Preattentive discrimination of across-category and within-category change in consonant–vowel syllable

Wichian Sittiprapaporn,1CA Mari Tervaniemi,2,3 Chittin Chindaduangratn1 and Naiphinich Kotchabhakdi1

1Neuro-Behavioural Biology Center, Institute of Science and Technology for Research and Development, Mahidol University, Salaya, Nakhonpathom, Thailand; 2Cognitive Brain Research Unit, Department of Psychology, University of Helsinki; 3Helsinki Brain Research Centre, Helsinki, Finland

CA Corresponding Author: stwst@mahidol.ac.th

Received 26 May 2005; accepted 15 June 2005

Event-related potentials to infrequently presented spoken deviant syllables /pi/ and /po/ among repetitive standard /pɔ/ syllables were recorded in Thai study participants who ignored these stimuli while reading books of their choices. The vowel across-category and within-category changes elicited a change-specific mismatch negativity response. The across-category and within-category change discrimination of vowels in consonant–vowel syllable was also assessed using the low-resolution electromagnetic tomography. The results of low-resolution electromagnetic tomography mismatch negativity generator analysis suggest that the within-category change perception of vowels is analyzed as the change in physical features of the stimuli, thus predominantly activating the right temporal cortex. In contrast, the left temporal cortex is predominantly activated in the across-category change perception of vowels, emphasizing the role of the left hemisphere in speech processing already at a preattentive processing level also in consonant–vowel syllables. The results support the hypothesis that a part of the superior temporal gyrus contains neurons specialized for speech perception. NeuroReport 16:1513–1518 © 2005 Lippincott Williams & Wilkins.

Key words: Categorical perception; Event-related potential; Low-resolution electromagnetic tomography; Mismatch negativity; Speech; Vowel

INTRODUCTION

Mismatch negativity (MMN), an index of preattentive processing of speech sounds, is an event-related potential (ERP) component elicited by deviant stimuli sequences of repetitive auditory stimuli [1]. The MMN component appears as a frontocentral negative wave, usually peaking between 100 and 300 ms from the onset of stimulus deviation. The MMN(m) in response to speech sounds can occur as a response to both within-category and across-category change of phonemes. The former is the change detection of physical feature(s) within the same phoneme boundary, such as changes in pitch, duration, and intensity. The latter is the change detection between independent phonemes [2]. For example, Thai language can be divided into front (i.e. /i/, /e/, /æ/), and back (i.e. /u/, /o/, /ɔ/), /a/) vowels or rounded (i.e. /u/, /o/, /ɔ/) and unrounded (i.e. /i/, /e/, /æ/, /a/) vowels. Consequently, the change detection between any combinations of these (i.e. /ɔ/ vs. /i/) is considered across-category change detection.

The MMN is traditionally thought to reflect the acoustic difference between two sounds, so that the larger the acoustic difference, the faster the behavioral responses of the listeners and the earlier and larger the MMN. However, if the deviant vowel is near the category boundary, then the familiarity of the stimuli modulates the reaction time [3]. Moreover, the reaction time results on the speech sound categorization and discrimination demonstrate that vowels were categorically coded [4,5]. This indicates that across-category discrimination performance is more accurate than within-category discrimination.

Because the vowel categorization is based on several simultaneous encoding mechanisms, numerous attempts have been made to establish the underlying cortical functions. Magnetoencephalographic (MEG) data of the auditory evoked field show that the different vowels elicited a distinct spatiotemporal pattern of cortical activity and were represented by different cortical mappings, as can be expected by their different spectral structures [6,7]. Vowels with mutually exclusive place of articulation features elicited separate centers of activation in the vicinity of the auditory cortex [7]. For example, the acoustically most dissimilar vowels /a/ and /i/ showed more distant location than the more similar vowels /e/ and /i/ [8]. The distance between sources representing /u/ and /i/ was larger than the distance between /a/ and /s/ [9]. Thus, the asymmetries in vowel perception evident in humans are likely to reflect general auditory mechanisms and can be considered indicative for a cortical map related to spectral dissimilarities of speech input [8,10].

Furthermore, the left and right hemispheres of the brain are not symmetrically involved in language processing. The human brain processes speech-relevant acoustic information in both the left and right temporal lobes especially in the superior temporal gyrus, which contains neural networks specialized for language processing [11]. Further, it is hypothesized that the two hemispheres may have two time courses for activation: (1) both hemispheres are active in...
parallel with the left hemisphere dominant for different phonemes throughout and (2) hemispheric dominance may alternate to play a specific role at different stages of linguistic processing [12]. A recent MEG study showed that both hemispheres were symmetrically activated in the preattentive across-category change perception of vowels while the within-category change of a vowel was predominantly activated in the right hemisphere. The study proposed that the dominance of the left auditory cortex in the preattentive speech processing might occur only at the level of perception of vowel across-category change [2]. However, the study utilizing only simple Japanese vowels did not address the issue of whether the functional representation of vowel stimuli in consonant–vowel syllables is specific to the level of perception of vowel across-category and within-category change.

The present study thus used low-resolution electromagnetic tomography analysis (LORETA) [13–15] to locate further multiple nondipolar sources in the across-category and within-category discrimination of vowel in consonant–vowel syllables in the MMN paradigm. Our purpose was to determine when and where in the brain different stages of different categories of vowel discrimination take place, with particular emphasis on right/left hemisphere comparisons.

**MATERIALS AND METHODS**

**Study participants:** Ten healthy right-handed (handedness assessed according to Oldfield [16]) native speakers of Thai (seven women; aged 18–35 years), with normal hearing sensitivity, gave their written, informed consent before participation in the study. The mean (±SD) age was 24.35 (±4.95) years.

**Stimuli and procedure:** Three consonant–vowel syllables of Thai were prepared: back rounded articulation /pɔi/, front unrounded: /pɛi/, and back rounded: /pɔ/. All stimuli were spoken by native female Thai speakers and digitally edited to have an equal peak energy level in decibel sound pressure level using the Cool Edit Pro v. 2.0 (Syntrillium Software Corporation, Phoenix, Arizona, USA) with 500 ms duration. The sounds were presented binaurally via headphones (Telephonic TDH-39-P) at 85 dB (determined using a Bruel and Kjaer 2230 sound level meter). All stimuli were identical at their suprasegmental (i.e. tone) unit, which was always ‘mid’ tone, thus eliminating any effect due to differences in frequency of occurrence of tones. The /pɛi/ (10%) (across-category vowel change) and /pɔi/ (10%) (within-category change vowel deviants) were presented among the /pɔ/ standard (90%) in random order (except that each deviant stimulus was preceded by at least one standard stimulus). The interstimulus interval was 1.25 s (offset–onset).

**Electroencephalographic recording:** Participants were seated in an electrically and acoustically shielded chamber, instructed to focus their attention on reading books of their own choice and to ignore any auditory signals. During the auditory stimulation, electric activity of the participants’ brain was continuously recorded with 21 active electrodes positioned according to the International 10/20 System of Electro-cap (Electro-cap International Inc., Eaton, Ohio, USA) and referred to linked mastoids. All 21 recording channels used for LORETA computation were the same as in previous studies [12,17], which corresponded to the standard electrode Talairach coordinate 10/20 system template [13]. A biologic Brain Atlas system amplified (bandpass 0.01–100 Hz), analog-digital converted (128 samples/s/channel) and stored the data. Averaged responses were digitally filtered offline with a bandpass of 1–30 Hz.

**Electroencephalographic data processing:** The recordings were filtered and carefully inspected for eye movement and muscle artifacts. ERPs were obtained by averaging epoch, which started 100 ms before the stimulus onset and ended 900 ms thereafter; the −100 to 0 ms interval was used as a baseline. Epochs with voltage variation exceeding ±100 µV at any electroencephalographic (EEG) channel were rejected from further analysis. The MMN was obtained by subtracting the response to the standard from that to the deviant stimulus. For each participant, the averaged MMN responses contained at least 125 accepted deviant trials. All responses were recalculated offline against average reference for further analysis.

**Spatial analysis:** The average MMN latency was defined as a moment of the global field power with an epoch of 40-ms time window related stable scalp-potential topography [18]. The individual momentary potential measures from 21 electrodes at the MMN latency were analyzed with LORETA to determine the MMN source loci [14]. These latencies were between 100 and 140 ms for across-category and within-category changes. LORETA calculated the current source density distribution in the brain, which contributed to the electrical scalp field, at each of 2395 voxels in the gray matter and the hippocampus of a reference brain (MNI 305, Brain Imaging Centre, Montreal Neurological Institute, Montreal, Quebec, Canada) on the basis of the linear weighted sum of the scalp electric potentials [14]. LORETA chooses the smoothest of all possible current density configurations throughout the brain volume by minimizing the total squared Laplacian of source strengths. This procedure only implicates that neighboring voxels should have a maximally similar electrical activity; no other assumptions were made. The applied version of LORETA used a three-shell spherical head model registered to the Talairach space and calculated the three-dimensional localization of the electrical sources contributing to the electrical scalp filed for all 10 participants and conditions, defining the regions of interest on the basis of local maxima of the LORETA distribution.

Stereotaxic coordinates of the voxels of the local maxima were determined within areas of significant relative change associated with the tasks. The anatomical localization of these local maxima was assessed with reference to the standard Stereotaxic atlas, and validation of this method of localization was obtained by superimposition of the statistical parametric maps (SPM) on a standard magnetic resonance image brain provided by the SPM99. Peaks located within the superior temporal gyrus were also identified by using published probability maps following a correction for the differences in the coordinate systems between the Talairach and
Tournoux atlas and the Stereotaxic space employed by SPM99.

**Statistical analysis:** The statistical significance of MMN was tested with the one-sample $t$-test. This was done by comparing the mean MMN amplitude against a hypothetical zero at the frontal (Fz) electrode site, where the MMN is most prominent. The paired-sample $t$-test was also made between the MMN amplitude of across-category and within-category change vowel.

**RESULTS**

The grand-averaged ERPs in Fig. 1 show that both across-category and within-category vowel changes elicited MMN between 100 and 140 ms with reference to the standard-stimulus ERPs. The MMN amplitude was statistically significant for both across-category and within-category vowel change (Table 1). The MMN amplitude differed significantly between the across-category and within-category vowel change conditions, being larger in amplitude after an across-category than a within-category vowel change ($t$-test; $t(9)=3.968; p<0.003$).

Table 2 demonstrates the $xyz$-values in Talairach space as calculated with LORETA in the time window 100–140. The across-category vowel change activated the left superior temporal gyrus ($-59, -32, 8; t$-value, $-0.0025$) more strongly. The within-category vowel change activated the

![Fig. 1](image)

**Table 1.** The mismatch negativity (MMN) mean amplitudes, standard deviations, and $t$-values for the different deviant stimuli used.

<table>
<thead>
<tr>
<th>Deviant stimulus</th>
<th>Mean MMN Amplitude ($\mu$V)</th>
<th>Standard deviation ($\mu$V)</th>
<th>$t$ values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Across-category</td>
<td>-3.53</td>
<td>0.90</td>
<td>12.05</td>
</tr>
<tr>
<td>Within-category</td>
<td>-2.58</td>
<td>0.24</td>
<td>35.39</td>
</tr>
</tbody>
</table>

$p<0.001$.

**Table 2.** Stereotaxic coordinates of the strongest activation foci during the across-category and within-category change discrimination.

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Coordinates (mm)</th>
<th>$t$ values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BA $x$ $y$ $z$</td>
<td>$t$ values</td>
</tr>
<tr>
<td>Across-category</td>
<td>22 $-59$ $-32$ 8</td>
<td>-0.0025</td>
</tr>
<tr>
<td></td>
<td>Left posterior superior temporal gyrus</td>
<td>$-0.0021$</td>
</tr>
<tr>
<td>Within-category</td>
<td>22 $60$ $-32$ 15</td>
<td>$-0.0021$</td>
</tr>
<tr>
<td></td>
<td>Right posterior superior temporal gyrus</td>
<td>$-0.0021$</td>
</tr>
</tbody>
</table>
right superior temporal gyrus (60, −32, 15; t-value, −0.0021) (Table 2 and Fig. 2) more strongly.

**DISCUSSION**

The main finding of our study indicates that the prominent response to across-category and within-category change of vowel in consonant–vowel syllable elicited MMN peaking at 100–140 ms from stimulus onset. The magnitude of the acoustic difference between the stimulus pairs was reflected by the MMN amplitude. The MMN amplitude was larger in across-category than within-category vowel in consonant–vowel syllable contrasts.

Furthermore, results of a tomographic source analysis show a significant bilateral hemispheric activity correlation with the different types of vowel change perception. The LORETA-MMN generator of across-category vowel change was predominantly active in the left auditory cortex while the within-category change of vowel was predominantly active in the right auditory cortex. The left-hemispheric distribution of the source location, estimated by the LORETA in this study, reflected the phonetic and acoustic contrast discrimination of the vowel. This can be the evidence for the left-hemispheric phoneme representation in the human auditory cortex. The converging evidence from factorial statistics of hemispheric effects on LORETA values around the MMN peaks provides further support for the significance of these preattentively activated hemispheric differences in vowel category discrimination.

For a long time, the EEG inverse problem could not find a satisfactory solution. As verified by point spread function simulations, sources could be reconstructed at best with some nonsystematic localization error. LORETA was found to have zero localization error [14]; hence, it finally solved.
exactly the EEG inverse solution problem (as far as localization error in the absence of noise is concerned). The present study thus used standard LORETA space consisting of 2594 voxels of size 7 × 7 × 7 mm. The spatial dispersion of the reconstruction with 21 electrodes is considerable with the zero localization. With respect to the usefulness of a new tomographic source analysis, the present data showed that the LORETA can locate multidipolar sources accurately to demonstrate the dynamic, alternating between left and right homologous regions for different category of vowels in language processing.

Moreover, using the current density measures, we demonstrate that different stages of vowel category processing in consonant–vowel syllable may be associated with left or right hemisphere prominence, in addition to bilateral activation in the superior temporal gyrus. Left-sided activation occurs predominantly to the across-category vowel change, while its right-sided activation occurs mainly for the within-category vowel change. This suggests that differential superior temporal gyrus activation is due to phonetic rather than lexical–semantic processes. Bilateral activation in the superior temporal gyrus might be involved in encoding combinations of temporal and spectral features as phonetic representations. The across-category change of vowel was specific to changes in spectral cues, but the within-category change of vowel was specific to changes in temporal cues. Early stages, associated with across-category change of vowel perception, thus were demonstrated by the left hemisphere. This was followed by activation of the homologous right hemisphere region for within-category vowel change [12,19,20]. The different patterns of activation to within-category and across-category changes of vowels support the view that information transfer in the right and left hemispheres is for analysis of different aspects: physical features within the same phoneme boundary in the right and independent phonemes in the left hemisphere.

Furthermore, when the vowels of the stimulus pair belonged to different categories, in other words, were in the across-category condition (/ɔ/–/i/ as in /pd/–/pi/), the assignment of the stimulus as standard or deviant could have created a contrast effect. The assignment, on the other hand, played no role in the same or within-category one (i.e. /ɔ/–/o/ as in /pd/–/po/). This might have resulted from the fact that the phonetic status was not obvious to the present participants because the area of the /pd/ as stimulus standard is near the /po/ as stimulus deviant. The results thus showed that the MMN depended primarily on the acoustic differences between the members within the vowel pairs. In other words, the MMN amplitude was larger for a vowel pair with a large acoustic distance than for a pair with a smaller distance in the acoustic space.

The present results support a model suggesting that in the intact brain each hemisphere relies on earlier processing stages in both hemispheres and on the evolving stimulus, which contributes to searching the lexicon. This model is compatible with separate right and left hemispheric distributions of neural networks, interconnected and using the processing of the other hemisphere’s output as input for its computations. The time scale of interhemispheric transfers suggests that information is transferred back and forth between left and right homologous regions. The extensive information exchanges between right and left homologous sites suggest that as a word is evolving over time, right and left hemisphere regions contribute and interact to create efficient linguistic parallel processing [12]. This may include the transfer of information to the right hemisphere to free resources in the left one to focus on the meanings [12,19] or to process further input as the word’s sounds progress over time [20]. Thus, the bilateral activation in the superior temporal gyrus of different category vowel change perception was compatible with previous findings [11] that two possible time courses may be involved: (1) both hemispheres are active in parallel with the left hemisphere dominant for different phoneme boundaries and (2) hemispheric dominance may alternate to play a specific role at different stages of linguistic processing. This supports the hypothesis that a part of the superior temporal gyrus contains neural networks specialized for language processing [11].

Our results are supported by the MEG studies showing that the processing of a longer vowel (600 ms) was mainly lateralized on the left hemisphere [21,22]. However, contradicting evidence was also found in previous reports that employed speech sounds with shorter duration: the left hemispheric predominant MMN was not systematically obtained [21,23–25]. It has been proposed that isolated semisynthetic vowels with short duration in a repetitive manner are not processed fully as phonemes in the participant’s brain [2]. Another possible reason for the discrepancy between our study and previous studies is the naturalness of stimuli. It has been hypothesized that the categorical perception of vowels is increased by the complexity of the synthesis and thus affected by the listener’s discrimination behavior [3]. By the use of natural speech of vowel in consonant–vowel syllable, the present data show that the speech-sound naturalness already affects the earlier, preattentive level of speech perception.

CONCLUSION

The vowel across-category change perception in consonant–vowel syllables elicited MMN peaking at 100–140 ms from stimulus onset. As assessed by LORETA, the MMN was predominantly generated in the left auditory cortex, thus emphasizing the role of the left hemisphere in the auditory preattentive processing of vowel across-category change perception in consonant–vowel syllables.

REFERENCES


Acknowledgement: This study was supported by the Thailand Research Fund (TRF) under Grant number MRG4780099. The authors sincerely thank the anonymous referees and the editor of NeuroReport for their constructive comments.