Journal of
Applied Sciences

Source Localization of Preattentive Processing for Different Vowel Duration Changes with Contour Tones in Monosyllabic Thai Words

Wichian Sittiprapaporn

Faculty of Medicine, Mahasarakham University, Mahasarakham, 44000, Thailand

Abstract: The objective of this study was to investigate that the Mismatch Negativity (MMN) can be used to provide an index of experience-dependent and long-term memory traces for different vowel duration changes with contour tones in monosyllabic Thai words. Twenty-two healthy right-handed adults participated in this study. It was found that the long-to-short vowel duration with falling and rising tone changes elicited a strong MMN bilaterally for both native and nonnative speakers of Thai, unlike short-to-long vowel duration with falling and rising tone changes. Source localization analyses demonstrated that sources were obtained in the Middle Temporal Gyrus (MTG) of the left hemisphere and in the Superior Temporal Gyrus (STG) of the right hemisphere for both subject groups.

Key words: Event-related potentials, mismatch negativity, sound, tone, vowel

INTRODUCTION

The human voice recognition and discrimination has been recently measured from the electrophysiological activity of the perceiver's brain (Naatanen, 2001). This was investigated by using an objective measure of pre-attentive sound discriminability, called the Mismatch Negativity (MMN), a component of the auditory Event-Related Potential (ERP) (Alho, 1995). Mismatch negativity can be used to investigate the neural processing of speech and language (Naatanen, 1992, 2001; Alho, 1995; Naatanen and Winkler, 1999; Pulvermuller *et al.*, 2001; Shtyrov *et al.*, 1998, 2000; Naatanen *et al.*, 1997) because it is considered to be a unique indicator of automatic cerebral processing of acoustic stimuli (Shtyrov and Pulvermuller, 2002). MMN, with its major source of activity in the supratemporal auditory cortex, is a brain response elicited in an oddball paradigm where a sequence of repetitive, 'standard', stimuli is interspersed with occasional 'deviant' stimuli that differ from the standard in one or several acoustical or temporal features (Alho, 1995). MMN is thus primarily a response to an acoustic change and an index of sensory memory. Importantly, the MMN can be elicited in the absence of the subjects attention (Naatanen, 1992).

The purpose of the present study was to use both an auditory MMN component of Event-Related Potential (ERP) recording and the Low Resolution Electromagnetic Tomography (LORETA) techniques to measure the degree of cortical activation and to localize the brain area contributing to the scalp recorded auditory MMN

component, respectively, during the passive oddball paradigm. Thus, the main objective of this study was to investigate that the mismatch negativity, a unique indicator of automatic cerebral processing of acoustic stimuli, can be used to provide an index of experience-dependent and long-term memory traces for vowel duration changes with contour tones in monosyllabic Thai words.

MATERIALS AND METHODS

Subjects: Twenty-two healthy right-handed adults with normal hearing and no known neurological disorders volunteered for participation: eleven Native Speakers (NS) of Thai, aged 18-35 (mean 24.2; five females) and eleven Nonnative Speakers (NonS), aged 23-35 (mean 31.2; seven females). The mean (±SD) age was 24.35 (±4.95) years. The approval of the institutional committee on human research and written consent from each subject were obtained.

Stimuli and procedure: Stimuli consisted of four pairs of monosyllabic, Thai words. Speech stimuli were digitally generated and edited to have equal peak energy level in decibels SPL with the remaining data within each of the stimuli scaled accordingly using the Cool Edit Pro v. 2.0 (Syntrillium Software Corporation). The sound pressure levels of speech stimuli were then measured at the output of the earphones (E-A-RTONE 3A, 50 Ω) in dBA using a Brüel and Kjaer 2230 sound-level meter. Four different stimuli were synthetically generated as follow:

- **Stimulus 1:** /kʰaam/-long vowel, falling tone
- **Stimulus 2:** /kʰam/-short vowel, falling tone
- **Stimulus 3:** /kʰaam/-long vowel, rising tone
- **Stimulus 4:** /kʰam/-short vowel, rising tone

The vowel-duration difference between stimulus (1) and (2) was 46 msec (546 vs. 500 msec) and between stimulus (3) and (4) was 56 msec (595 vs. 539 msec) with the same intensity used in each stimulus. Five NS listened to the synthesized words and evaluated them all as natural sounding.

The Standard (S)/Deviant (D) pairs for each experiment which was randomized across subjects, were shown:

Experiment 1: Standard (1)-Deviant (2):
(Stimulus 1: /kʰaam/-long vowel, falling tone)-(Stimulus 2:/kʰam/-short vowel, falling tone)

Experiment 2: Standard (2)-Deviant (1):
(Stimulus 2: /kʰam/-short vowel, falling tone)-(Stimulus 1:/kʰaam/-long vowel, falling tone)

Experiment 3: Standard (3)-Deviant (4):
(Stimulus 3: /kʰaam/-long vowel, rising tone)-(Stimulus 4:/kʰam/-short vowel, rising tone)

Experiment 4: Standard (4)-Deviant (3):
(Stimulus 4: /kʰam/-short vowel, rising tone)-(Stimulus 3:/kʰaam/-long vowel, rising tone)

The sounds were presented binaurally via headphones (Telephonic TDH-39-P) at 85 dB. The inter-stimulus interval (ISI) was 1.25 sec (offset-onset). Deviant stimuli appeared randomly among the standards at 10% probability. Each experiment included 125 trials (10% D). The stimuli were binaurally delivered using SuperLab software (Cedrus Corporation, San Pedro, USA) via headphones (Telephonic TDH-39-P). EEG signal recording was time-locked to the onset of a word. Subjects were instructed not to pay attention to the stimuli presented via headphones but rather to concentrate on a self-selected silent, subtitled movie.

Electroencephalographic recording: For EEG/ERP recording, the standard 20 locations of the 10-20 system, EEG is recorded via an Electro-cap (Electrocap International) from 20 active electrodes (Fp1, Fp2, F7, F3, Fz, F4, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, Oz, O2) positioned according to the 10-20 International System of Electrode Placement, plus Oz and Ground are applied, pre-mounted in an elastic Electro-Cap. Reference

electrodes are manually applied to left and right mastoids, where the Fp1 and Fp2 electrodes are used for ocular artifact detection. Horizontal eye movements are monitored with electrodes at the left and right outer canthi and vertical eye movements are monitored at Fp1 and Fp2. EEG is amplified with a gain of 30,000 and filtered with a bandpass of 0.1-30 Hz. EEGs are acquired as continuous signals and are subsequently segmented into epochs of 1 sec (a 100 msec pre-stimulus baseline and a 900 msec post-stimulus epoch).

EEG data processing: The recordings were filtered and carefully inspected for eye movement and muscle artifacts. ERPs were obtained by averaging epoch which started 100 msec before the stimulus onset and ended 900 msec thereafter; -100-0 msec interval was used as a baseline. Epochs with voltage variation exceeding $\pm100\ \mu V$ at any EEG channel were rejected from further analysis. The MMN was obtained by subtracting the response to the standard from that to the deviant stimulus. 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 (GFP) with an epoch of 40 msec time window related stable scalp-potential topography (Pascual-Marqui *et al.*, 1994). In the next step, low-resolution electromagnetic tomography (LORETA) was applied to estimate the current source density distribution in the brain which contributes to the electrical scalp field (Pascual-Marqui *et al.*, 1994). Maps were computed with LORETA. LORETA computed the smoothest of all possible source configurations throughout the brain volume by minimizing the total squared Laplacian of source strengths.

Data analysis: During the auditory stimulation, electric activity of the subjects' brain was continuously recorded. The MMN was obtained by subtracting the response to the standard from that to the deviant stimulus. The statistical significance of MMN was tested with one sample t-test. An across-experiment ANOVA was carried out so as to make cross-linguistic comparisons.

RESULTS

The grand-averaged ERPs show that both long-to-short and short-to-long vowel duration changes perception elicited MMN between 172-264 msec with reference to the standard-stimulus ERPs. An ANOVA comparing MMN amplitudes of the S and D yield a main effect of conditions in Experiments 1 ($F_{(3,40)} = 8.61$,

Table 1: Mean amplitude (µV)±SD of MMN elicited by a vowel duration changes with contour tones perception in NS and NonS

Vowel	NS	NonS
Long (falling)-to-short (falling)	-1.61±0.18	-1.12±0.59
Short (falling)-to-long (falling)	-1.16±0.24	-1.09±0.24
Long (rising)-to-short (rising)	-1.22±0.31	-0.30±0.21
Short (rising)-to-long (rising)	-0.93±0.25	-0.11±0.49

NS: Native speaker, NonS: Nonnative speaker

Table 2: Stereotaxic coordinates of activation foci during the vowel duration changes with level tone perception

Vowel duration	Coordinates (mm)				
	BA	x	y	z	t-values
Native Speaker (NS)					
Long (Falling)-to-Short (Falling)	39	-52	-67	8[a]	4.15
	22	53	-60	15[b]	2.69
Short (Falling)-to-Long (Falling)	39	-52	-67	8[a]	3.58
	22	53	-60	15[b]	1.89
Long (Rising)-to-Short (Rising)	21	-52	-20	3[a]	3.64
	37	46	-67	8[c]	2.79
Short (Rising)-to-Long (Rising)	21	-52	-20	3[a]	2.39
	22	53	-60	15[b]	2.06
Nonnative Speaker (NonS)					
Long (Falling)-to-Short (Falling)	39	-52	-67	8[a]	3.60
	22	53	-60	15[b]	2.59
Short (Falling)-to-Long (Falling)	39	-52	-67	8[a]	2.73
	37	46	-67	8[c]	1.59
Long (Rising)-to-Short (Rising)	39	-52	-67	8[a]	3.38
	22	53	-60	15[b]	1.65
Short (Rising)-to-Long (Rising)	39	-52	-67	8[a]	2.81
	21	60	-32	6[b]	2.12

[a]Left middle temporal gyrus (MTG), [b]Right superior temporal gyrus (STG), [c]Right middle temporal gyrus (MTG)

$p < 0.0001$) and 3 ($F_{(3,40)} = 23.62$, $p < 0.0001$). In Experiment 2 and 4, however, the S-D differences were not significant (e.g., $F_{(3,40)} = 1.22$, $p = 0.2511$ in Experiment 2, n.s.; $F_{(3,40)} = 0.52$, $p = 0.615$ in Experiment 4, n.s., for the main effect of conditions). The result showed that long-to-short duration changes with falling and rising tones both elicited a strong MMN bilaterally for NS and NonS, unlike short-to-long duration changes with falling and rising tone changes (Table 1). Furthermore, an across-experiment ANOVA demonstrated an interaction and main effects. The significant difference in MMN amplitudes was observed between groups across experiments ($F_{(7,80)} = 45.61$, $p < 0.0001$).

Source localization analyses were performed using LORETA-KEY (Pascual-Marqui *et al.*, 1994). Table 2 demonstrates the xyz-values in Talairach space as calculated with LORETA in the time window 172-264 msec. In Experiment 2 and 3, a single source was estimated to be located in the Middle Temporal Gyrus (MTG) of each hemisphere for both subjects groups. In Experiment 1 and 4, sources were obtained in the MTG of the Left Hemisphere (LH) and in the Superior Temporal Gyrus (STG) of the Right Hemisphere (RH) for both subject groups. No hemispheric difference was discovered in this study (Table 2, Fig. 1-8).
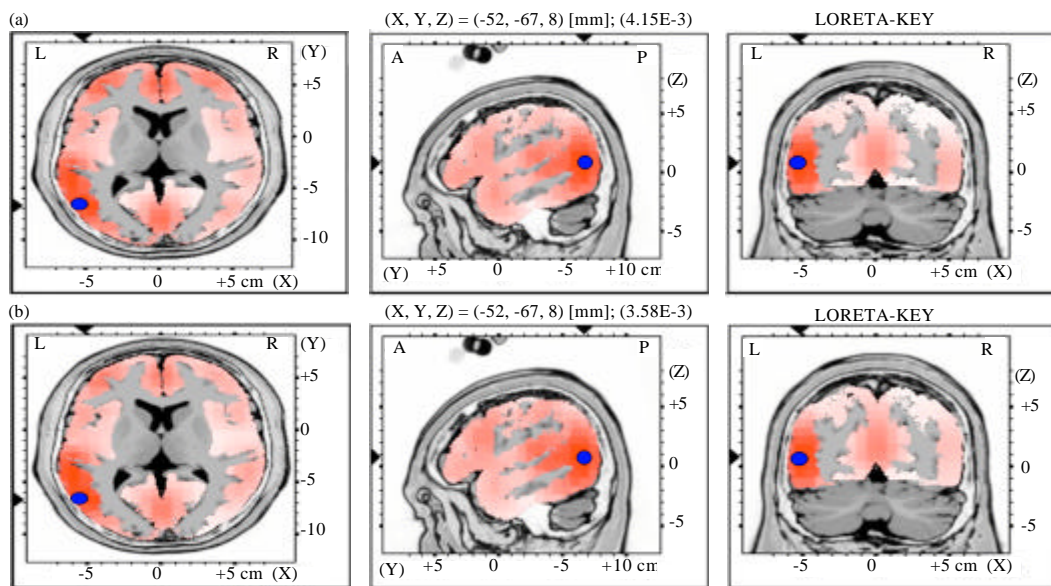

Fig. 1(a-b): LORETA graph t-statistic comparing the event-related potentials (ERPs) for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for, (a) Long-to-short and (b) Short-to-long duration changes of vowels with falling tone of native speaker (NS) activated in left hemisphere (LH), Red color: Local maxima of increased electrical activity for different duration of vowel change responses in an axial, a sagittal and a coronal slice through the reference brain, Blue dots: The center of significantly increased electric activity
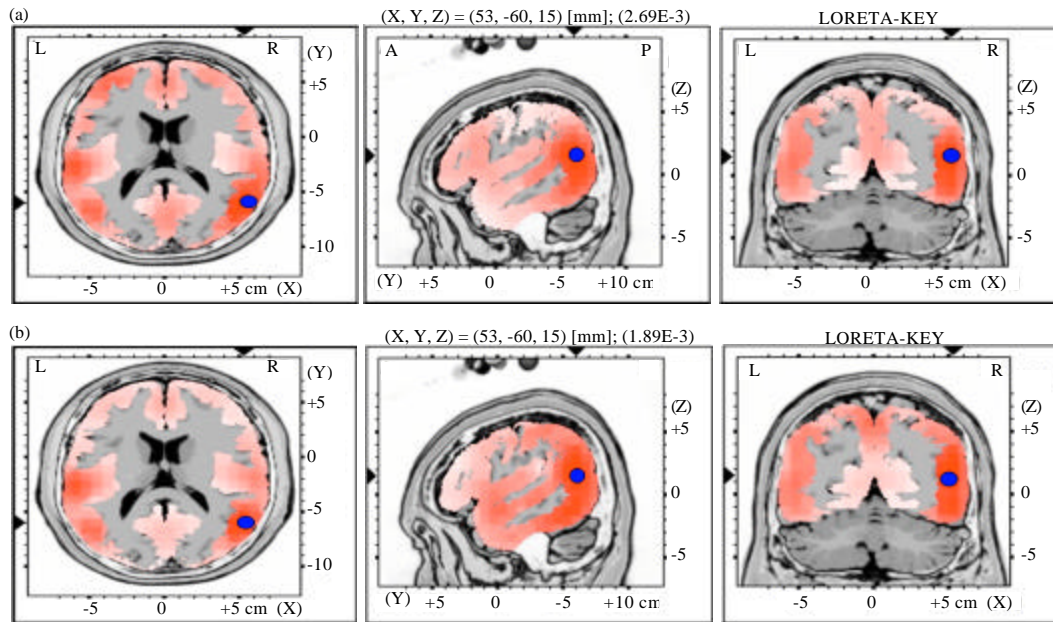
Fig. 2(a-b): LORETA graph t-statistic comparing the event-related potentials (ERPs) for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for, (a) Long-to-short and (b) Short-to-long duration changes of vowels with falling tone of NS activated in RH
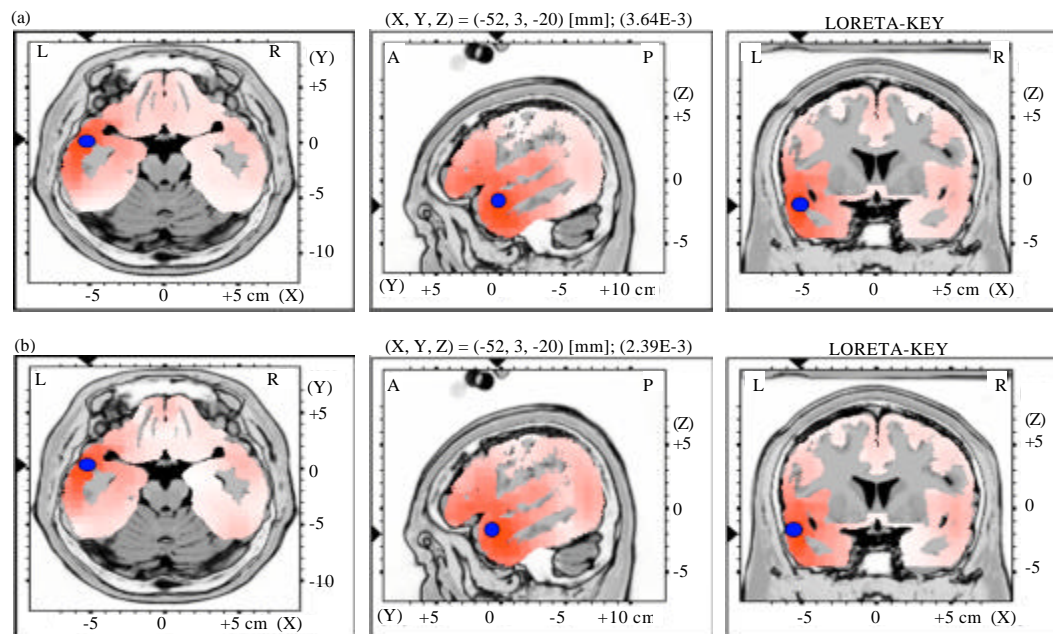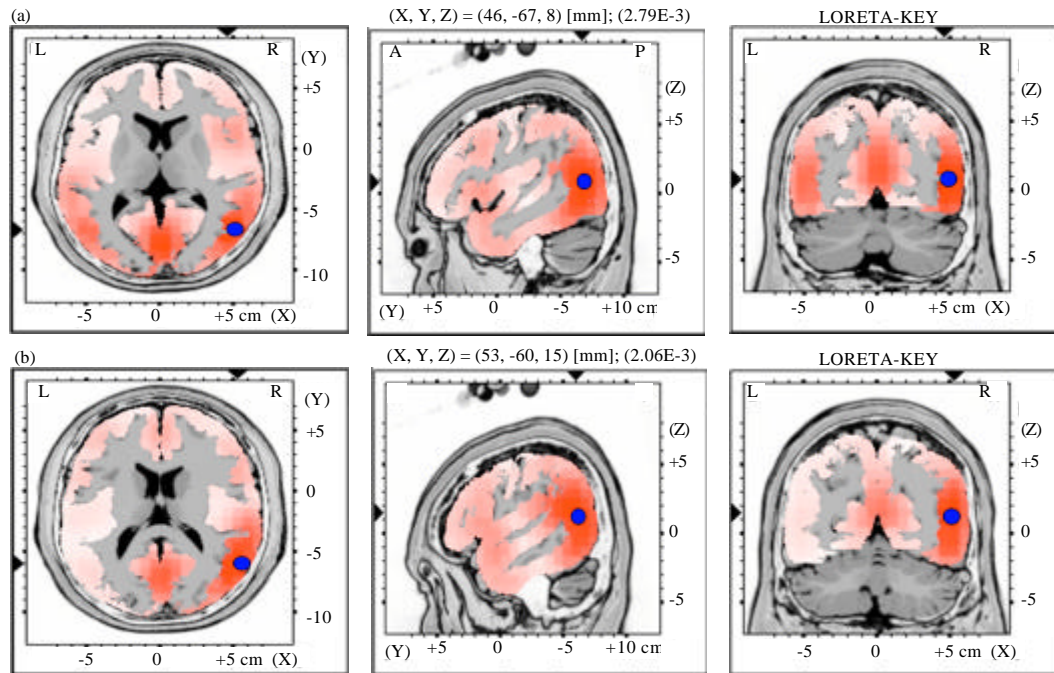

Fig. 3(a-b): LORETA graph t-statistic comparing the event-related potentials (ERPs) for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for, (a) Long-to-short and (b) Short-to-long duration changes of vowels with rising tone of NS activated in LH

Fig. 4(a-b): LORETA graph t-statistic comparing the event-related potentials (ERPs) for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for, (a) Long-to-short duration and (b) Short-to-long duration changes of vowels with rising tone of NS activated in RH
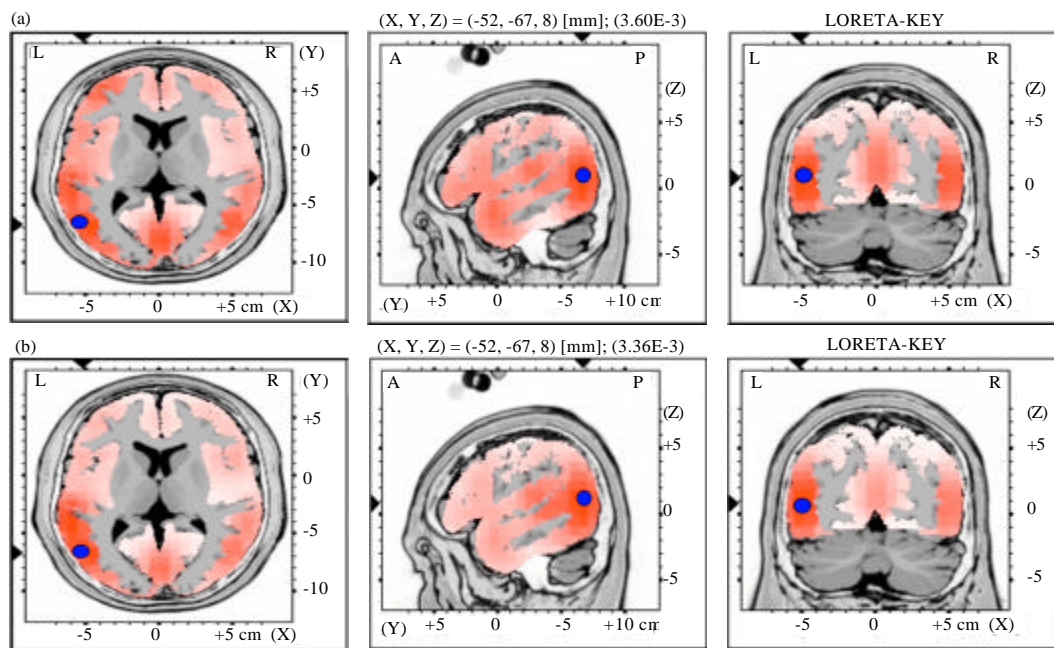

Fig. 5(a-b): LORETA graph t-statistic comparing the event-related potentials (ERPs) for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for, (a) Long-to-short duration and (b) Short-to-long duration changes of vowels with falling tone of NonS activated in LH
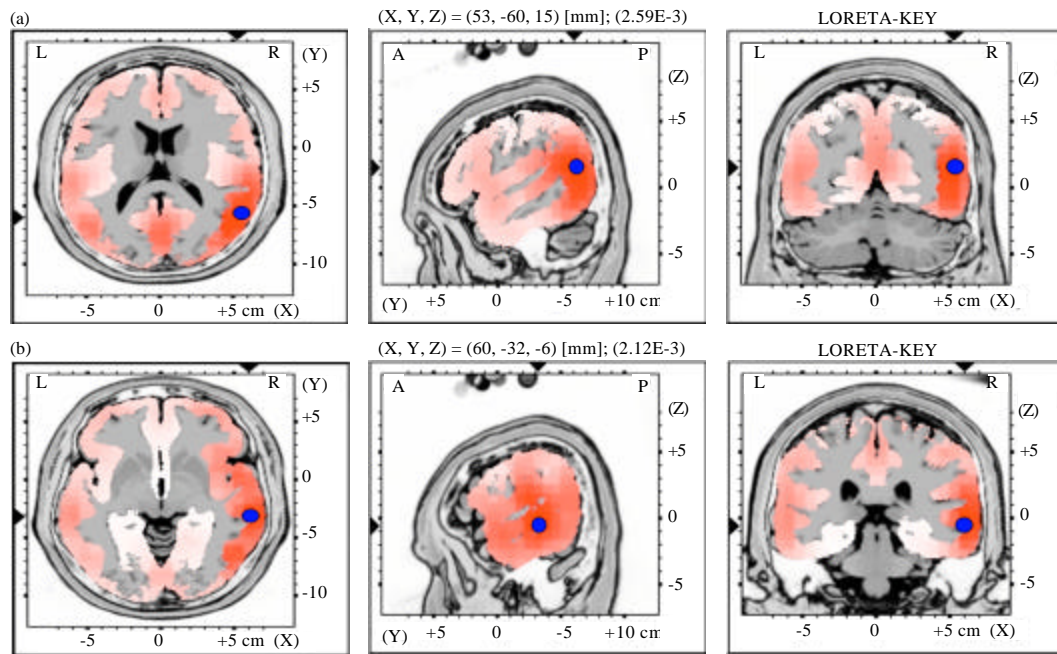
Fig. 6(a-b): LORETA graph t-statistic comparing the event-related potentials (ERPs) for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for, (a) Long-to-short duration and (b) Short-to-long duration changes of vowels with falling tone of NonS activated in RH
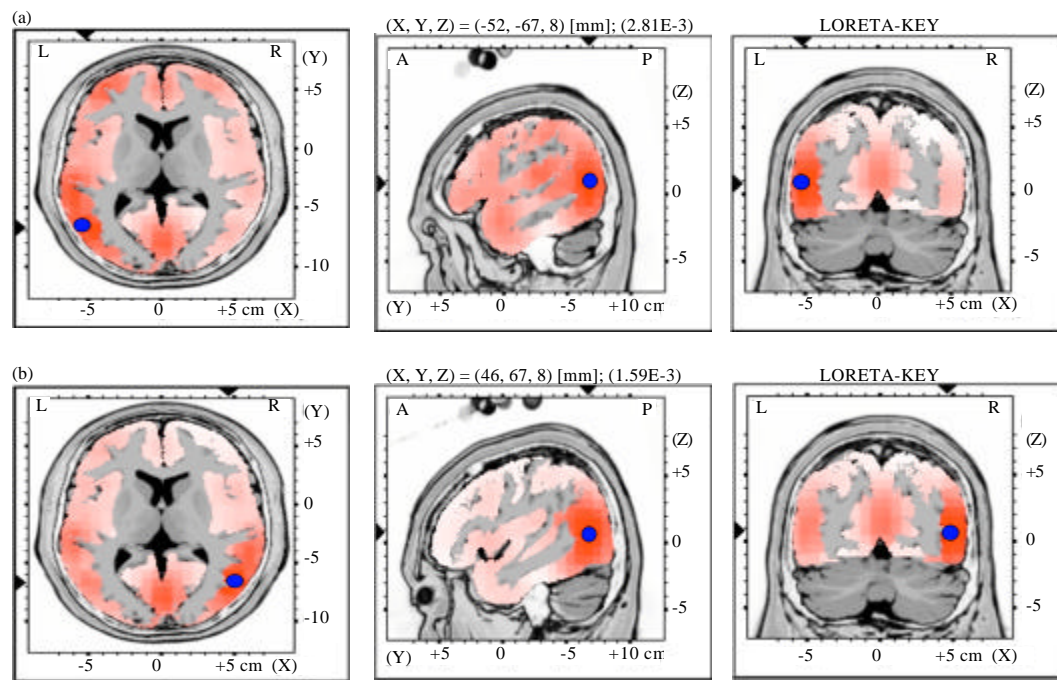

Fig. 7(a-b): LORETA graph t-statistic comparing the event-related potentials (ERPs) for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for, (a) Long-to-short duration and (b) Short-to-long duration changes of vowels with rising tone of NonS activated in LH
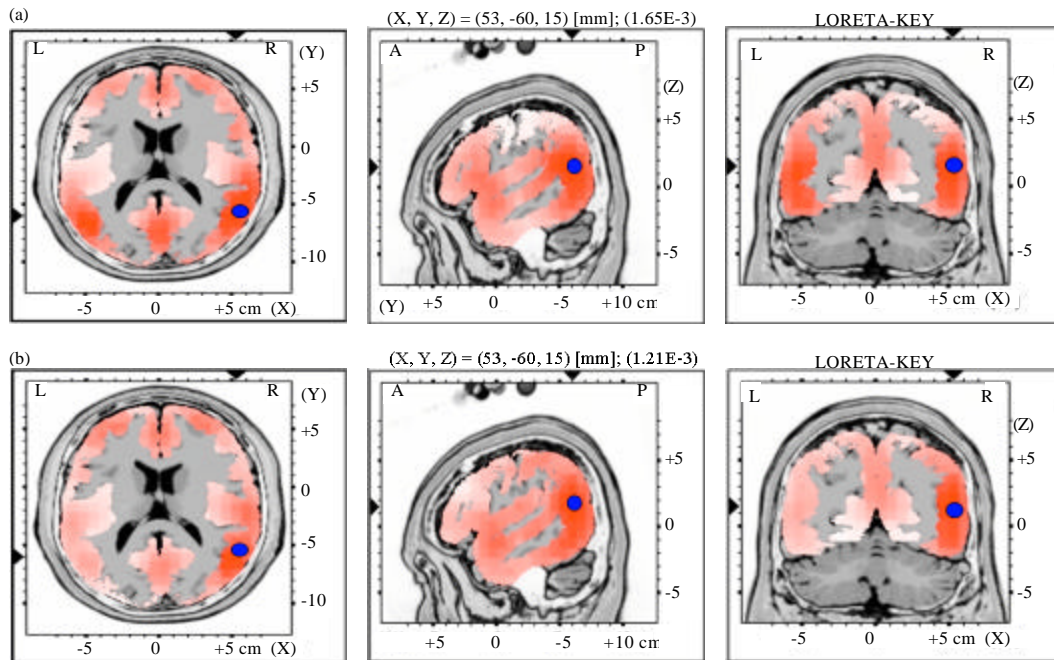
Fig. 8(a-b): LORETA graph t-statistic comparing the event-related potentials (ERPs) for mismatch negativity (MMN) responses at the time point of the individual peak over Fz for, (a) Long-to-short duration and (b) Short-to-long duration changes of vowels with rising tone of NonS activated in RH

DISCUSSION

The magnitude of the acoustic difference between the stimulus pairs was reflected by the MMN amplitude. The different vowel duration changes with falling and rising tone perception elicited MMN between 172-264 msec with reference to the standard-stimulus ERPs. The long-to-short duration changes of vowel with falling and rising tone elicited a strong MMN bilaterally for both native and nonnative speakers of Thai, unlike short-to-long duration with changes of vowel falling and rising tone.

Source localization analyses using LORETA-KEY demonstrated that sources were obtained in the Middle Temporal Gyrus (MTG) of the left hemisphere and in the Superior Temporal Gyrus (STG) of the right hemisphere for both subject groups. While the right hemisphere is predominant in the perception of the non-native speech sounds, the left hemisphere is predominant in the perception of the native speech sounds. The hemispheric pattern of the cortical activity related to native speech-sound discrimination is different from that involved in non-native speech-sound discrimination processing: the left-hemispheric dominance in speech perception can be observed already at its early stage and may be explained by pre-existing long-term memory traces (the acoustic templates of perceived signals) for the native speech sounds formed in the dominant hemisphere.

As the MMN presumably reflect the early stage of speech processing in the human brain, the MMN reflects an early, pre-attentive, automatic speech processing (Naatanen *et al.*, 1997). So, from the known early auditory-cortex responses to sounds, only the mismatch negativity seems to be sensitive to the hemispheric lateralization of the speech function. An advantage of the possible application of MMN as a measure of speech lateralization is that it can be used, unlike behavioral measures, with any subject groups, including patients unable to communicate or concentrate on a test task. The MMN might be of potential interest as a technique of evaluating speech-processing lateralization, since its measurement is non-invasive, relatively inexpensive (especially in case of the EEG) and applicable to any subjects or patients (Naatanen, 1992, 2001; Naatanen *et al.*, 1997; Naatanen and Winkler, 1999).

The current findings are supported by the MEG studies showing that the processing of a longer vowel (600 msec) was mainly lateralized on the left hemisphere (Eulitz *et al.*, 1995; Obleser *et al.*, 2001). However, contradicting evidence was also found in previous

reports which employed speech sounds with shorter duration: the left hemispheric predominant MMN was not systematically obtained (Eulitz *et al.*, 1995; Aaltonen *et al.*, 1994; Tervaniemi *et al.*, 1999; Vihla and Salmelin, 2003). It has been proposed that isolated semisynthetic vowels with short duration in a repetitive manner are not processed fully as phonemes in the subject's brain (Kasai *et al.*, 2001). Another possible reason for the discrepancy between present 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 (Shtyrov *et al.*, 2000; Savela *et al.*, 2003). 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 e.g., preattentive level of speech perception.

CONCLUSION

The long-to-short duration with falling and rising tone changes elicited a strong MMN bilaterally for both native and nonnative speakers of Thai, unlike short-to-long duration with falling and rising tone changes. Source localization analyses demonstrated that sources were obtained in the Middle Temporal Gyrus (MTG) of the left hemisphere and in the Superior Temporal Gyrus (STG) of the right hemisphere for both subject groups.

REFERENCES

Aaltonen, O., O. Eerola, A.H. Lang, E. Uusipaikka and J. Tuomainen, 1994. Automatic discrimination of phonetically relevant and irrelevant vowel parameters as reflected by mismatch negativity. J. Acoustic Soc. Am., 96: 1489-1493.

Alho, K., 1995. Cerebral generators of Mismatch Negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. Ear Hearing, 16: 38-51.

Eulitz, C., E. Diesch, C. Pantev, S. Hampson and T. Elbert, 1995. Magnetic and electric brain activity evoked by the processing of tone and vowel stimuli. J. Neurosci., 15: 2748-2755.

Kasai, K., H. Yamada, S. Kamio, K. Nakagome and A. Iwanami *et al.*, 2001. Brain lateralization for mismatch response to across-and within-category change of vowels. Neuroreport, 12: 2467-2471.

Naatanen, R. and I. Winkler, 1999. The concept of auditory stimulus representation in cognitive neuroscience. Psychol. Bull., 125: 826-859.

Naatanen, R., 1992. Attention and Brain Function. Lawrence Erlbaum, Hillsdale, USA.

Naatanen, R., 2001. The perception of speech sounds by the human brain as reflected by the Mismatch Negativity (MMN) and its Magnetic Equivalent (MMNm). Psychophysiology, 38: 1-21.

Naatanen, R., A. Lehtokoski, M. Lennes, M. Cheour and M. Huotilainen *et al.*, 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. Nature, 385: 432-434.

Obleser, J., C. Eulitz, A. Lahiri and T. Elbert, 2001. Gender differences in functional hemispheric asymmetry during processing of vowels as reflected by the human brain magnetic response. Neurosci. Lett., 314: 131-134.

Pascual-Marqui, R.D., C.M. Michel and D. Lehmann, 1994. Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain. Int. J. Psychophysiol., 18: 49-65.

Pulvermuller, F., T. Kujala, Y. Shtyrov, J. Simola and H. Tiitinen *et al.*, 2001. Memory traces for words as revealed by the Mismatch Negativity (MMN). NeuroImage, 14: 607-616.

Savela, J., T. Kujala, J. Tuomainen, M. Ek, O. Aaltonen and R. Naatanen, 2003. The mismatch negativity and reaction time as indices of the perceptual distance between the corresponding vowels of two related languages. Cognitive Brain Res., 16: 250-256.

Shtyrov, Y. and F. Pulvermuller, 2002. Memory traces for inflectional affixes as shown by mismatch negativity. Eur. J. Neurosci., 15: 1085-1091.

Shtyrov, Y., T. Kujala, J. Ahveninen, M. Tervaniemi, R.J. Ilmoniemi and R. Naatanen, 1998. Background acoustic noise and the hemispheric lateralization of speech processing in the human brain: Magnetic mismatch negativity study. Neurosci. Lett., 251: 141-144.

Shtyrov, Y., T. Kujala, S. Palva, R.J. Ilmoniemi and R. Naatanen, 2000. Discrimination of speech and of complex nonspeech sounds of different temporal structure in the left and right cerebral hemispheres. Neuroimage, 12: 657-663.

Tervaniemi, M., A. Kujala, K. Alho, J. Virtanen, R.J. Ilmoniemi and R. Naatanen, 1999. Functional specialization of the human auditory cortex in processing phonetic and musical sounds: A Magnetoencephalographic (MEG) study. NeuroImage, 9: 330-336.

Vihla, M. and R. Salmelin, 2003. Hemispheric balance in processing attended and non-attended vowels and complex tones. Cognitive Brain Res., 16: 167-173.