PROCESSING OF SPEECH AND NON-SPEECH TONAL INFORMATION BY NATIVE AND NONNATIVE TONE LANGUAGE SPEAKERS: AN EVENT-RELATED ELECTROPHYSIOLOGICAL STUDY

Yang Zhang[1], Angela Cooper[2], and Yue Wang[2]
[1]Dept. of Speech-Language-Hearing Sciences, University of Minnesota, Minneapolis, MN, 55455, USA
zhang470@umn.edu
[2]Dept. of Linguistics, Simon Fraser University, Burnaby, BC V5A 1S6, Canada; yuew@sfu.ca

INTRODUCTION

One long-deliberated question in speech perception is the extent to which it involves language-specific mechanisms. Some research findings support the independence of speech and nonspeech processing (*1-3*). Others indicate general auditory mechanisms shared by speech and nonspeech processing (*4-7*). Lexical tones present an interesting case at the suprasegmental level to study brain organization of auditory and linguistic processing as a function of learning experience. Imaging data for lexical tones have shown right-hemisphere dominance for auditory processing and left-hemisphere dominance for linguistic processing (*8-10*). However, the dynamic interplay between the two processes and the temporal courses have not been fully addressed (*6*). The present investigation employed high-density ERPs to examine two basic issues: (a) whether auditory processing of pitch patterns fundamentally differs from lexical tone processing, and (b) whether there is automatic transfer of learning for tonal pattern processing from speech to nonspeech in tone language speakers but not in non-tone language users.

METHODS

The participants included 14 native Mandarin Chinese listeners, and 16 native English listeners with no tone language experience. They were all right-handed healthy young adults. Five participants were excluded (2 native and 3 nonnative) due to incomplete data or excessive noise in EEG recording.

The stimuli for the speech condition included two Mandarin monosyllables (*ju, ci*) produced with two tones (rising, falling) by a female native Mandarin speaker. For the nonspeech condition, the same speaker hummed these two syllables with rising and falling tones.

A modified oddball paradigm was employed *(3)*. The oddball sequences contained both syllables (*ju, ci*) in alternation. Each condition consisted of four different sequences, varying in syllable order (*ju-ci* or *ci-ju*) and deviant syllable (*ju* or *ci*), producing a total of eight different blocks. For each condition, 800 stimuli were presented (20% deviant, 80% standard). A 128-electrode light mesh (Geodesic net) was applied, with the vertex (Cz) electrode used as a reference with the ground at the nasion. The stimuli were presented via speakers.

The raw EEG data for each participant were analyzed and averaged with common average reference in BESA software (Brain Electrical Source Analysis) with a passband of 0.5~40 Hz. Two windows, 100~300 ms and 300~ 600 ms, were selected after inspection of the grand mean data to search mismatch negativity (MMN) peak in the subtracted waveform (deviant – standard) (*11*). MMN quantification used the mean of sample values in a window of 40 ms centered around the MMN peak *(6)*. Global field power for the subtracted waveforms was calculated for each participant and each stimulus condition. Furthermore, point-to-point t-tests were performed to see the temporal evolution of significant differences for key factors of interest (*12*). Topographical maps were calculated for selected peak points. Finally, repeated measures MANOVA was performed for the MMNs at F3 site (electrodes 19, 23, 24) and F4 site (electrodes 3, 4, 124) *(6, 11)*.

RESULTS

In the speech condition (Fig. 1), natives showed one early MMN (corresponding to the consonant portion of the syllable) and one late MMN (corresponding to the vowel portion) in both hemispheres. Nonnatives did not show point-to-point significances in the left site, and the right hemisphere site had a later MMN response than the first MMN of the native group. MANOVA results confirmed the significant MMN differences between the native and nonnative listeners for the early MMN $[F(1,23) = 5.02, p < 0.05]$.

In the nonspeech condition (Fig. 2), both native and nonnative groups showed an early MMN response in both hemispheres followed by P3a response. The native group showed earlier and larger MMN and P3a than the nonnative speakers ($p < 0.05$; MANOVA results).

To avoid channel-selection bias, GFPs *(12* for all the 129 electrodes*)* were calculated for the MMN waveforms (Fig. 3). The results confirmed the MMN results at F3 and F4 sites, showing that native listeners had earlier and stronger MMN responses for both speech and nonspeech conditions. Interestingly, the nonspeech condition elicited larger MMN responses than the speech condition in the native listeners ($p < 0.05$). The MMNs for the speech and nonspeech conditions were comparable for the nonnative listeners.

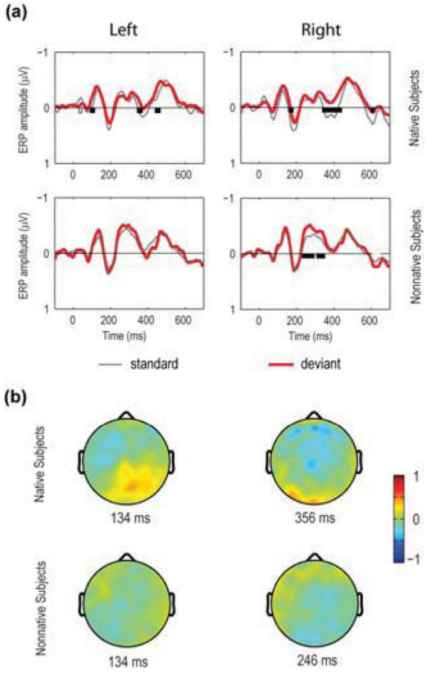

Fig. 1. ERPs for the speech stimuli in the two participant groups. (a) Grand average ERPs for the standard and deviant stimuli from F3 (electrodes 19, 23, 24) and F4 sites (electrodes 3, 4, 124). Point-to-point significant differences were indicated by the black/white bars on x-axis ($p < 0.01$). (b) Topographical maps for the peaks in MMN waveforms.

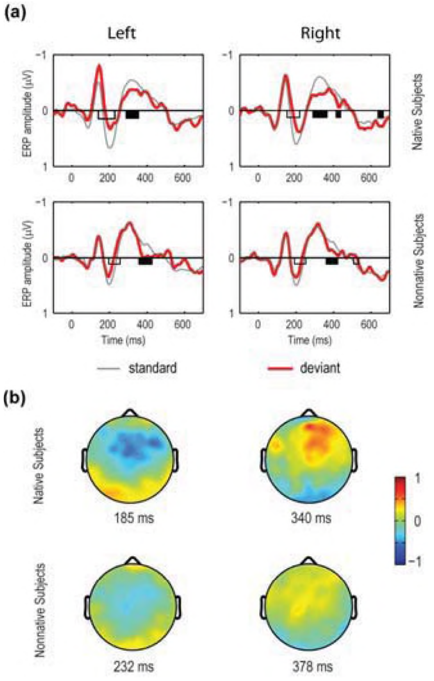

Fig. 2. ERPs for the nonspeech stimuli in the two participant groups. (a) Grand average ERPs for the standard and deviant stimuli from F3 and F4 sites. Point-to-point significant differences are indicated by the black/white bars on x-axis ($p < 0.01$). (b) Topographical maps for the peaks in MMN waveforms.

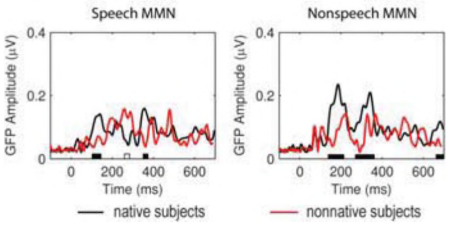

Fig. 3. Global field power of MMN waveforms in all the 129 electrodes for the speech and nonspeech conditions. Significant point-to-point differences between the native and nonnative groups are indicated by the black/white bars on x-axis ($p < 0.01$).

DISCUSSION

Despite different patterns of MMN elicitation in the speech and nonspeech conditions, the MMNs were present for both native and nonnative listeners, but the magnitude was much smaller in the nonnative listeners for the speech stimuli. These data are consistent with previous studies *(3, 6, 7,11)*. Unlike a previous report (3), the native listeners showed earlier MMN responses than nonnative listeners in both speech and nonspeech conditions. Furthermore, the MMNs for nonspeech were larger than those of the speech condition only in native listeners, suggesting that the effects of language experience may automatically transfer to non-speech tonal pattern extraction *(6)*. The results indicate a high degree of interconnectivity for speech and nonspeech in terms of neural sensitivity to pitch categories.

ACKNOWLEDGMENTS

This research is supported by a Discovery Grant to YW from the Natural Sciences and Engineering Research Council of Canada (312457-2006). YZ is supported by two Brain Imaging Research Awards from UMN to work on the project. The authors acknowledge assistance from Mathieu Dovan, Anish Jain, Amy Lee, David Potter, Rebecca Simms, and Marcus Watson.

REFERENCES

1. A. M. Liberman, D. H. Whalen, *Trends Cogn. Sci.* **4**, 187 (2000).
2. K. Miyawaki *et al.*, *P&P* **18**, 331 (1975).
3. Y. Zhang, P. K. Kuhl, T. Imada, M. Kotani, Y. Tohkura, *Neuroimage* **26**, 703 (Jul 1, 2005).
4. A. L. Francis, V. Ciocca, B. K. Ng, *Percept. Psychophys.* **65**, 1029 (Oct, 2003).
5. L. L. Holt, A. J. Lotto, R. L. Diehl, *J. Acoust. Soc. Am.* **116**, 1763 (Sep, 2004).
6. J. Xi, L. J. Zhang, H. Shu, Y. Zhang, P. Li, *Neuroscience,* (In press).
7. E. Sussman *et al.*, *Hear. Res.* **190**, 128 (Apr, 2004).
8. R. J. Zatorre, J. T. Gandour, *Philosophical Transactions of the Royal Society B: Biological Sciences* **363**, 1087 (2008).
9. Y. Wang, J. A. Sereno, A. Jongman, J. Hirsch, *J. Cogn. Neurosci.* **15**, 1019 (Oct 1, 2003).
10. P. C. Wong *et al.*, *Cereb. Cortex* **18**, 828 (Apr, 2008).
11. E. Kaan, R. Wayland, M. Bao, C. M. Barkley, *Brain Res.* **1148**, 113 (May 7, 2007).
12. Y. Zhang *et al.*, *Dev. Sci,* (In press).