Neuropsychologia 51 (2013) 2238–2244

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Hemispheric lateralization for early auditory processing of lexical tones: Dependence on pitch level and pitch contour

Xiao-Dong Wang [a,b], Ming Wang [a], Lin Chen [a,c,*]

[a] CAS Key Laboratory of Brain Function and Diseases, School of Life Sciences, University of Science and Technology of China, Hefei 230027, China
[b] School of Humanities and Social Sciences, Nanyang Technological University, 637332 Singapore, Singapore
[c] Auditory Research Laboratory, University of Science and Technology of China, Hefei 230027, China

ARTICLE INFO

Article history:
Received 9 October 2012
Received in revised form
20 July 2013
Accepted 24 July 2013
Available online 1 August 2013

Keywords:
Lexical tone
Pitch level
Pitch contour
Mandarin Chinese
Mismatch negativity
Hemispheric lateralization

ABSTRACT

In Mandarin Chinese, a tonal language, pitch level and pitch contour are two dimensions of lexical tones according to their acoustic features (i.e., pitch patterns). A change in pitch level features a step change whereas that in pitch contour features a continuous variation in voice pitch. Currently, relatively little is known about the hemispheric lateralization for the processing of each dimension. To address this issue, we made whole-head electrical recordings of mismatch negativity in native Chinese speakers in response to the contrast of Chinese lexical tones in each dimension. We found that pre-attentive auditory processing of pitch level was obviously lateralized to the right hemisphere whereas there is a tendency for that of pitch contour to be lateralized to the left. We also found that the brain responded faster to pitch level than to pitch contour at a pre-attentive stage. These results indicate that the hemispheric lateralization for early auditory processing of lexical tones depends on the pitch level and pitch contour, and suggest an underlying inter-hemispheric interactive mechanism for the processing.

1. Introduction

Pitch is the basic and most general acoustic element of a spoken language. Tonal languages, such as Mandarin Chinese, use pitch for the determination of word meaning (Howie, 1976). More than 60% of the languages in the world are tonal (Yip, 2002). In Mandarin Chinese, there are four lexical tones which can be described phonetically as level, rising, dipping, and falling (tones 1 to 4, respectively) (Howie, 1976). The average fundamental frequency (*F*0) contours of time-normalized Mandarin Chinese tones are shown in Fig. 1A (Xu, 1997). In recent years, auditory cognitive processing of Chinese lexical tones has gained increasingly more attention in the field (Gu et al., 2012; Lee & Lee, 2010; Lee, Dutton, & Ram, 2010; Lee, Tao, & Bond, 2010; Luo et al., 2006; Ren, Yang, & Li, 2009; Wang, Gu, He, Chen, & Chen, 2012; Wang, Jongman, & Sereno, 2001; Wong, 2002).

According to acoustic perceptual dimensions, the pitch pattern of a Mandarin Chinese lexical tone comprises two separate features (Chandrasekaran, Gandour, & Krishnan, 2007; Gandour, 1983): The first one is the relative height of the pitch which is reflective of the pitch level and the second one is the shape and direction of pitch (level, rising, dipping or falling) which is reflective of the pitch contour (Fig. 1B). Both of these dimensions are critical for the recognition of a lexical tone. The change of a pitch level is referred to as a step change in *F*0 from the sound onset (similar to music scales) whereas the change of a pitch contour is referred to as a consecutive change in *F*0 within tens to hundreds of milliseconds. In Mandarin Chinese, a change in pitch level is a within category change whereas that in pitch contour is an across category change in lexical tones. For the same lexical tone, different speakers pronounce variable pitch levels and similar pitch contours.

Generally speaking, music is believed to be preferentially processed in the right hemisphere of the brain and speech preferentially processed in the left hemisphere. Lexical tones are linguistic pitches that sound melodic, but define the meanings of words. There has long been an issue regarding the hemispheric lateralization of lexical tones. Current literature mainly provides data regarding the hemispheric lateralization for lexical tone processing in general, including the overall and simultaneous processing of pitch level and pitch contour. However, relatively little is known about how each dimension contributes to the hemispheric lateralization for the central auditory processing of lexical tones. For instance, the neural mechanisms of overall lexical tone processing have been intensively investigated using neuroimaging techniques such as functional MRI (fMRI) or positron emission tomography (PET) (Gandour et al., 2002; Klein, Zatorre, Milner, & Zhao, 2001; Wong, Parsons, Martinez, & Diehl, 2004). These neuroimaging studies demonstrate that in tonal languages, a lexical tone that carries semantic information is preferentially processed in the left hemisphere of native speakers, as revealed

* Corresponding author at: CAS Key Laboratory of Brain Function and Diseases, School of Life Sciences, University of Science and Technology of China, Hefei 230027, China. Tel.: +86 551 6360 7623.
E-mail address: linchen@ustc.edu.cn (L. Chen).

0028-3932/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.neuropsychologia.2013.07.015

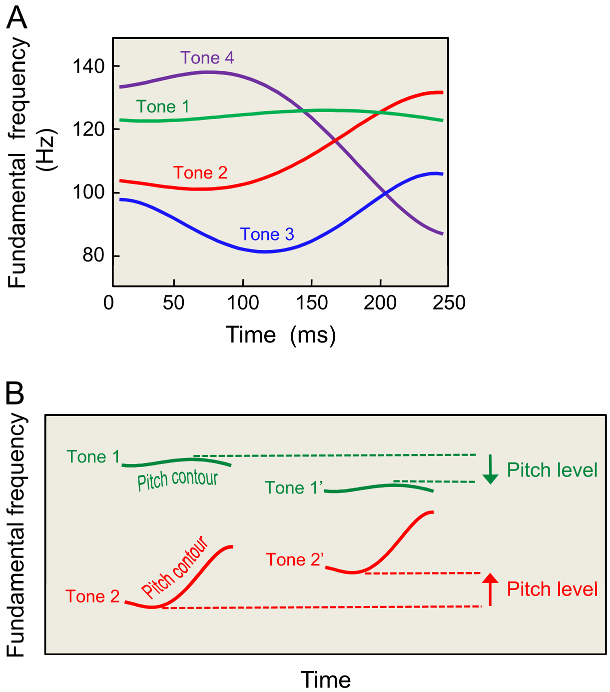

Fig. 1. (A) Average fundamental frequency contours of time-normalized Mandarin Chinese tonal stimuli adapted from Xu (1997). Tone 1 (Flat tone), Tone 2 (Rising tone), Tone 3 (Dipping tone) and Tone 4 (Falling tone) denote the four Mandarin tones. Of those four tones, Tone 1 and Tone 2 were used in this experiment. (B) Schematic illustrating the concept of pitch contour and pitch level. Tone 1 and Tone 2 have different pitch contours because the fundamental frequency of Tone 1 remains relatively constant over time whereas that of Tone 2 rises over time. Similarly, Tone 1′ and Tone 2′ have different contours, too. Although Tone 1 and Tone 1′ have the same pitch contour, the pitch level of Tone 1′ is lower than that of Tone 1. Similarly, the pitch level of Tone 2′ is higher than that of Tone 2 although their pitch contours are the same.

by fMRI or PET, which measure the temporally aggregated neural events, including those at an attentive stage of auditory processing. Recently, a number of electrophysiological studies have been conducted investigating the neural mechanisms in tonal processing at a pre-attentive stage (Chandrasekaran, Krishnan, & Gandour, 2007a, 2007b, 2009a; Gu et al., 2012; Kaan, Barkley, Bao, & Wayland, 2008; Luo et al., 2006). Converging evidence from EEG studies demonstrates that the auditory processing of lexical tones, at a pre-attentive stage, is lateralized to the right hemisphere (Luo et al., 2006; Ren et al., 2009), suggesting that the auditory processing of lexical tones is shaped mainly by acoustic properties at a pre-attentive processing stage (Chandrasekaran, Krishnan, & Gandour, 2009b). A recent study investigating the pre-attentive processing of pitch contour and pitch height in Cantonese lexical tones, demonstrated that the size and latency of the mismatch negativity (MMN) were sensitive to the size of the pitch level change, while the latency of P3a captured the presence of pitch contour change (Tsang, Jia, Huang, & Chen, 2011). This indicates that the pitch contour and pitch height are two important dimensions in early auditory processing of lexical tones. However, the hemispheric lateralization for the processing of pitch contour and pitch height has not been addressed yet.

In the present study, we focused on the hemispheric lateralization for, and the time course of, the early auditory processing of each of the two dimensions (pitch level and pitch contour) by measuring the MMN in response to separate violations of the two dimensions of a lexical tone. The MMN is an index of automatic auditory change detection and that happens regardless of whether the subject is paying attention (Näätänen, Paavilainen, Rinne, & Alho, 2007; Picton, Alain, Otten, Ritter, & Achim, 2000). It has also been suggested that MMN is a marker for an individual′s state dependent sensitivity to emotionally relevant changes such as voice pitch (Schirmer & Escoffier, 2010). The MMN is an efficient tool for investigating the brain′s processing of tonal languages since changes in the pitch of repetitive sounds elicit the MMN (Chandrasekaran et al., 2009a, 2007; Kaan et al., 2008; Luo et al., 2006). In the current study, synthesized stimuli were applied to strictly control the acoustic attributes of the lexical tones. A multiple standard stimuli design in the block of pitch contour contrast was used (Fig. 2C). This enabled us to investigate the central auditory processing of pitch contour dimension without the contamination of the pitch level dimension and other acoustic features such as a formant change. The stimuli used in this study were the Chinese vowel /a/ with the flat tone, presented in seven pitch levels, and the vowel /a/ with the rising tone. We applied the synthesized stimuli to ensure that, except for the $F0$, all other features of the stimuli were kept constant, such as intensity, first formant (F1), second formant (F2) and duration (Fig. 2). We sought to determine the hemispheric lateralization for, and the time course of, the early auditory processing of the two dimensions since they differed in both acoustic properties and linguistics.

2. Materials and methods

2.1. Subjects

Twelve native speakers of Mandarin Chinese (6 males and 6 females) aged 21–27 years old, from the University of Science and Technology of China, participated in the study. Subjects had no musical training and no history of neurological or psychiatric impairment. All subjects were right-handed according to an assessment using the Chinese version of the Edinburgh Handedness Inventory (Oldfield, 1971) and all subjects reported normal hearing. The protocols and experimental procedures deployed in the present study were reviewed and approved by the Biomedical Research Ethics Committee of the University of Science and Technology of China. All subjects provided written informed consent.

2.2. Stimuli

The synthesized vowels /a/, with seven flat and one rising pitch contour, were used as stimuli in the experiment. The $F0$ contours of the stimuli were adapted from Xu (1997). Speech waveforms were generated by a high quality speech synthesizer STRAIGHT (Speech Transformation and Representation using Adaptive Interpolation of weiGHTed spectrum) (Kawahara, Masuda-Katsuse & de Cheveigné, 1999), using a source-filter model. A periodic excitation sequence was used to simulate the vocal tract filter to produce vowels. Line Spectral Pairs (LSP) were utilized to model the vocal tract filter (McLoughlin, 2008), and the LSP parameters were extracted from a speaker′s utterance using Linear Predictive Coding (LPC) (Wouter & Macon, 2001) frame by frame. In order to keep the formant structure stable, we selected representative frame LSP parameters from the extracted LSPs for all of the frames of the vowels during the synthesis process, so that the only difference between the generated vowels came from the $F0$. Time domain short waveform was generated and overlap-added at each pitch-synchronous point to generate speech waveforms using Time-Domain Pitch-Synchronous Overlap-Add techniques (TD-PSOLA) (Moulines & Verhelst, 1995). The vowel /a/ with the flat pitch contour was generated in seven $F0$ levels, 94.6 Hz, 101.6 Hz, 108.6 Hz, 115.6 Hz, 122.6 Hz, 129.6 Hz and 136.6 Hz of the onset $F0$ value. Vowel /a/ with the rising pitch contour was generated with its $F0$ value rising continuously, with time, from 102 Hz to 130 Hz. All stimuli were 250 ms long with rise and fall time of 5 ms each. Except for the $F0$ dimensions, all other parameters were standard and identical for all stimuli. The F1 and F2 were synthesized as 830 Hz and 1300 Hz, respectively.

2.3. Procedure

MMN was recorded using an auditory odd-ball paradigm (Näätänen & Escera, 2000). During the experiment, subjects were instructed to ignore the auditory stimulus and watch a silent movie. The stimuli were delivered binaurally at 60 dB above each subject′s threshold through headphones (TDH-39; Telephonics, Farmingdale, NY, USA) in an electrically shielded, soundproofed room with a stimulus onset asynchrony of 800 ms. There was a total of 2100 trials for both the pitch level contrast condition and pitch contour contrast condition. In each condition, the standard stimuli were presented in 87.5% of the trials and the deviant stimuli were presented in 12.5%. For the pitch level contrast condition, there were (flat /a/-level 2, flat /a/-level 6), (flat /a/-level 6, flat /a/-level 2) oddball blocks (Fig. 2C, right panel). For the pitch contour contrast condition, the standard