Sound perception affected by nonlinear variation of accuracy in memory trace

Hirooki Yabe,CA Ren Asai, Tomiharu Hiruma, Takeyuki Sutoh, Sachiko Koyama,[1] Ryusuke Kakigi,[1]
Mari Tervaniemi[2] and Sunao Kaneko

[1]Department of Neuropsychiatry, Hirosaki University School of Medicine, Hirosaki 036-8562; Department of Integrative Physiology, National Institute for Physiological Sciences, Okazaki, Japan; [2]Cognitive Brain Research Unit, Department of Psychology, University of Helsinki, Finland

CACorresponding Author: yabe@cc.hirosaki-u.ac.jp

Received 13 October 2004; accepted 4 November 2004

The mismatch negativity (MMN) reflects the neural representation of the acoustic environment stored in sensory memory. The short phase of sensory memory corresponds to the temporal window of integration (TWI) which integrates the neighboring sounds into a unitary event. We measured the magnetic MMN (MMNm) and the discriminative reaction time (RT) responding to an omitted segment incorporated into a complex sound. Consequently, for the late omitted segments, the MMNm amplitudes were decreased, and those latencies and the RTs were prolonged. The percentages of the correct responses were also reduced for the late omitted segments. In sum, the discriminative sensitivity nonlinearly declined toward the end of the TWI, suggesting that the time-wise accuracy of the neural trace nonlinearly varies in sensory memory. *NeuroReport* 15:2813–2817 © 2004 Lippincott Williams & Wilkins.

Key words: Auditory perception; MEG; Memory trace; Mismatch negativity; Sensory memory; Temporal window of integration

INTRODUCTION

Automatic change-detection system in the human brain was revealed in audition by investigating the mismatch negativity (MMN) [1] and its magnetic counterpart (MMNm) [2]. The generator of the MMN is located in the vicinity of the primary auditory cortex [3]. The discriminative system as reflected by the MMN requires the storage of the previous state of the acoustic environment for detecting an incoming deviating sound [4,5]. The sensory memory underlying this detection system maintains the information on various elements of repetitive sounds, such as frequency [6,7], intensity [6,8], duration [9], language [10] and omission [11–14]. The MMN amplitude was large and the MMN latency was short when a discrimination performance was good [4]. Furthermore, the improvement of the discrimination performance correlated with the enlarged amplitude and the shortened latency of the MMN. Näätänen interpreted this finding so that the improved trace, that is, the increased accuracy of the sensory information encoded in the trace, made it easier for the (sensory memory) system to detect stimulus deviation against the background formed by the neural trace [4]. In other words, the discriminative sensitivity as reflected by the amplitude and latency of the MMN can be regarded as an index of the accuracy of the trace in sensory memory.

Cowan's [15] short phase of sensory memory appears to correspond to the temporal window of integration (TWI) which integrates the neighboring sounds into single information units [4,16]. Recent MMN studies have provided physiological evidence for the TWI [11–13,17–18]. The duration of this short phase of sensory memory was estimated to be 160–170 ms [14]. The results of a previous study [19] suggested, however, that automatic discriminative sensitivity might not be uniform in TWI of auditory sensory memory. The similar phenomenon has been reported as an auditory induction [20].

The purpose of the present study is to estimate the time-wise accuracy of the sound trace represented in sensory memory, as reflected by the time-course of discriminative sensitivity inside the TWI, and then to examine the influence of this varying accuracy on the discrimination performance. To this end, we measured the MMNm responses for the sequences of complex sounds containing an omitted segment. In a separate experiment, we measured the discriminative reaction times (RTs) of age-matched subjects for the same stimulus sequences and compared them with the results of MMNm.

MATERIALS AND METHODS

Magnetoencephalography (MEG) sessions: The MMNm of 21 healthy subjects (13 males, mean age 32.1 ± 4.98 years) were measured with MEG. During the experiment, subjects were instructed to concentrate on watching a self-selected silenced movie and to ignore the sounds. Subjects gave informed consent after the purpose of the study was explained to them.

Auditory stimuli were presented with a constant SOA of 500 ms to the subject's left ear via a plastic tube from outside the magnetically and electrically shielded sound-attenuated chamber. The movie was projected onto the screen in the chamber from outside. The number of stimuli was 4000. The stimulus sequences were randomized from seven types of sounds as shown in the left column of Fig. 1: standard stimuli (probability of 70%), and 6 different deviants of stimuli each at a probability of 5%. The standard stimulus consisted of 8 tone burst segments with a frequency of 1000 Hz and an intensity of 80 dB SPL as follows: 0–22 ms, 22–44 ms, 44–66 ms, 66–88 ms, 88–110 ms, 132–154 ms, and 154–176 ms (each with an 18 ms plateau and 2 ms rise/fall time). The six deviant stimuli were formed by omitting one of six middle segments in time.

Data were measured with dual 37-channel magnetometers (Magnes, Biomagnetic Technologies Inc., San Diego, CA). The gradiometers were arranged in a uniformly distributed array in concentric circles over a spherically concave surface. The device was 144 mm in diameter. The probe was centered at around the C3 and C4 positions of the International 10-20 System which covered the left and right hemispheres. Each coil of 20 mm diameter was connected to a superconducting quantum interference device (SQUID).

Data were filtered with a 1–50 Hz bandpass filter, and digitized at a sampling rate of 520.8 Hz. The analysis time was 50 ms before and 400 ms after stimulus onset, and DC was offset using the pre-stimulus period as the baseline. A spherical model was fitted to the digitized head shape of each subject, and the location, orientation and amplitude of the best-fitted single equivalent current dipole (ECD) were estimated for each time point [21]. Magnetic resonance imaging (MRI) was performed using a GE Signa 1.0 T system. The statistical analysis was carried out by using Geisser-Greenhouse (ε) corrected analysis of variance (ANOVA) with repeated measures and Tukey multiple comparison test. The root mean square (RMS) values of all channels in the right hemisphere were used in the analysis of the amplitudes. Two types of peak latencies were used in the analysis: the absolute peak latency was defined as the latency measured from the onset of the complex stimulus, while the relative peak latency was defined as the latency measured from the onset of each omitted segment.

Psychophysics sessions: Twenty-one age-matched healthy subjects (13 females; mean age 31.67 ± 6.45 years) were instructed to respond by pressing a key as rapidly as possible upon detecting these stimuli. The reaction times to the deviant stimuli were measured. Subjects gave informed consent.

The same stimulus sequences as these used in the MEG sessions were presented with a constant SOA of 1000 ms to the subject's left ear. The number of stimuli was 1000. The subjects were instructed to respond as rapidly as possible to deviant stimuli by pressing a response key. The time window for an acceptable response was defined as 200–1000 ms. Two types of reaction time were used in the analysis: the absolute reaction time was defined as the interval measured from the onset of the complex stimulus, whereas the relative reaction time was defined as the interval measured from the onset of each omitted segment.

RESULTS

Figure 1 presents the difference waveforms of a representative subject to the six deviant stimuli. The distinct deflections (identified as the MMNm) in response to each omitted segment in the 1st, 2nd, 3rd, 4th, 5th, and 6th showed different peak latencies for the different deviant stimuli. The grand-averaged waveforms for the 21 subjects calculated by RMS values of all channels are shown in Fig. 2a. The peak amplitudes of these deflections were determined by calculating the peak of RMS values of all channels from the onset of the complex stimulus. The MMNm peak amplitudes were smaller for the omitted segments at the later part of the complex sound (i.e., deviant 4th, 5th, and 6th) than at the early part of that (i.e., deviant 1st, 2nd, and 3rd), as shown in Fig. 2b. A repeated-measures ANOVA also showed a statistically significant difference between these amplitudes measured by the peak of RMS ($F(5,100)=0.479$, $p<0.01$, $\varepsilon=0.649$).

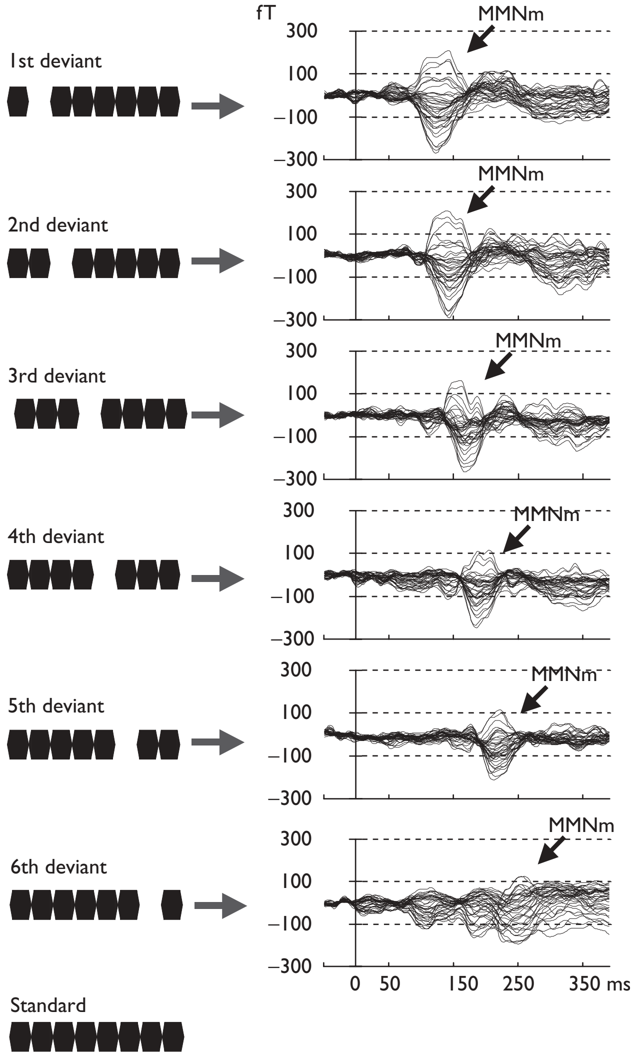

Fig. 1. Thirty-seven superimposed difference waveforms for the 1st, 2nd, 3rd, 4th, 5th and 6th deviant stimuli (response to the standard stimuli subtracted from each).

The absolute MMNm latency: A repeated-measures ANOVA showed a statistically significant difference between the absolute latencies measured by the peak of RMS ($F(5,100)=725.713$, $p<0.001$, $\varepsilon=0.649$). The MMNm peak latencies were significantly shorter the earlier was the

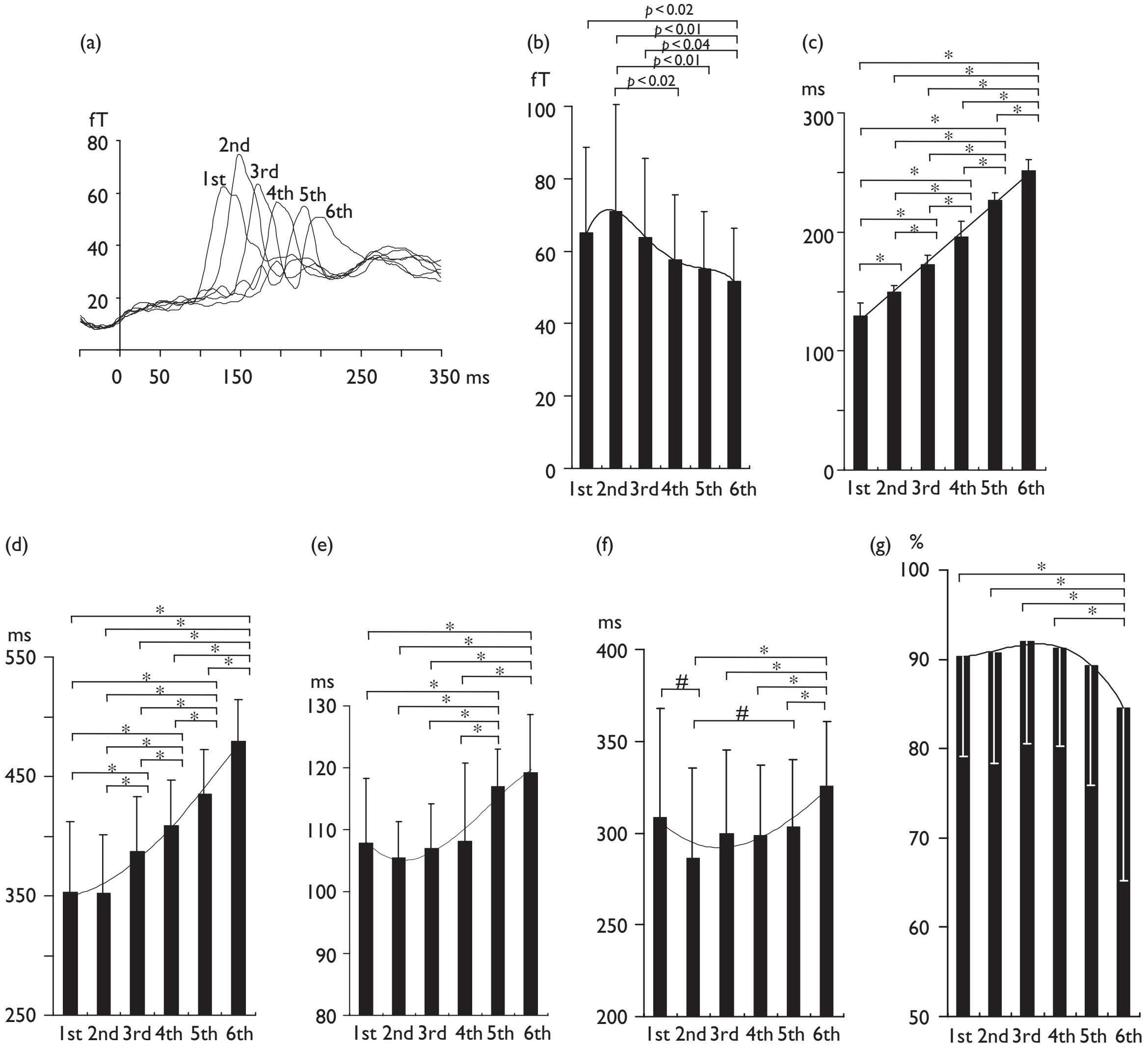

Fig. 2. (**a**) The grand-averaged waveforms for the 21 subjects calculated by RMS values of all channels for MMNm responses recorded from the right hemisphere of all subjects. (**b**) The peak amplitudes of RMS values for MMNm responses recorded from the right hemisphere of all subjects. The curved line shows the approximate function of $y=-0.399 \times 4+6.2233 \times 3-33.8 \times 2+69.643x+23.475$. (**c**) Absolute peak-latencies of RMS values for the MMNm responses recorded from the right hemisphere in all subjects. The line shows the approximate function of $y=24.65x+101.52$. (**d**) Absolute reaction times for the different deviant stimuli. The curved line shows the approximate function of $y=-0.306 \times 3+6.5713 \times 2-7.3297x+351.29$. (**e**) Relative peak latencies of RMS values for the MMNm responses recorded from the right hemisphere of all subjects. The curved line shows the approximate function of $y=-0.3064 \times 3+6.5713 \times 2-7.3297x+351.29$. (**f**) Relative reaction times for the deviant stimuli 1–6. The curved line shows the approximate function of $y=-0.3064 \times 3+6.5713 \times 2-29.33x+329.29$. (**g**) Percentages of the correct responses. The curved line shows the approximate function of $y=-0.1481 \times 3+0.8719 \times 2-0.8847x+90.476$. ($x=1, 2, 3, 4, 5$, and 6 (corresponding to the first omitted section (onset at 44 ms), second (66 ms), third (88 ms), fourth (110 ms), fifth (132 ms), and sixth (154 ms) deviant stimulus). The error bar indicates s.d. (Tukey multiple comparison test: $*p < 0.01$).

omitted section in the stimuli (Fig. 2c). The relationship between these latencies was linear. The time differences between these peak latencies were nearly proportional to the multiple intervals between the consecutive segments, which was a multiple of 22 ms.

In the psychophysics sessions, a repeated-measures ANOVA showed a statistically significant difference between all absolute reaction times ($F(5,100)=142.745$, $p<0.001$, $\varepsilon=0.439$). The absolute reaction times for deviant stimuli 1–6 were significantly different from each other in this order (Fig. 2d). The relationship between all absolute reaction times was almost linear as were also the absolute peak latencies of the MMNm. The time differences between these peak latencies were nearly proportional to the multiple intervals between the consecutive segments, which was a multiple of 22 ms.

The relative MMNm latency and reaction time: The relative peak latencies of the MMNm were longer for the omitted segments at the later part of the complex stimuli (i.e., the 5th) and especially the 6th segments (omitted) than at the early part of the complex stimuli (i.e., the 2nd and 3rd omitted segments), as shown in Fig. 2e. A repeated-measures ANOVA showed a statistically significant difference between these latencies as measured from the peak of

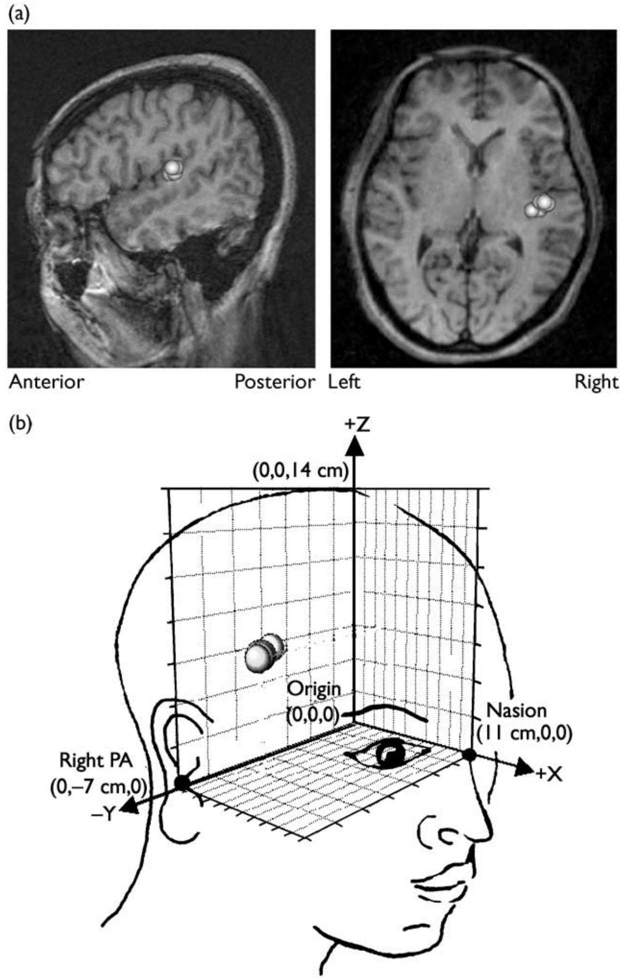

Fig. 3. (**a**) The ECD locations to MMNm in a representative subject on the MRI calculated for the lst, 2nd, 3rd, 4th, 5th, and 6th deviant stimuli (gray filled circle). (**b**) Averaged ECD locations to MMNm for all subjects calculated for the lst, 2nd, 3rd, 4th, 5th, and 6th deviant stimuli (gray filled circle).

RMS ($F(5,100)=11.379$, $p<0.001$, $\varepsilon=0.649$). The relationship between these latencies was nonlinear. The relative reaction times were longer for the omission of the 6th than of the 2nd segment, as shown in Fig. 2f. This result was similar to that for the relative peak latencies. A repeated-measures ANOVA showed a statistically significant difference between these relative reaction times measured from the peak of RMS ($F(5,100)=9.798$, $p<0.001$, $\varepsilon=0.395$). The relationship between the relative reaction times was nonlinear as was also that between the relative peak-latencies.

The percentages of the correct responses: The percentages of the correct responses were low for the omitted segments at the later part of the complex stimuli, (especially deviant 6), as shown in Fig. 2g. A repeated measures ANOVA showed a statistically significant difference between these percentages ($F(5,100)=4.009$, $p<0.03$, $\varepsilon=0.417$). The relationship between these percentages was nonlinear.

MMNm generator loci: The ECDs for the peak deflections (MMNm) for all deviant stimuli were calculated with a goodness of fit of $>90\%$. The ECDs of a representative subject on the MRI are shown in Fig. 3a. The ECDs for the MMNm responses for all subjects are also shown in Fig. 3b. The origin of the head-based coordinate system was the midpoint between the preauricular points. The x-axis indicated the coronal plane with a positive value toward the left preauricular point, and the z-axis lay on the transverse plane perpendicular to the x-y line with a positive value toward the upper side. These ECDs suggested that the dominant neural activity generating these responses lay within the supratemporal auditory cortex. There was no significant difference between the ECD locations for the different deviant stimuli (x-position: $F(5,85)=0.185$, n.s., $\varepsilon=0.533$, y-position: $F(5,85)=0.254$, n.s., $\varepsilon=0.642$, z-position: $F(5,85)=0.313$, n.s., $\varepsilon=0.563$).

DISCUSSION

The present study evaluated the time-wise accuracy of the neural trace represented in sensory memory by measuring MEG responses to the omitted segments placed on different time periods within the postulated duration of the TWI. All these 6 different deviant stimuli elicited distinct MMNm components. The ECDs of these components in the supratemporal auditory cortex corresponded to the other MEG-based localizations of the MMN [2].

The MMNm peak amplitudes were largest for the omitted segment of 2nd deviant and were smallest for the omitted segment of 6th deviant. The approximate function curve of the peak amplitudes nonlinearly decreased in the time order of the omitted segments toward the end of complex sound. The MMN amplitude usually decreases when the magnitude of deviation is decreased [4]. However, this is not the case in the present study, for the magnitude of deviation is equal among the 6 deviant stimuli employed. Therefore, the MMNm-amplitude difference of between the early and late segments indicates the varying accuracy of the sensory-memory trace.

The absolute peak latencies of the MMNms for the different deviants were sequentially delayed in the order of the omitted segments (deviants 1–6). The latency delays were almost proportional to the multiple of the 22 ms of the segment duration, that is, the time points at which the changes, i.e., the omitted segments, were initiated. These findings were supported by the previous studies of MMNs to changes in sound duration [9,22]. Thus, the MMNm should begin as soon as the change is detected. Furthermore, the activation of the change-detector system underlying the MMNm generation is triggered by the initiation of change, but this process still lasts after change detection [19]. The absolute RT showed a tendency of sequential delay similar to that for the absolute peak latency of the MMNm. It implies that the brain processes related to automatic perception and perceptional performance may be initiated proportionally when the change begins.

The approximate curves of the absolute peak latency of the MMNm and the absolute reaction times are superimposed by another nonlinear component. To extract the nonlinear tendency, the relative MMNm peak latencies and relative reaction times were calculated. Both were shortened at the early part of the complex sound and were prolonged at the later part of the complex sound. The difference in the relative peak latency and in the relative reaction time between the early and late segments reflects the varied accuracy of the trace in sensory memory. The approximate curves of the relative peak latency of the MMNm and the relative reaction time showed the phase-reversal in relation to those of the MMNm peak amplitudes. Furthermore, the percentages of the correct responses were also reduced for the omission of the later segments of the complex sounds.

The results involving the peak amplitude and relative peak latency of the MMNm indicate that automatic discriminative sensitivity peaked at around 50 (44–66)ms from the onset of the complex sound and gradually declined toward the end of this sound whose duration corresponds to the supposed TWI in auditory sensory memory. As described above, the discriminative sensitivity can be regarded as an index of the accuracy of the neural trace in sensory memory. The time-wise accuracy of the sound trace is not uniform over sensory memory. Furthermore, as suggested by the findings involving the relative reaction time and the percentage of the correct responses, the declining accuracy at the later part of sensory memory resulted in the fall of behavioral performance. That is to say, the perceptional performance is affected by the varying accuracy of memory trace.

CONCLUSION

The time-wise sensitivity of the neural trace nonlinearly varies in sensory memory, resulting in a corresponding fall of the perceptional performance. In conclusion, sound perception should be affected by nonlinear variation of accuracy in sensory memory.

REFERENCES

1. Näätänen R, Gaillard AWK and Mäntysalo S. Early selective attention effects on evoked potential reinterpreted. *Acta Psychol* 1978; **42**:313–329.
2. Alho K. Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear Hear* 1995; **16**:38–51.
3. Scherg M, Vajsar J and Picton TW. A source analysis of the late human auditory evoked potentials. *J Cogn Neurosci* 1989; **5**:363–370.
4. Näätänen R. *Attention and Brain Function*. Hillsdale, NJ: Erlbaum; 1992.
5. Brattico E, Winkler I, Naatanen R, Paavilainen P and Tervaniemi M. Simultaneous storage of two complex temporal sound patterns in auditory sensory memory. *Neuroreport* 2002; **13**:1747–1751.
6. Paavilainen P, Degerman A, Takegata R and Winkler I. Spectral and temporal stimulus characteristics in the processing of abstract auditory features. *Neuroreport* 2003; **14**:715–718.
7. Sams M, Paavilainen P, Alho K and Näätänen R. Auditory frequency discrimination and event-related potentials. *Electroencephalogr Clin Neurophysiol* 1985; **62**:437–448.
8. Näätänen R, Paavilainen P, Alho K, Reinikainen K and Sams M. The mismatch negativity to intensity changes in an auditory stimulus sequence. In: Johnson R, Rohrbaugh JW and Parasuraman R (eds). *Current Trends in Event-Related Brain Potential Research*. Amsterdam: Elsevier; 1987, pp. 125–131.
9. Kaukoranta E, Sams M, Hari R, Hämäläinen M and Näätänen R. Reactions of human auditory cortex to changes in tone duration. *Hear Res* 1989; **41**:15–22.
10. Näätänen R, Lehtokoski A, Lennes M, Cheour M, Huotilainen M, Iivonen A *et al*. Language-specific phoneme representation revealed by electric and magnetic brain responses. *Nature* 1997; **385**:432–434.
11. Tervaniemi M, Saarinen J, Paavilainen P, Danilova N and Näätänen R. Temporal integration of auditory information in sensory memory as reflected by the mismatch negativity. *Biol Psychol* 1994; **38**:157–167.
12. Yabe H, Tervaniemi M, Reinikainen K and Näätänen R. The temporal window of integration in auditory system as revealed by omission MMN. *Neuroreport* 1997; **8**:1971–1974.
13. Yabe H, Tervaniemi M, Sinkkonen J, Huotilainen M, Ilmoniemi RJ and Näätänen R. Temporal window of integration of auditory information in the human brain. *Psychophysiology* 1998; **34**:615–620.
14. Yabe H, Sato Y, Sutoh T, Hiruma T, Shinozaki N *et al*. The duration of integrating window in auditory sensory memory. In: Barbar C, Celesia GG, Hashimoto I and Kakigi R (eds). *Functional Neuroscience: Evoked Potentials and Magnetic Fields*. Amsterdam: Elsevier; 1999, pp. 166–169.
15. Cowan N. Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychol Psychol Bull* 1988; **104**:163–191.
16. Bregman AS. *Auditory scene analysis: the perceptual organization of sound*. Cambridge: MIT Press; 1990.
17. Winkler I, Reinikainen K and Näätänen R. Event-related brain potentials reflect of echoic memory in humans. *Percept Psychophys* 1993; **53**:443–449.
18. Sussman E, Winkler I, Ritter W, Alho K and Näätänen R. Temporal integration of auditory stimulus deviance as reflected by the mismatch negativity. *Neurosci Lett* 1999; **264**:161–164.
19. Yabe H, Koyama S, Kakigi R, Gunji A, Tervaniemi M, Sato Y and Kaneko S. Automatic discriminative sensitivity inside temporal window of sensory memory as a function of time. *Cogn Brain Res* 2001; **12**:39–48.
20. Warren RM. *Auditory Perception: A New Analysis and Synthesis*. Cambridge: Cambridge University Press; 1999.
21. Sarvas J. Basic mathematical and electromagnetic concepts of biomagnetic inverse problem. *Phys Med Biol* 1987; **32**:11–22.
22. Paavilainen P, Jiang D, Lavikainen J and Näätänen R. Stimulus duration and the sensory memory trace: an event-related potential study. *Biol Psychol* 1993; **35**:139–152.

Acknowledgements: The authors are very grateful to Professor R. Näätänen for the helpful comments, and to the staff of Department of Integrative Physiology (Okazaki). This study was supported by the grant-in-aid for scientific research (12670920, 15390346) from the Japan Society for the Promotion of Science (JSPS).