is perceived remain to be determined, and we will focus on them in the further experiments.

Acknowledgements

This research was supported by Grants from the National Natural Science Foundation of China (31100732 and 31271091) and Specialized Research Fund for the Doctoral Program of Higher Education (20112136120003).

Author details

Gui-Qin Ren¹, Yi-Yuan Tang², Xiao-Qing Li³ and Xue Sui¹

*Address all correspondence to: renguiqin@126.com

1 School of Psychology, Liaoning Normal University, Dalian, China

2 Institute of Neuroinformatics and Laboratory for Body and Mind, Dalian University of Technology, Dalian, China

3 State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

References

- [1] Andersen, T, Tiippana, S, Laarni, K, Kojo, J, & Sams, I. M. ((2009). The role of visual spatial attention in audiovisual speech perception. *Speech Communication*, , 51, 184-193.
- [2] Astheimer, L, & Sanders, L. (2009). Listeners modulate temporally selective attention during natural speech processing. *Biological. Psychology*, 80 (1), 23-34.

- [3] Astheimer, L, & Sanders, L. (2012). Temporally selective attention supports speech processing in 3- to 5-year-old children. *Developmental Cognitive Neuroscience*, , 2, 120-128.
- [4] Chao, Y. R. (1968). *A Grammar of Spoken Chinese*, Berkeley: University of California Press, Berkeley, CA.
- [5] Cruttenden, A. (1997). *Intonation*. (2nd ed.). Cambridge, England: Cambridge University Press.
- [6] Gandour, J, Dziedzic, M, Wong, D, Lowe, M, Tong, Y. X, Hsieh, L, et al. (2003). Temporal integration of speech prosody is shaped by language experience: An fMRI study. *Brain and Language*, , 84, 318-336.
- [7] Gandour, J, Ponglorpisit, S, Potisuk, S, Khunadorn, F, Boongird, P, & Dechongkit, S. (1997). Interaction between tone and intonation in Thai after unilateral brain damage. *Brain and Language*, , 58, 174-196.
- [8] Gandour, J, Tong, Y, Wong, D, Talavage, T, Dziedzic, M, Xu, Y, et al. (2004). Hemispheric roles in the perception of speech prosody. *NeuroImage*, , 23, 344-357.
- [9] Gandour, J, Wong, D, Lowe, M, Dziedzic, M, Sathamnuwong, N, Long, Y, et al. (2002). Neural circuitry underlying perception of duration depends on language experience. *Brain and Language*, , 83, 268-290.
- [10] Hugdahl, K, Thomsen, T, Ersland, L, Rimol, L, & Niemi, M. J. ((2003). The effects of attention on speech perception: an fMRI study. *Brain and Language*, , 85, 37-48.
- [11] Jones, J. A, & Munhall, K. G. (1997). The effects of separating auditory and visual sources on audiovisual integration of speech. *Canadian Acoustics*, , 25, 13-19.
- [12] Klein, D, Zatorre, R, Milner, B, & Zhao, V. PET study of tone perception in Mandarin Chinese and English speakers. *NeuroImage*, , 13, 646-653.
- [13] Levänen, S, Ahonen, A, Hari, R, Mcevoy, L, & Sams, M. (1996). Deviant auditory stimuli activate human left and right auditory cortex differently. *Cerebral Cortex*, 6, 288-296.
- [14] Liu, Y, & Perfetti, C. A. (2003). The time course of brain activity in reading English and Chinese: An ERP study of Chinese bilinguals. *Human Brain Mapping*, , 18, 167-175.
- [15] Luo, H, Ni, J. T, Li, Z. H, Li, X. O, Zhang, D. R, Zeng, F. G, et al. (2006). Opposite patterns of hemisphere dominance for early auditory processing of lexical tones and consonants. *Proceedings of the National Academy of Sciences of the United States of America*, , 103, 19558-19563.
- [16] Marco-pallarés, L, Grau, C, & Ruffini, G. (2005). Combined ICA-LORETA analysis of mismatch negativity. *NeuroImage*, , 25, 471-477.

- [17] Molholm, S, Martinez, A, Ritter, W, Javitt, D. C, & Foxe, J. J. (2005). The neural circuitry of pre-attentive auditory change-detection: An fMRI study of pitch and duration mismatch negativity generators. *Cerebral Cortex*, 15, 545-551.
- [18] Mulert, C, Leicht, G, Pogarell, O, Mergl, R, Karch, S, Juckel, G, et al. (2007). Auditory cortex and anterior cingulate cortex sources of the early evoked gamma-band response: relationship to task difficulty and mental effort, *Neuropsychologia*, , 45, 2294-2306.
- [19] Navarra, J, Alsius, A, Soto-faraco, S, & Spence, C. (2010). Assessing the role of attention in the audiovisual integration of speech. *Information Fusion*, , 11, 4-11.
- [20] Näätänen, R, & Escera, C. (2000). Mismatch negativity: clinical and other applications. *Audiol Neurootol*, , 5, 105-110.
- [21] Näätänen, R, Kujala, T, Escera, C, Baldeweg, T, Kreegipuu, K, Carlson, S, et al. (2012). The mismatch negativity (MMN)- A unique window to disturbed central auditory processing in ageing and different clinical conditions. *Clinical Neurophysiology*, , 123, 424-458.
- [22] Näätänen, R, Paavilainen, P, Rinne, T, & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, 118, 2544-2590.
- [23] Näätänen, R, Pakarinen, S, Rinne, T, & Takegata, R. (2004). The mismatch negativity (MMN): towards the optimal paradigm. *Clinical Neurophysiology*, 115, 140-144.
- [24] Näätänen, R, Tervaniemi, M, Sussman, E, Paavilainen, P, & Winkler, I. (2001). Primitive intelligence' in the auditory cortex. *Trends in Neurosciences*, , 24, 283-288.
- [25] Pascual-marqui, R. D, Esslen, M, Kochi, K, & Lehmann, D. (2002). Functional imaging with low resolution brain electromagnetic tomography (LORETA): review, new comparisons, and new validation. *Japanese Journal of Clinical Neurophysiology*, 30, 81-94.
- [26] Pascual-marqui, R. D, Michel, C. M, & Lehmann, D. (1994). Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *International Journal of Psychophysiology*, 18, 49-65.
- [27] Pell, M. D. (2006). Cerebral mechanisms for understanding emotional prosody in speech. *Brain and Language*, , 96, 221-234.
- [28] Pell, M, & Baum, S. (1997). Unilateral brain damage and the acoustic cues to prosody: Are prosodic comprehension deficits perceptually based? *Brain and Language*, , 57, 195-214.
- [29] Pulvermüller, F, & Shtyrov, Y. (2006). Language outside the focus of attention: the mismatch negativity as a tool for studying higher cognitive processes. *Progress in Neurobiology*, 79, 49-71.

- [30] Pulvermüller, F, & Shtyrov, Y. (2003). Automatic processing of grammar in the human brain as revealed by the mismatch negativity. *NeuroImage*, 20, 159-172.
- [31] Pulvermüller, F, Shtyrov, Y, Ilmoniemi, R. J, & Marslen-wilson, W. D. (2006). Tracking speech comprehension in space and time. *NeuroImage*, 31, 1297-1305.
- [32] Ren, G. Q, Han, Y. C, Zou, Y. L, & Ren, Y. T. (2011). Early cortical processing of Mandarin intonation, *Acta Psychologica Sinica (in Chinese)*, 43, 241-248.
- [33] Ren, G. Q, Liu, Y, & Han, Y. C. (2009 a). Phonological activation in Chinese reading: an event-related potential study using low-resolution electromagnetic tomography. *Neuroscience*, 164, 1623-1631.
- [34] Ren, G. Q, Yang, Y, & Li, X. (2009 b). Early cortical processing of linguistic pitch patterns as revealed by the mismatch negativity. *Neuroscience*, 162, 87-95.
- [35] Shen, J. (1992). Hanyu yudiao moxing chuyi [On Chinese intonation model]. *Yuwen Yanjiu*, 45, 16-24.
- [36] Talairach, J, & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. (Rayport, M., Translator.), New York: Thieme Medical Publishers.
- [37] Tong, Y, Gandour, J, Talavage, T, Wong, D, Dziedzic, M, Xu, Y, Li, X, et al. (2005). Neural circuitry underlying sentence-level linguistic prosody. *NeuroImage*, 28, 417-428.
- [38] Tsang, Y-K, Jia, S, Huang, J, & Chen, H-C. (2011). ERP correlates of pre-attentive processing of Cantonese lexical tones: the effects of pitch contour and pitch height. *Neuroscience Letters*, 487, 268-272.
- [39] Warrier, C. M, & Zatorre, R. J. (2004). Right temporal cortex is critical for utilization of melodic contextual cues in a pitch constancy task. *Brain*, 127, 1616-1625.
- [40] Wong, P. C. M. (2002). Hemispheric specialization of linguistic pitch patterns. *Brain Research Bulletin*, 59, 83-95.
- [41] Wang, Y, Jongman, A, & Sereno, J. A. (2001). Dichotic perception of Mandarin tones by Chinese and American listeners. *Brain and Language*, 78, 332-348.
- [42] Xi, J, Zhang, L. J, Shu, H, Zhang, Y, & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience*, 170, 223-231.
- [43] Xu, Y. (2005). Speech melody as articulatorily implemented communicative functions. *Speech Communication*, 46, 220-251.
- [44] Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, 25, 61-83.
- [45] Yuan, J. (2004). Intonation in Mandarin Chinese: acoustics, perception, and computation modeling, *Cornell University*, dissertation.

- [46] Zatorre, R. J, & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, 11, 946-953.
- [47] Zatorre, R. J, Belin, P, & Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends in Cognitive Science*, , 6, 37-46.