Neuroscience 170 (2010) 223–231

CATEGORICAL PERCEPTION OF LEXICAL TONES IN CHINESE REVEALED BY MISMATCH NEGATIVITY

J. XI,[a,b] L. ZHANG,[c] H. SHU,[a]* Y. ZHANG[d]* AND P. LI[e]

[a]*State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, PR China*

[b]*State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, PR China*

[c]*College of Chinese Studies, Beijing Language and Culture University, Beijing, PR China*

[d]*Department of Speech-Language-Hearing Sciences and Center for Neurobehavioral Development, University of Minnesota, MN 55455, USA*

[e]*Department of Psychology and Center for Language Science, Pennsylvania State University, PA 16802, USA*

Abstract—The present study investigated the neurophysiological correlates of categorical perception of Chinese lexical tones in Mandarin Chinese. Relative to standard stimuli, both within- and across-category deviants elicited mismatch negativity (MMN) in bilateral frontal-central recording sites. The MMN elicited in the right sites was marginally larger than in the left sites, which reflects the role of the right hemisphere in acoustic processing. At the same time, relative to within-category deviants, the across-category deviants elicited larger MMN in the left recording sites, reflecting the long-term phonemic traces of lexical tones. These results provide strong neurophysiological evidence in support of categorical perception of lexical tones in Chinese. More important, they demonstrate that acoustic and phonological information is processed in parallel within the MMN time window for the perception of lexical tones. Finally, homologous nonspeech stimuli elicited similar MMN patterns, indicating that lexical tone knowledge influences the perception of nonspeech signals.

Key words: mismatch negativity (MMN), categorical perception, lexical tones, Mandarin Chinese.

Categorical perception is one of the most extensively studied phenomena in speech perception. It refers to the ability that human listeners can perceive continuous acoustic signals as discrete linguistic representations. While most of the early categorical perception studies focused on segmental features (consonants and vowels) (Fry et al., 1962; Harnad, 1987; Liberman et al., 1957, 1961), there has been a recent surge of interests in whether categorical perception applies to the suprasegmental level, including vowel duration contrasts of quantity languages[1] (Nenonen et al., 2003; Ylinen et al., 2005) and lexical tone contrasts of tonal languages (Francis et al., 2003; Hallé et al., 2004; Xu et al., 2006). For example, behavioral results of lexical tone perception show that continua ranging from one tone to another are not perceived categorically unless the tonal continua involve contour tones (Abramson, 1979; Francis et al., 2003; Hallé et al., 2004; Wang, 1976; Xu et al., 2006). In Mandarin Chinese there are four lexical tones, only one of which (high level tone, Tone1) does not involve a contour (Howie, 1976). Thus, native perception of Chinese tonal continua is expected to be categorical with better sensitivity to across-category distinction relative to within-category differences.

Advances in brain imaging and neurophysiological studies have allowed researchers to address fundamental questions about the neural mechanisms underlying behavioral responses for lexical tone perception (Chandrasekaran et al., 2007, 2009; Gandour et al., 2000, 2002, 2004; Gandour, 2006; Hsieh et al., 2001; Krishnan et al., 2009; Ren et al., 2009; Wang et al., 2003; Wong et al., 2004). The neural correlates of categorical perception of lexical tones, however, have not been carefully examined. The only exception was Chandrasekaran et al. (2009), in which across-category (Tone1/Tone2) and within-category (Tone2/Tone2L) tonal contrasts were investigated with an electrophysiological measure, and a larger across-category than within-category mismatch negativity (MMN) was found. However, Chandrasekaran's study used stimuli from different tonal continua and the physical differences for the two contrasts in comparison were not equated, and therefore, their conclusions must remain tentative.

The neurophysiological study of categorical perception of lexical tones has particular theoretical significance for the understanding of the dynamic interplay between acoustic and phonological processing at the suprasegmental level. Recent neuroimaging studies have demonstrated that the two hemispheres play different roles in lexical tone perception; specifically, the right hemisphere is involved in auditory/acoustic processing, whereas the left hemisphere is responsive to linguisic/phonological functions (Gandour et al., 2000, 2002, 2004; Gandour, 2006; Hsieh et al., 2001; Wang et al., 2003; Wong et al., 2004). But the interplay between auditory and phonological processing has not been addressed in these imaging studies. There are at least two fundamental questions of interest: (a) whether the lower-level auditory processing precedes the higher-level phonological processing, and (b) whether the two processes interact with each other during processing. Zhao et al. (2008) provided some functional magnetic

*Correspondence to: Y. Zhang, Department of Speech-Language-Hearing Sciences & Center for Neurobehavioral Development, University of Minnesota, MN 55455, USA. Tel: +1-612-6247818; fax: +1-612-6243322. Tel: +86-10-58808186 (H. Shu).
E-mail addresses: zhang470@umn.edu (Y. Zhang); shuh@bnu.edu.cn (H. Shu).
Abbreviations: EEG, electroencephalogram; ERP, event-related potential; fMRI, functional magnetic resonance imaging; MMN, mismatch negativity.

[1] In a quantity language, vowel duration gives rise to phonological contrast.

0306-4522/10 $ - see front matter © 2010 IBRO. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuroscience.2010.06.077

resonance imaging (fMRI) evidence on the second question, but due to the coarse temporal resolution of fMRI, they did not address the first question as that is related to the time course of auditory versus phonological processing for lexical tone perception.

Auditory neurophysiology provides one optimal approach towards examining the time course of neural responses and the interplay between acoustic and phonological processing of lexical tones. Luo et al. (2006) found right-lateralized MMN activation for Chinese lexical tone perception, and suggested a two-stage model according to which acoustic information is initially processed in the right hemisphere and then linguistic information is mapped in the left hemisphere. However, other studies have failed to find right lateralization of MMNs for lexical tones (Chandrasekaran et al., 2007, 2009). The discrepancies between studies may be due to confounds in the design of the relevant studies. As we pointed out earlier, in Chandrasekaran et al.'s study (2009), one confounding factor was that the physical differences for the two contrasts were not equated (across-category physical changes were much larger than the within-category physical changes), and the different MMN responses, therefore, could reflect both acoustic and phonological processing.

Categorical perception of lexical tones is primarily determined by language experience (Hallé et al., 2004; Xu et al., 2006). Research has further demonstrated that the experience-dependent pitch processing skills extend to the processing of nonspeech sounds (Bent et al., 2006; Xu et al., 2006). For example, Bent et al. found that Chinese listeners more often misidentified flat and rising pitch contours of simple nonspeech sound than English listeners, which was attributed to the influence of long-term categorical representations of lexical tones. Some studies have also investigated the neural substrates for the perception of nonspeech homologues of lexical tones by native speakers (Chandrasekaran et al., 2007; Krishnan et al., 2009), and showed that native speakers exhibit stronger categorical representations of nonspeech material relative to nonnative speakers. In these studies, the perceptual advantage of native speakers was attributed to more efficient auditory processing of pitch, although it is equally possible that categorical perception of nonspeech stimuli is influenced by long-term phonological representations of speech material in the native language.

To sum up, although the neural mechanisms underlying the perception of lexical tones have been extensively investigated, the neural correlates of categorical perception of lexical tones, however, have not been carefully examined. By studying categorical perception of lexical tones, we can identify the dynamic interplay between acoustic and phonological processing, particularly with reference to whether the two processes take place serially or in parallel. Furthermore, we can also identify the influences that phonological representations of lexical tones might have on the perception of nonspeech homologues. Considering methodological confounds in previous studies, in the present study we used well-controlled synthetic tonal continua with equal physical intervals to investigate categorical perception of lexical tones in order to differentiate acoustic and phonological processing of both speech and nonspeech stimuli.

EXPERIMENTAL PROCEDURES

Participants

Sixteen neurologically healthy volunteers (eight females, eight males; aged 20–26, mean age=22) with normal hearing and minimal musical experience (less than 1 year of total musical training and no musical training within the past 5 years) participated in the study. All participants were native speakers of Chinese, and all were college students at Beijing Normal University. They were all right-handed according to a handedness questionnaire adapted from a modified Chinese version of the Edinburgh Handedness Inventory (Oldfield, 1971). Participants gave written consent before they took part in the experiment. The experiment was approved by the ethics review board at Beijing Normal University's Imaging Center for Brain Research. Data from two female participants were excluded from further analyses due to insufficient number of acceptable trials as a result of strong electroencephalogram (EEG) artifacts including excessive blinking.

Stimuli

Two sets of stimuli were constructed for this experiment: speech and nonspeech. The speech stimuli were Chinese monosyllables /pa/ that differed in their lexical tones (the high rising tone, Tone2 and the falling tone, Tone4). The original stimuli were recorded at a sampling rate of 44.1 kHz from a female native Chinese speaker. The monosyllables were digitally edited using SoundForge (SoundForge9, Sony Corporation, Japan), each having a duration of 200 ms. In order to isolate the lexical tones and keep the rest of the acoustic features identical, pitch tier transfer was performed using the Praat software (*http://www.fon.hum.uva.nl/praat/*). This procedure generated two stimuli, /pa2/ and /pa4/, which were identical with each other except for the pitch contour difference. The /pa2/ and /pa4/ stimuli were then taken as the endpoint stimuli to create a 10-interval lexical tone continuum. A morphing technique was performed in Matlab (Mathworks Corporation, USA) using STRAIGHT (Kawahara et al., 1999) in 10 equal intervals. All 11 stimuli were normalized in RMS intensity. The 11 stimuli in the /pa2/-/pa4/ lexical tone continuum were examined in a behavioral test (identification and discrimination tasks) on 16 native Chinese speaking adults who did not participate in the subsequent event-related potential (ERP) study. The identification task required participants to identify each stimulus. In the session, 40 trials for each stimulus were randomly presented in isolation. The AX discrimination task required participants to judge whether pairs of presented stimuli were the same or different. Stimulus presentation was randomized and both directions of presentation order were tested, each for 20 trials. Based on the behavioral test results (Fig. 1), an across-category stimulus pair (3 and 7) and a within-category stimulus pair (7 and 11) were chosen for the ERP oddball paradigm experiment. In particular, stimulus 7 in the continuum was used as the standard stimulus, and stimuli 3 (an across-category deviant) and 11 (a within-category deviant) were used as deviants.[2]

The nonspeech stimuli were harmonic tones with the same pitch, amplitude, and duration parameters as the speech stimuli.

[2] A behavioral test was carried out on the ERP participants after the ERP experiment in which the participants were asked to identify the three (3, 7, 11) stimuli as one of the four lexical tones. All stimuli were correctly identified by all the participants (0% error rate, 10 occurrences of each stimuli). Therefore, the ERP participants indeed perceived 3 vs. 7 as an across-category contrast and 7 vs. 11 as a within-category contrast.

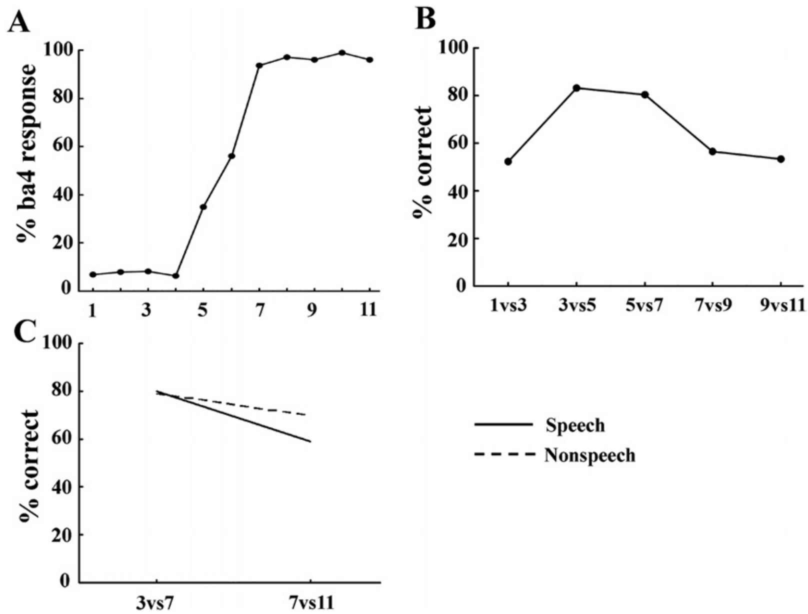

Fig. 1. Behavioral results. (A) Identification task for the speech stimuli. (B) Two-step discrimination task for the speech stimuli. (C) Continua 3 vs. 7 and continua 7 vs. 11 discrimination task for both the speech and nonspeech stimuli.

The only difference between the speech and nonspeech stimuli was in the spectral components. The nonspeech stimuli were composed of six equal amplitude harmonics (1, 3, 6, 7, 8, 12) of the F0. Harmonics 2, 4, 5, 9, 10, and 11 were omitted to increase perceptual dissimilarity with the speech stimuli. All the stimuli were normalized to have equal average RMS intensity (Fig. 2).

ERP procedure

Two oddball blocks were presented to each participant with counterbalanced sequence among the participants. The within-category deviant and across-category deviant occurred pseudo-randomly among standards with a probability of 10% respectively, and any two adjacent deviants were separated by at least three standards for a total presentation of 1000 stimuli in each presentation block. The stimulus-onset-asynchrony (SOA) was 800 ms.

Participants were seated comfortably in an acoustically and electrically shielded chamber and instructed to ignore the presented sounds while watching a self-selected movie. The movie was presented in mute mode with subtitles on a screen of 15 cm×10 cm positioned approximately 1 m in front of the seat. Participants were informed to complete a questionnaire about the movie after the experiment. Sound stimuli were presented binaurally via an insert Sony earphone. The right and left acoustic channels of the insert earphone were calibrated for equal and comfortable loudness (70 dB SPL) prior to the experiment. The experimental session lasted approximately 1.5 h including preparation, data acquisition, breaks, and cleanup.

Electroencephalogram (EEG) recording

Continuous EEG was recorded using a HydroCel Geodesic Sensor Net (HCGSN) consisting of 128 electrodes evenly distributed

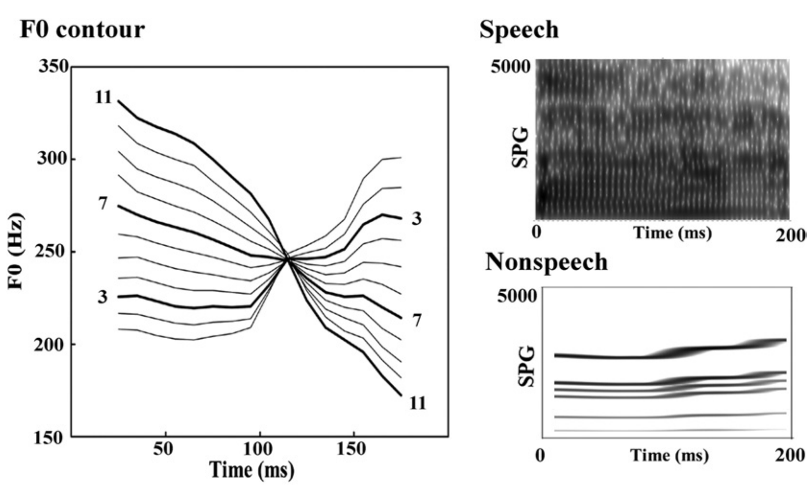

Fig. 2. Schematic illustration of the stimuli. The left panel shows the F0 continuum which the speech stimuli are based on (continua 3, 7 and 11 are marked with thick lines). The top-right panel shows a broadband spectrogram of continuum 3 of the speech stimulus. The bottom-right panel shows a narrowband spectrogram of its nonspeech homologue.

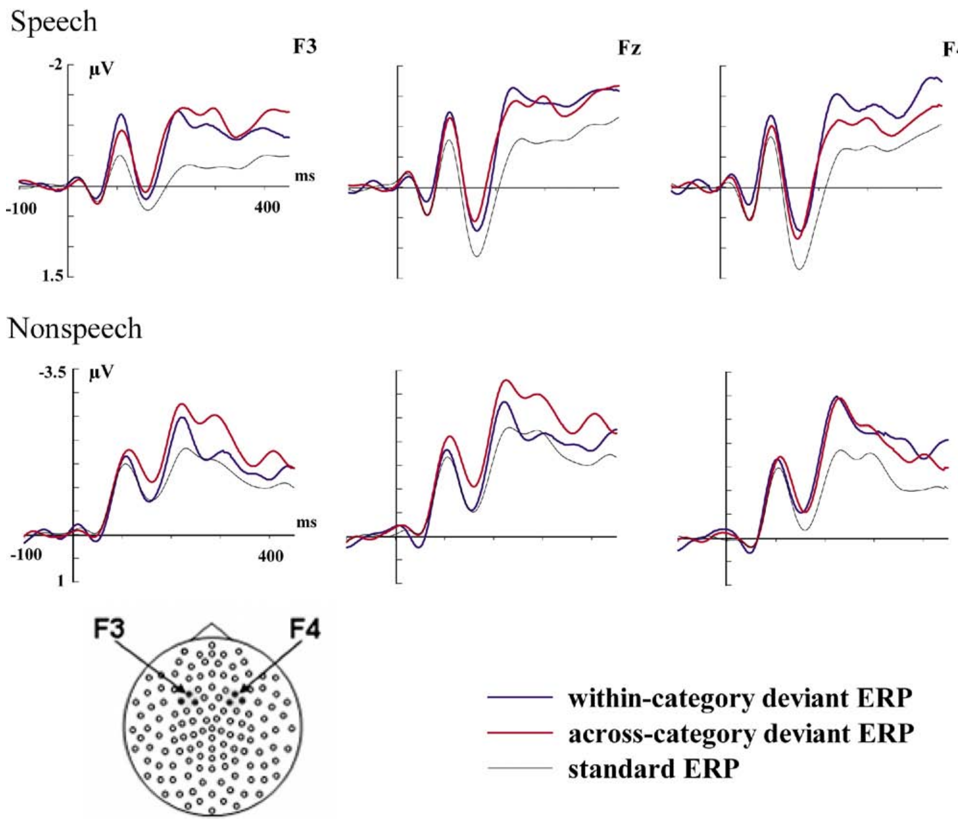

Fig. 3. Grand average waveforms elicited by the across-category deviants (P=10%), within-category deviants (P=10%) and the standards (P=80%) in speech and nonspeech conditions.

across the scalp and referenced against the vertex electrode (Tucker, 1993). The GSN also includes electrodes next to, and below, the eyes for recording horizontal and vertical eye-movements. The impedance of each electrode was kept below 5 kΩ.

Data analysis

Off-line signal processing was carried out using Netstation software (Version 4.2). The raw data were first digitally filtered with a 0.3~20 Hz bandpass filter and segmented for 700 ms starting 100 ms prior to the onset of stimuli. Data were then re-referenced to the average of all the electrodes, and baseline corrected. Recorded trials with eye blinks or other activities beyond the range of −50~50 μV were rejected.

Based on previous studies and visual inspection of the grand-average waveforms, two recording sites were selected for statistical analyses: left frontal (sites F3, channels 19, 23 and 24) and right frontal (sites F4, channels 3, 4 and 124) (Díaz et al., 2008; Kaan et al., 2007; Kirmse et al., 2008; Luo et al., 2006; Oken and Chiappa, 1986). Only the standard before the deviant was used for averaging and subtraction. Difference waves for MMN were obtained by subtracting the averaged standard from the averaged deviant. The MMN peak latency for each subject was found within a 60 ms time window that was defined by the grand-average waveforms at Fz. The calculation of mean amplitude was conducted with a moving window technique: first the negative peak within the 60 ms time window was found for each subject, then the value of a time window which extended ±40 ms surrounding the MMN peak was averaged. Statistical analysis only included those participants with at least 80 accepted deviant trials in each condition.

RESULTS

The grand average waveforms elicited by the standards and the deviants are shown in Fig. 3. Negative peaks were observed in the deviant-minus-standard difference waves for the speech (lexical tone) and nonspeech (harmonic tone) conditions (see Fig. 4). Two three-way condition (lexical tone/harmonic tone)×deviant type (within-category/across-category)×hemisphere (left/right) repeated measures ANOVAs were conducted for mean amplitude and peak latency respectively. For all analyses, degrees of freedom were adjusted according to the method of Greenhouse–Geisser when appropriate.

MMN peak latency

The mean peak latencies at the electrodes (F3, F4) were plotted in the upper panel of Fig. 5. Results from the omnibus ANOVA yielded a significant main effect of condition (F(1, 13)=17.196, P=0.001, nonspeech>speech), showing that speech was processed faster than nonspeech. The main effect of deviant type or hemisphere failed to reach significance. None of the two- or three-way interactions reached significance.

MMN mean amplitude

For the MMN amplitudes (lower panel of Fig. 5), the main effect of condition was not significant (F(1, 13)=1.498,

$P=0.243$). The main effect of deviant type was significant ($F(1, 13)=7.382$, $P=0.018$, across-category>within-category) and the main effect of laterality was marginally significant ($F(1, 13)=3.693$, $P=0.077$, right hemisphere>left hemisphere). There was also a significant interaction between deviant type and hemisphere ($F(1, 13)=5.790$, $P=0.032$), indicating that the amplitude of the response to across-category deviant was greater than response to within-category deviant in the left hemisphere ($F(1, 13)=20.80$, $P=0.001$), but not in the right hemisphere ($F(1, 13)=0.16$, $P=0.694$). No other effects reached statistical significance.

DISCUSSION

The present study examined the neurophysiological, particularly MNN, indices of categorical perception of lexical tones in native speakers of Mandarin Chinese. It was found that across-category contrast elicited a larger MMN than within-category distinction for Chinese listeners. Because physical intervals for the two contrasts were equated, the study provided strong neurophysiological evidence in support of categorical perception of lexical tones. Previous studies using a similar design showed that the MMN response is enhanced by the listener's detection of across-category differences in vowels (Winkler et al., 1999), place of articulation (Dehaene-Lambertz, 1997), voice onset time (VOT) (Sharma and Dorman, 2000) and vowel length of quantity languages (Nenonen et al., 2003). Our results are consistent with these studies, supporting the view that when listeners process a speech signal to recover the phonological

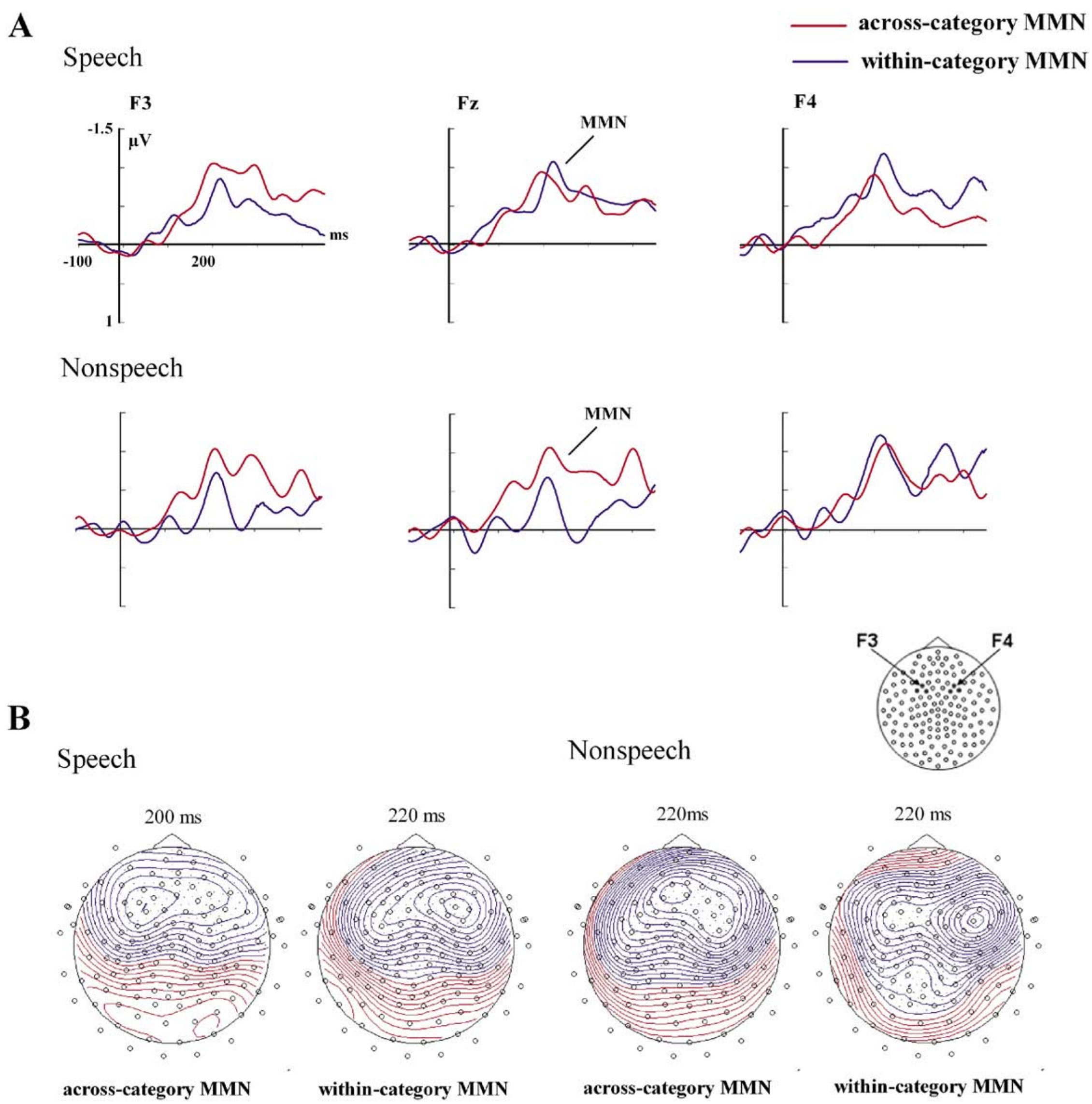

Fig. 4. Mismatch negativities (MMNs) recorded in speech and nonspeech conditions. (A) Grand average traces of MMN evoked by within- and across-category changes from the F3, Fz, and F4 electrode locations; (B) Maps display the topographic distribution of the mean amplitudes in the MMN analysis window.

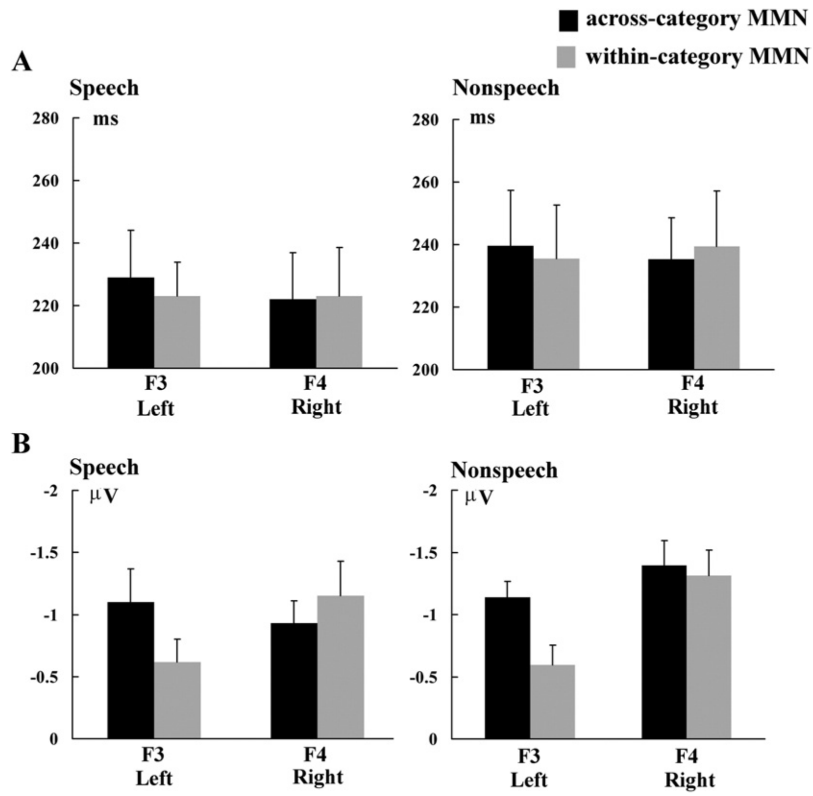

Fig. 5. MMN peak latency (A) and mean amplitude (B) values from F3 and F4 electrodes (vertical bars represent one standard error).

representations of their native language, they automatically extract the categorical linguistic information and show reduced sensitivity towards within-category acoustic differences.

With regard to laterality of lexical tone perception, two competing hypotheses have been proposed in the literature. The acoustic hypothesis claims that speech perception is cue-dependent and that pitch is mediated by the right hemisphere regardless of linguistic functions (Klouda et al., 1988; Zatorre and Belin, 2001). In contrast, the functional hypothesis claims that speech perception is task-dependent and that pitch patterns carrying linguistic functions are lateralized to the left hemisphere while those carrying no linguistic functions lateralized to the right hemisphere (Van Lancker, 1980; Wong, 2002). Both hypotheses have found empirical support in the past (Gandour et al., 2000, 2002; Wong et al., 2004; Zatorre and Belin, 2001). In recent years, a more comprehensive model that integrates the acoustic hypothesis and the functional hypothesis has been put forward by Gandour and his colleagues (Gandour et al., 2004; Tong et al., 2005; Zatorre and Gandour, 2008), according to which the right hemisphere is sensitive to low-level acoustic processing and the left hemisphere is sensitive to high-level linguistic processing.

Data from the present study provided neurophysiological evidence in testing these models. We found that relative to standard stimuli, both within- and across-category deviants elicited larger MMNs in the right frontal-central recording sites, which is consistent with previous studies (Luo et al., 2006; Ren et al., 2009). The right lateralized response presumably reflects the role of the right hemisphere in acoustic processing. At the same time, relative to within-category deviants, the across-category deviants elicited larger MMN, which probably reflects the long-term phonemic traces of lexical tones because the ERP topography data showed that the enhancement was dominant in the left frontal-central recording sites (Näätänen et al., 1997). These results are consistent with the hypothesis of Gandour et al. (2004) with regard to the differential functional roles of the two hemispheres. Given the low spatial resolution of ERP, however, the results should be considered preliminary. Magnetoencephalography (MEG) study with high temporal resolution and acceptable spatial resolution is needed in order to clearly identify the temporal dynamics of cortical activation in different brain regions involved in the processing of Chinese lexical tones.

Previous models, however, have not addressed the time course issue with regard to whether auditory/acoustic processing precedes high-level linguistic/phonological processing or whether the two types of processing occur in parallel. Our data provided insights into this issue. Findings from our study indicate that acoustic processing and linguistic processing interact at an early stage, in that acoustic and phonological information is activated in parallel for lexical tone perception within the short MMN time window. This finding provides counter evidence to the two-stage model for lexical tone processing (Luo et al., 2006), according to which lower-level auditory processing precedes higher-level phonological processing.

Another issue that our data help to resolve is the question of how linguistic experience affects the perception of nonspeech sounds. One view is that there exist separate modules for speech and nonspeech processing, and that exposure to the sound structure of a particular language has no influence on nonspeech auditory processing (Burnham et al., 1996; Liberman and Mattingly, 1989; Miyawaki et al., 1975; Zhang et al., 2005). The contrasting view is that there are shared speech and nonspeech processing mechanisms, and that language learning exerts strong influence on the perception of nonspeech sounds (Bent et al., 2006; Francis et al., 2003). Such discrepant views could be due to the fact that speech contrasts such as lexical tone categories, vowels, and approximants in the cited studies are cued by markedly different spectral and temporal features—some acoustic features in speech may be directly analyzed on the basis of general auditory processing capacity, and others may reflect language-specific processing of the neural system. Our findings on lexical tone processing are consistent with the second view—the MMN amplitude data for nonspeech stimuli were similar to the results for speech stimuli. Our study accords well with studies of other languages, for example, by Sussman et al. (2004) who found similar patterns of MMN for speech and nonspeech counterparts for Finnish speakers.[3] Despite the MMN amplitude similarity between speech stimuli and nonspeech control, our analysis of MMN latencies revealed differences in speech vs. nonspeech processing. The MMN latency data showed that speech was processed faster than nonspeech. This latency effect could be attributed to listeners' linguistic processing or familiarity with the phonetic segments (i.e. /p/ and /a/) of the speech stimuli. Although speech and nonspeech stimuli in our study have the same pitch, amplitude, and duration parameters, the difference in spectral components clearly distinguishes the two. Therefore listeners' phonetic processing of the /pa/ syllable might have led to faster change detections.

An alternative view is that categorical perception of pitch contours results from short-term categorical memory, which runs parallel to fine-grain sensory encoding (Xu et al., 2006). The MMN difference between the within- and across-category contrasts, therefore, might be due to the rising versus falling pitch patterns in terms of feature abstraction rather than high-level phonological representation.[4] This view can also potentially account for our findings that the across-category deviants elicited larger MMN than the within-category deviants for both speech and nonspeech sounds. According to this view, for nonnative listeners, across-category deviants should also elicit larger MMN than within-category deviants, especially for the nonspeech stimuli.

Evidence from a separate ERP experiment that we conducted does not agree with this view. Using the same stimuli and procedure as in the current study, we tested Korean listeners who had no experience with Chinese lexical tones and found that both within- and across-category deviants elicited larger MMN in the right frontal-central recording sites, similar to the results of the Chinese participants. However, these listeners showed no difference in the MMNs elicited by the within- and across-category deviants, which is different from the results of the Chinese native listeners (see the supplementary materials for further details). These results therefore provide more convincing evidence that the MMN differences elicited by the within- and across-category deviants for the Chinese listeners in the present study could be attributed to high-level phonemic processing rather than low-level auditory processing. In recent fMRI studies, it has been found that categorical perception of phonemic sound patterns is mediated by the left middle superior temporal sulcus (Joanisse et al., 2007; Liebenthal et al., in press). Future fMRI studies on categorical perception of Chinese lexical tones could further help to clarify whether it is attributable to long-term phonological representations or to short-term categorical memory.

[3] Tervaniemi et al (2006) showed larger MMN responses to duration change of nonspeech sounds for Finnish speakers than for German speakers.

[4] We are grateful to an anonymous reviewer for pointing out this issue.

CONCLUSION

In conclusion, the present study used neurophysiological (MMN) indices to identify the neural correlates of categorical perception of Chinese lexical tones. Our findings have shown spatial (hemispheric) differences in auditory processing versus phonological processing but no temporal differences between the two types of processing—the two types of processing occur in parallel within the MMN time window and show early interaction between the two hemispheres. Our data also suggest common mechanisms underlying the processing of speech and nonspeech materials for native speakers, in which language-specific tonal experience with the perception of lexical tones influences the processing of nonspeech signals.

Acknowledgments—The research was supported by the Program for Changjiang Scholars and Innovative Research Team at Beijing Normal University, by grants from the Natural Science Foundation of Beijing (7092051) to HS and in part from the National Science Foundation of USA (BCS#0642586) to PL. YZ was supported by a Single Semester Leave Award and two CLA Brain Imaging Research Awards from the University of Minnesota to work on the project. The authors thank four anonymous reviewers for suggestions to improve the manuscript.

REFERENCES

Abramson AS (1979) The noncategorical perception of tone categories in Thai. In: Frontiers of speech communication research, (Lindblom B, Öhman S, eds), pp 127–134. London: Academic Press.

Bent T, Bradlow AR, Wright BA (2006) The influence of linguistic experience on the cognitive processing of pitch in speech and nonspeech sounds. J Exp Psychol Hum Percept Perform 32: 97–103.

Burnham D, Francis E, Webster D (1996) The development of tone perception: cross-linguistic aspects and the effect of linguistic context. In: Pan-Asiatic linguistics: proceedings of the 4th International Symposium, ILCRDMUS, Mahidol University.

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

Chandrasekaran B, Krishnan A, Gandour J (2009) Relative influence of musical and linguistic experience on early cortical processing of pitch contours. Brain Lang 108:1–9.

Dehaene-Lambertz G (1997) Electrophysiological correlates of categorical phoneme perception in adults. Neuroreport 8:919–924.

Díaz B, Baus C, Escera C, Costa A, Sebastián-Gallés N (2008) Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language. Proc Natl Acad Sci U S A 105:16083–16088.

Francis AL, Ciocca V, Ng BK (2003) On the (non)categorical perception of lexical tones. Percept Psychophys 65:1029–1044.

Fry DB, Abramson AS, Eimas PD, Liberman AM (1962) The identification and discrimination of synthetic vowels. Lang Speech 5:171–189.

Gandour J (2006) Brain mapping of Chinese speech prosody. In: Handbook of east Asian psycholinguistics, Vol. 1 (Li P, Tan L, Bates E, Tzeng O, eds), pp 308–319. Cambridge, UK: Cambridge University Press. [Chinese].

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 G (2000) A crosslinguistic PET study of tone perception. J Cogn Neurosci 12:207–222.

Gandour J, Wong D, Lowe M, Dzemidzic M, Satthamnuwong N, Tong Y, Lurito J (2002) Neural circuitry underlying perception of duration depends on language experience. Brain Lang 83:268–290.

Hallé PA, Chang YC, Best CT (2004) Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. J Phon 32:291–453.

Harnad S, ed (1987) Categorical perception: the groundwork of cognition. New York, NY: Cambridge University Press.

Howie JM (1976) Acoustical studies of mandarin vowels and tones. New York, NY: Cambridge University Press.

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

Joanisse MF, Zevin JD, McCandliss BD (2007) Brain mechanisms implicated in the preattentive categorization of speech sounds revealed using fMRI and short-interval habituation trial paradigm. Cereb Cortex 17:2084–2093.

Kaan E, Wayland R, Bao M, Barkley CM (2007) Effects of native language and training on lexical tone perception: an event-related potential study. Brain Res 1148:113–122.

Kawahara H, Masuda-Katsuse I, de Cheveigne A (1999) Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction: possible role of a repetitive structure in sounds. Speech Commun 27:187–207.

Kirmse U, Ylinen S, Tervaniemi M, Vainio M, Schröger E, Jacobsen T (2008) Modulation of the mismatch negativity (MMN) to vowel duration changes in native speakers of Finnish and German as a result of language experience. Int J Psychophysiol 67:131–143.

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

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

Liberman AM, Harris KS, Hoffman HS, Griffith BC (1957) The discrimination of speech sounds within and across phonemic boundaries. J Exp Psychol 54:358–368.

Liberman AM, Harris KS, Kinney JA, Lane H (1961) The discrimination of relative onset-time of the components of certain speech and nonspeech patterns. J Exp Psychol 61:379–388.

Liberman AM, Mattingly IG (1989) A specialization for speech perception. Science 243:489–494.

Liebenthal E, Desai R, Ellingson MM, Ramachandran B, Desai A, Binder JR (in press) Specialization along the left superior temporal sulcus for auditory categorization. Cereb Cortex, in press.

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.

Miyawaki K, Strange W, Verbrugge R, Liberman A, Jenkins J, Fujimura O (1975) An effect of language experience: the discrimination of /r/ and /l/ by native speakers of Japanese and English. Percept Psychophys 18:331–340.

Näätänen R, Lehtokoski A, Lennes M, Cheour M, Huotilainen M, Iivonen A, Vainio M, Alku P, Ilmoniemi RJ, Luuk A, Allik J, Sinkkonen J, Alho K (1997) Language-specific phoneme representations revealed by electric and magnetic brain responses. Nature 385:432–434.

Nenonen S, Shestakova A, Huotilainen M, Näätänen R (2003) Linguistic relevance of duration within the native language determines the accuracy of speech-sound duration processing. Cogn Brain Res 16:492–495.

Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9:97–113.

Oken BS, Chiappa KH (1986) Statistical issues concerning computerized analysis of brainwave topography. Ann Neurol 19:493–494.

Ren GQ, Yang Y, Li X (2009) Early cortical processing of linguistic pitch patterns as revealed by the mismatch negativity. Neuroscience 162:87–95.

Sharma A, Dorman MF (2000) Neurophysiological correlates of cross-language phonetic perception. J Acoust Soc Am 107:2697–2703.

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.

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.

Tucker DM (1993) Spatial sampling of head electrical fields: the geodesic sensor net. Electroencephalogr Clin Neurophysiol 87:154–163.

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

Wang WSY (1976) Language change. Ann N Y Acad Sci 208:61–72.

Wang Y, Sereno JA, Jongman A, Hirsch J (2003) fMRI evidence for cortical modification during learning of mandarin lexical tone. J Cogn Neurosci 15:1019–1027.

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

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

Winkler I, Lehtokoski A, Alku P, Vainio M, Czigler I, Csepe V, Aaltonen O, Raimo I, Alho K, Lang H, Iivonen A, Näätänen R (1999) Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations. Brain Res Cogn Brain Res 7: 357–369.

Xu Y, Gandour J, Francis A (2006) Effects of language experience and stimulus complexity on categorical perception of pitch direction. J Acoust Soc Am 120:1063–1074.

Ylinen S, Shestakova A, Alku P, Huotilainen M (2005) The perception of phonological quantity based on durational cues by native speakers, second-language users and nonspeakers of Finnish. Lang Speech 48:313–338.
Zatorre RJ, Belin P (2001) Spectral and temporal processing in human auditory cortex. Cereb Cortex 11:946–953.
Zatorre R, Gandour J (2008) Neural specializations for speech and pitch: moving beyond the dichotomies. Philos Trans R Soc Lond B Biol Sci 363:1087–1104.
Zhang Y, Kuhl PK, Imada T, Kotani M, Tohkura Y (2005) Effects of language experience: neural commitment to language-specific auditory patterns. Neuroimage 26:703–720.
Zhao J, Shu H, Zhang L, Wang X, Gong Q, Li P (2008) Cortical competition during language discrimination. Neuroimage 43:624–633.

APPENDIX

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroscience.2010.06.077.

(Accepted 28 June 2010)
(Available online 13 July 2010)