Auditory Preattentive Processing of Speech Prosody in Lexical Tone: a Functional Imaging with Low Resolution Brain Electromagnetic Topography (LORETA)

Wichian Sittiprapaporn, Chittin Chindaduangratn, and Naiphinich Kotchabhakdi

Clinical and Research Electroneurophysiology Laboratory  
Neuro-Behavioural Biology Center  
Institute of Science and Technology for Research and Development  
Mahidol University  
wichian_s@thailand.com

Abstract

The present study investigated the preattentive perception of speech prosody changes in lexical tone as revealed by the low-resolution electromagnetic tomography (LORETA). The mismatch negativity (MMN), elicited by consonant-vowel (CV) syllable change of both native and non-native speech sounds in tonal languages was assessed by measuring the electric sources of MMN in both hemispheres. Our results suggest that the left hemisphere was predominant in the perception of native speech sounds, whereas the non-native speech sound was dominated by the right hemisphere, which may be explained by the specialization in processing the prosodic and emotional components of speech formed in this hemisphere.

1. Introduction

It is widely accepted that the left cerebral hemisphere is considered to be dominant in the perception and production of speech in most humans, while the right one is known to be specialized in processing the prosodic and emotional components of speech [1]. Attempts to find in the human brain a mechanism specific for the processing of speech and distinct from other sound processing structures have been conducted. One of the most popular methods of noninvasive assessment of speech lateralization in the brain has been dichotic listening. Data obtained in dichotic listening experiments with speech sounds have shown in most people a perceptual right-ear advantage (REA) which is considered to be an indicator of left-hemisphere dominance in speech processing. A recent study has evaluated the differences between dichotic listening and mismatch negativity (MMN) as measures of speech lateralization in the human brain [2]. The MMN component of the auditory event-related brain potential (ERP) is elicited by an infrequently change in a repetitive sound sequence and can be used to investigate the neural processing of speech and language [2-7] because it is considered to be a unique indicator of automatic cerebral processing of acoustic stimuli. It has been found that both dichotic listening and MMN methods indicated slightly stronger left-hemisphere dominance in speech processes [2]. In contrast to these approaches, we used a tomographic method that is also Talairach based, that is, low resolution electromagnetic tomography (LORETA), which is now widely used in the analysis of evoked potentials and magnetoencephalography (MEG) [8-10]. The majority of this study considered the MMN responses and applied LORETA to grand mean auditory MMN response, resulting in images produced by electroencephalography (EEG). Thus, the aim of the present study was to localize the brain activity underlying the MMNs response elicited during the preattentive processing of speech prosody in lexical tones. Anatomical location of active sources is revealed by projection of LORETA solutions into a grey-matter structures model and evaluation of the effects of native and non-native speech sound in tonal languages on these source distributions.

2. Materials and methods

2.1. Subjects

Nine healthy right-handed volunteers (native Thai speakers aged 18-35 years) participated in the study. All subjects had normal hearing sensitivity and gave their written informed consent before participation in the study.

2.2. Language

Thai has five contrastive lexical tones [11], traditionally labeled mid ( ), low ( ), falling ( ), high ( ), and rising ( ); for example, /kâa/ ‘stuck’ /kâa/ ‘galangal’ /kâa/ ‘kill’ /kâa/ ‘trade’ /kâa/ ‘leg’. The midtone can be described phonologically as mid-level with a final drop, low tone as low-falling, falling tone as high-falling, high tone as high-rising, and rising tone as low-rising. The primary acoustic correlated of Thai tones is voice fundamental frequency [12].

2.3. Stimuli

Two stimuli of speech sounds, each consisting of a CV syllable were prepared to elicit MMN in response to native and non-native speech sound changes (standard, Thai word with falling tone /kâa/; deviant, Chinese morpheme with a similar tone /ta/ and vice versa) All stimuli with 500 ms in duration were digitally edited to have an equal peak energy level in dB SPL using the Cool Edit Pro v. 2.0 (Syntrillium Software Corporation). The sounds were presented binaurally via headphones (Telephonic TDH-39-P) at a comfortable listening level of ~85 dB (determined using a Bruel & Kjaer 2230 sound level meter).

2.4. Acoustic stimulation

(1) The Thai word /kâa/ deviant (10%) was presented among the Chinese morpheme /ta/ standard (90%) and (2) the Chinese morpheme /ta/ deviant (10%) was presented among the Thai word /kâa/ standard (90%) in random order (except that each deviant stimulus was preceded by at least one standard stimulus). The inter-stimulus interval (ISI) was 1.25 second (offset-onset).

2.5. Electroencephalographic recording and processing

Subjects were seated in an electrically and acoustically shielded chamber, instructed to read a book of their own choices and to ignore any auditory signals. During the auditory stimulation, electric activity of the subjects’ brain was continuously recorded with 21 active electrodes positioned according to the 10-20
International System of Electro-cap and referred to linked mastoids with an electrode between Fz and Fpz connected to ground. A biologic Brain Atlas system amplified (Band-pass 0.01-100 Hz); analog-digital converted (128 samples/s/channel) and stored the data. Averaged responses were digitally filtered with a bandpass of 1-30 Hz. ERPs were then obtained by averaging epoch, which started 100 ms before the stimulus onset and ended 900 ms thereafter, the –100 – 0 ms interval was used as a baseline. Epochs with voltage variation exceeding ±100 µV at any EEG channels were rejected from further analysis. MMN was obtained by subtracting the response to the standard from that to the deviant stimulus. The averaged MMN responses contained at least 125 accepted deviant trials.

2.6. Spatial analysis
ERP voltages were transformed into reference-independent values by re-computing the voltages against average reference [13]. To quantify the amount of “spatial relief” or hilliness of the topographic field, the scalp electric potential power or global field power (GFP) was calculated. The amplitude variability across the 21 electrodes was calculated by averaging the data, time point by time point, in the difference waveforms. The point of maximal GFP in the time segment between 144 and 204 ms after stimulus onset was detected. LORETA was then applied to estimate the current source density distribution in the brain, which contributes to the electrical scalp field [14]. LORETA computes current density for each cortical voxel as the linear, weighted sum (power value) of the scalp electric potentials, applying the constraint of maximal smoothness to the ‘inverse solution’ without need for a priori-knowledge of the number of sources, using a 7-mm spatial resolution for the digitized Talairach human brain atlas provided by the Neurological Institute, Montreal, and computing 2394 voxels in cortical gray matter and hippocampus.

3. Results
There was MMN and it was significant in both conditions. The native (/kʰæ/ deviant) condition yielded higher MMN amplitudes than the non-native (/tʰæ/ deviant) condition (respectively mean amplitude -2.41 µV; GFP = 1.89 and -0.98 µV; GFP = 1.32, respectively). LORETA revealed that MMNs response to native speech sound was associated with activity in the left superior frontal gyrus (the maximum of the t = -2.30, Brodmann area (BA) 10; x, y, z = -10, 66, 22), while there was associated with more activity in the right anterior cingulate gyrus (t = -3.47, BA 36; x, y, z = 18, -32, -13) in non-native speech sound (Fig. 1).
representations leads to the enhancement of the cortical mismatch response [4]. Therefore, the significant of electric source of MMNs response in prelexical perceptual process can be accounted for by the subjects’ previous language experience and the degree to which the task is learned or automatic. In other word, the electric source of MMNs response to which the native speech sound is engaged in these auditory discrimination tasks reflects different cognitive strategies that native Thai listeners’ employ depending on their language experiences. This implies that the brain might be capable of automatic lexical classification of the incoming speech signals already at very early stages of speech processing [2, 17]. This also suggests that auditory parameters of the speech signal are not only encoded in higher cortical areas by their complex acoustic properties, but also by their linguistic relevance in particular language [18].

5. Conclusion

The present study was conducted to determine the perception of speech prosody in lexical tones. The present study elicited the difference in cerebral lateralization between the auditory preattentive perception of native and non-native-speech sounds in tonal languages. This revealed that spoken words from the subjects’ native language elicited electric sources of mismatch responses, indicating a source contribution in the left auditory cortex and in the right auditory cortex for non-native speech sounds. These responses can be activated in the absence of active attention to the auditory input and are probably available at the early stages of cerebral speech processing.

6. Acknowledgement

This study was supported by the University Development Communities (UDC) grant from Mahidol University, Thailand.

7. References


