Electrophysiological evidence of categorical perception of Chinese lexical tones in attentive condition

Linjun Zhang[a], Jie Xi[b], Han Wu[c], Hua Shu[c] and Ping Li[d]

Previous electrophysiological evidence supports categorical perception of Chinese lexical tones at the preattentive stage (Xi and colleagues). In this study, we examined participants' attentive responses to tonal continua in an event-related potential experiment that recorded their N2b and P3b oddball responses. We found that for both the N2b and the P3b component, the responses elicited by the within-category deviants were similar in the left and the right recording sites. However, the across-category deviants elicited larger responses in the left recording sites than in the right sites, reflecting conscious phonological processing of lexical tones. These results provide electrophysiological correlates of categorical perception of Chinese lexical tones in later stages associated with controlled processes.

***NeuroReport* 23:35–39 © 2011 Wolters Kluwer Health | Lippincott Williams & Wilkins.**

NeuroReport 2012, **23**:35–39

Keywords: categorical perception, Chinese lexical tones, mismatch negativity, N2b, P3a, P3b

[a]Department of Chinese Linguistics, College of Chinese Studies, Beijing Language and Culture University, [b]Laboratory for Higher Brain Function, Institute of Psychology, Chinese Academy of Sciences, [c]State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China and [d]Department of Psychology and Center for Language Science, Pennsylvania State University, University Park, Texas, USA

Correspondence to Jie Xi, PhD, Laboratory for Higher Brain Function, Institute of Psychology, Chinese Academy of Sciences, No. 4, Datun Road, Chaoyang District, Beijing 100101, China
Tel: +8610 64836689; fax: +8610 58806154; e-mail: yangqiustar@gmail.com

Received 17 September 2011 **accepted** 17 October 2011

Introduction

In speech perception, the phenomenon of categorical perception has been extensively studied for more than half a century [1,2]. It refers to the ability of human listeners to perceive continuous acoustic signals as discrete, categorical, linguistic representations. Therefore, listeners are more likely to notice the differences between categories than within categories. Most of the early studies have focused on segmental features such as voice onset time and formant transitions [1–3], and there is a consensus that consonantal features are perceived in a categorical manner. In recent years, however, there has been increasing interest in suprasegmental features, especially in lexical tone contrasts of Mandarin Chinese. More behavioral results have indicated that native perception of Chinese lexical tones is categorical because any Chinese tonal continuum from one tone to another involves contour tones [4–6].

In addition to behavioral methods, electrophysiological tools have allowed to address fundamental questions concerning speech perception. Specifically, the event-related potential (ERP) components such as mismatch negativity (MMN), N2b, and P3 index various cognitive processes taking place during stimulus perception, identification, and evaluation that underlie speech perception. The MMN component, which is most often identified at fronto-central electrode locations and peaks approximately 100–250 ms from stimulus onset, reflects early and automatic change detection in the context of a passive auditory oddball paradigm [7]. The N2b and P3 components are often elicited together in similar oddball contexts, but only when the stimuli are consciously attended to [8,9]. The N2b component is an increased negativity that follows MMN in time and peaks at about 200–300 ms from stimulus onset. It is most typically identified in centro-parietal sites and indexes attentional deviance detection [10,11]. The P3 component can be differentiated into two subcomponents: P3a originates from stimulus-driven fronto-central activity, peaking at about 300–400 ms from stimulus onset, and P3b originates from temporal–parietal activity associated with the context updating of working memory representations, peaking approximately 350–450 ms from detection of the deviation [8,12]. The MMN and N2b/P3 components have been used to examine acoustic or phonological processing in speech perception at different processing stages: at early, automatic, and preattentive stages (indexed by MMN), as well as at later, active, and attentive stages (indexed by N2b/P3). The MMN and P3b components have been shown to be sensitive to phonemic changes; therefore, they can distinguish neural processing associated with phonological representations [9,13]. Studies of categorical perception of consonants and vowels have also demonstrated larger MMN and P3b amplitudes for across-category deviants than within-category deviants, showing that the components are clear indicators of categorical perception [14–18].

Compared with extensive studies on categorical perception of segmental features (consonants and vowels), electrophysiological investigations of categorical perception of Chinese lexical tones are scarce. In a previous ERP study using a passive oddball paradigm, Xi *et al.* [6] observed that across-categorical tonal deviants elicited larger MMN than within-category tonal deviants in the

0959-4965 © 2011 Wolters Kluwer Health | Lippincott Williams & Wilkins DOI: 10.1097/WNR.0b013e32834e4842

left hemisphere, which provided strong electrophysiological evidence in support of categorical perception of lexical tones in Chinese at the early, preattentive stage. In the present study, we investigated the electrophysiological correlates of categorical perception of Chinese lexical tones at later, attentive stages of processing. We presented stimuli in an active oddball paradigm, similar to Xi *et al.*'s [6] MMN study, except that the participants were asked to press a button to indicate the across-category and within-category tonal deviants. We hypothesized that the N2b/P3b components involved in active discrimination of stimuli differences should be sensitive to the deviance. More important, across-category and within-category tonal deviants should elicit different patterns of N2b/P3b responses, differentiating acoustic and phonological processing.

Methods

Participants

We tested 14 neurologically healthy volunteers (eight women; mean age 24, range 20–28) with normal hearing and minimal musical experience. All participants were native speakers of Mandarin Chinese and were right-handed. Participants gave written consent before participating in the experiment. The experiment was approved by the ethics review board at Beijing Normal University's Imaging Center for Brain Research. The ERP data of one female participant were excluded from further analyses due to insufficient number of acceptable trials as a result of strong electroencephalogram artifacts including excessive blinking.

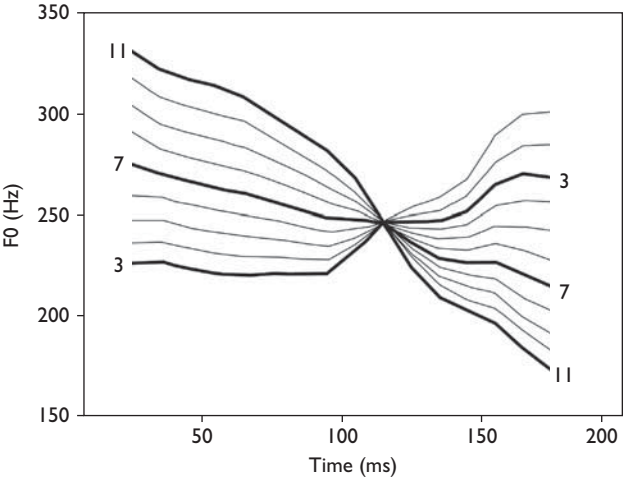
Stimuli

The stimuli were previously used in Xi *et al.*'s MMN study [6]. They were chosen from the Chinese lexical tonal continuum from the high-rising tone (tone 2) to the falling tone (tone 4) (Fig. 1), which form an across-category stimulus pair (3 and 7) and a within-category stimulus pair (7 and 11). Because the physical intervals between the across-category pair and the within-category pair were equated, the key difference between the stimulus pairs is that the former involves a change to a different phonological category, whereas the latter involves only an acoustic change. As in Xi *et al.*'s [6] MMN study, in the present study, we used stimulus 7 as the standard stimulus and stimuli 3 and 11 as the across-category and within-category deviants, respectively.

Procedure

The within-category and across-category deviants were presented pseudorandomly 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 an oddball block. The stimulus-onsetasynchrony was 800 ms. Participants were seated comfortably in an acoustically and electrically shielded chamber. They were instructed to pay attention to the stimuli presented and to press a button when they detected a different sound. Sound stimuli were presented binaurally through an insert Sony earphone (Sony Corporation, Tokyo, Japan). The right and left acoustic channels of the insert earphone were calibrated for equal and comfortable loudness (70 dB SPL) before the experiment. The experimental session lasted approximately 1 h including preparation, data acquisition, and cleanup.

Fig. 1



Tone contours of the continuum from /ba2/ to /ba4/. Continua 3, 7, and 11 are marked with thick lines.

Electrophysiological recording and analysis

Continuous electroencephalogram was recorded using a HydroCel Geodesic Sensor Net (Electrical Geodesics Incorporated Company, Eugene, Oregon, USA) consisting of 128 electrodes evenly distributed across the scalp and referenced against the vertex electrode [19]. The Geodesic Sensor Net also includes electrodes next to, and below, the eyes for recording horizontal and vertical eye-movements. The impedance of each electrode was maintained below 5 kΩ. Off-line signal processing was carried out using Netstation software (version 4.2; Electrical Geodesics Incorporated Company). The raw data were first digitally filtered using a 0.3–20 Hz bandpass filter and segmented for 1000 ms starting 100 ms before 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 to 50 μV were rejected. Two recording sites were selected for N2b statistical analyses: left central (sites C1, channels 30, 36, and 37) and right central (sites C2, channels 87, 104, and 105). Another two recording sites were selected for P3b statistical analyses: left parietal (sites P1, channels 59, 60, and 66) and right parietal (sites P2, channels 84, 85, and 91). Only the standards before the deviant were obtained for averaging and subtraction. Difference waves for N2b and P3b were obtained by subtracting the averaged

standard from the averaged deviant. The N2b peak latency for each participant was found within a 60 ms time window that was defined by the grand-average waveforms at Cz. The P3b peak latency for each participant was found within a 100 ms time window that was defined by the grand-average waveforms at Pz. The mean amplitude was calculated using a moving window technique: first, the negative peak within a specific time window (for N2b 60 ms and for P3b 100 ms) was found for each participant and then the values of a time window that extended ± 60 ms surrounding the N2b peak and ± 100 surrounding the P3b peak were averaged. Statistical analysis only included those participants with at least 80 accepted deviant trials in each condition.

Results

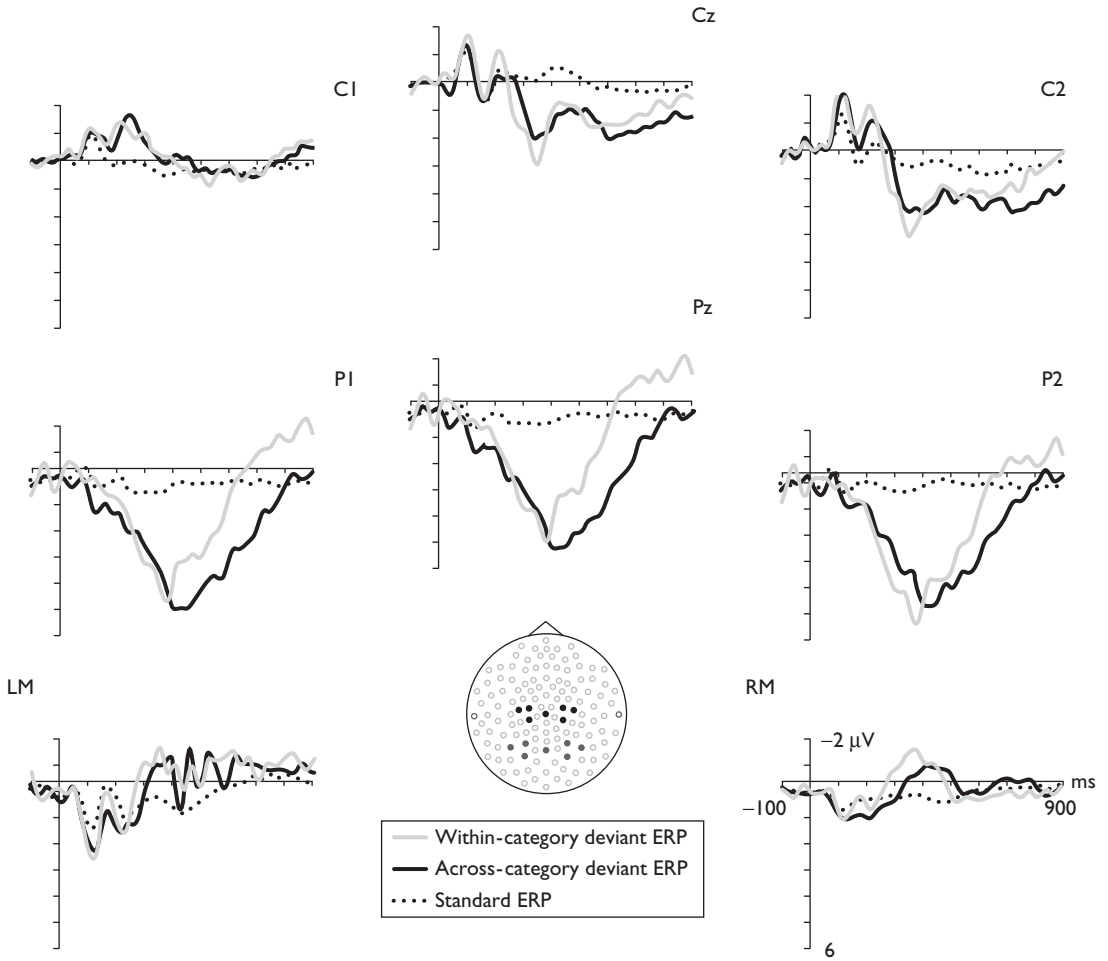
Behavioral data

The false alarm rate was very low (below 3%). The difference between the hit rate for the across-category deviants ($M = 97\%$, SD = 6%) and that for the within-category deviants ($M = 82\%$, SD = 20%) was statistically significant [$t(12) = 2.422$, $P = 0.032$], reflecting the categorical perception effect.

Event-related potential data

Compared with the standard, the within-category and across-category deviants elicited more obvious ERP responses (Fig. 2). Difference waves were obtained by subtracting deviant from standard waveforms (Fig. 3). Four separate two-way (deviant type × hemisphere) repeated measures analysis of variances were conducted for N2b and P3b mean amplitudes and peak latencies. For the N2b amplitude, neither the main effect of deviant type nor the main effect of hemisphere was statistically significant ($P > 0.10$). However, there was a significant interaction between deviant type and hemisphere [$F(1,12) = 27.110$, $P < 0.001$], indicating that the amplitude of the response to across-category deviant was greater in the left

Fig. 2



Grand average waveforms elicited by the across-category deviants (*P*=10%), within-category deviants (*P*=10%), and the standards (*P*=80%) at frontal, central, and posterior scalp sites. ERP, event-related potential; LM, left mastoid; RM, right mastoid.

continua. The electrophysiological results are consistent with previous findings showing that the N2b and P3b components are often elicited when an infrequent stimulus is attentively detected from the frequently repeating stimuli, which reflects conscious discrimination of an event [8,9].

The ERP components associated with controlled processing in the attentive condition (N2b and P3b) are often elicited together in response to target detection and can act as indicators of phonological processing and categorical perception [9,17]. For example, Maiste *et al.* [17] have found that conscious classification of a change in phonemic categories is accompanied by larger N2 and P3 responses. In the present study, the across-category and within-category tonal deviants elicited different patterns of N2b and P3b responses: only the across-category deviants elicited larger N2b and P3b responses in the left hemisphere than the right hemisphere, whereas the within-category deviants elicited similar N2b and P3b responses in both hemispheres. It is obvious that these results reflected acoustic and phonological processing involved in attentional detection of within-category and across-category deviance, respectively, which is consistent with Maiste *et al.*'s [17] findings.

A number of studies have shown that increasing task difficulty leads to greater decrements in P3b amplitude ([20]; for a review, see [21]). In the present study, the main effects of deviant type were not significant for the N2b and P3b amplitudes, although it was more difficult for the participants to detect the within-category deviants than the across-category deviants, which may be due to the relative ease in detecting the within-category deviants (reflected by the high discrimination rate). If the within-category difference had been more difficult to hear, the P3b amplitude may have decreased further.

The present set of results, together with the results of Xi *et al.*'s [6] MMN study, suggest that the ERP components including MMN, N2b, and P3b reflect categorical perception of Chinese lexical tones, but they play roles in different stages and cognitive processes. Given the low spatial resolution of ERP, more fine-grained analysis of the neural correlates of categorical perception may be pursued in the future with methods such as magnetoencephalography, so that we can more clearly identify the temporal dynamics of cortical activation in different brain regions involved in different processes of categorical perception of lexical tones in Chinese and other tonal languages.

Conclusion

In summary, this study shows that the N2b and P3b responses elicited by the across-category tonal deviants differed from the within-category deviants in amplitude and scalp topography, reflecting conscious phonological processing of Chinese lexical tones. The results provide electrophysiological correlates of categorical perception of Chinese lexical tone at later, attentive stages, which complements the previous MMN study.

Acknowledgements

The research was supported by a grant from the Humanities and Social Sciences Foundation (Projects for Young Scholars) of the Chinese Ministry of Education (10YJCZH223) and a grant from the China Postdoctoral Science Foundation (20100480509).

Conflicts of interest

There are no conflicts of interest.

References

1 Fry DB, Abramson AS, Eimas PD, Liberman AM. The identification and discrimination of synthetic vowels. *Lang Speech* 1962; **5**:171–189.
2 Liberman AM, Harris KS, Hoffman HS, Griffith BC. The discrimination of speech sounds within and across phonemic boundaries. *J Exp Psychol* 1957; **54**:358–368.
3 Lisker L. Is it VOT or a first-formant transition detector? *J Acoust Soc Am* 1975; **57**:1547–1551.
4 Hallé PA, Chang YC, Best CT. Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *J Phon* 2004; **32**:291–453.
5 Xu Y, Gandour J, Francis A. Effects of language experience and stimulus complexity on categorical perception of pitch direction. *J Acoust Soc Am* 2006; **120**:1063–1074.
6 Xi J, Zhang L, Shu H, Zhang Y, Li P. Categorical perception of lexical tones in Chinese revealed by mismatch negativity (MMN). *Neuroscience* 2010; **170**:223–231.
7 Näätänen R, Paavilainen P, Rinne T, Alho K. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin Neurophysiol* 2007; **118**:2544–2590.
8 Näätänen R. The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive functions. *Behav Brain Sci* 1990; **13**:201–288.
9 Sussman E, Kujala T, Halmetoja J, Lyytinen H, Alku P, Näätänen R. Automatic and controlled processing of acoustic and phonetic contrasts. *Hear Res* 2004; **190**:128–140.
10 Novak GP, Ritter W, Vaughan HG, Wiznitzer ML. Differentiation of negative event-related potentials in an auditory discrimination task. *Electroencephalogr Clin Neurophysiol* 1990; **75**:255–275.
11 Perrault N, Picton T. Event-related potentials recorded from the scalp and nasopharynx. I. N1 and P2. *Electroencephalogr Clin Neurophysiol* 1984; **59**:177–194.
12 Polich J. Updating P3: an integral theory of P3a and P3b. *Clin Neurophysiol* 2007; **118**:2128–2148.
13 Hisagi M, Shafer VL, Strange W, Sussman ES. Perception of a Japanese vowel length contrast by Japanese and American English listeners: behavioral and electrophysiological measures. *Brain Res* 2010; **11**:89–105.
14 Dehaene-Lambertz G. Electrophysiological correlates of categorical phoneme perception in adults. *Neuroreport* 1997; **8**:919–924.
15 Frenck-Mestre C, Meunier C, Espesser R, Daffner K, Holcomb P. Perceiving nonnative vowels: the effect of context on perception as evidenced by event-related brain potentials. *J Speech Lang Hear Res* 2005; **48**:1496–1510.
16 Kasai K, Yamada H, Kamio S, Nakagome K, Iwanami A, Fukuda M, *et al.* Brain lateralization for mismatch response to across- and within-category change of vowels. *Neuroreport* 2001; **12**:2467–2471.
17 Maiste AC, Wiens AS, Hunt MJ, Scherg M, Picton TW. Event-related potentials and the categorical perception of speech sounds. *Ear Hear* 1995; **1**:68–90.
18 Sharma A, Dorman F. Cortical auditory evoked potential correlates of categorical perception of voice-onset time. *J Acoust Soc Am* 1999; **106**:1078–1083.
19 Tucker DM. Spatial sampling of head electrical fields: the geodesic sensor net. *Electroencephalogr Clin Neurophysiol* 1993; **87**:154–163.
20 Scheffers MK, Johnson R Jr. Recognition memory and search for attended letters: an event-related potential study. *J Psychophysiol* 1994; **9**:328–347.
21 Kok A. On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology* 2001; **38**:557–577.