Mismatch negativity (MMN) reveals sound grouping in the human brain

Naoko Shinozaki,[CA] Hirooki Yabe, Yasuharu Sato, Takeyuki Sutoh, Tomiharu Hiruma, Tadayoshi Nashida and Sunao Kaneko

Department of Neuropsychiatry, Hirosaki University School of Medicine, Zaifu-cho 5, Hirosaki 036-8563, Japan

[CA]Corresponding Author

Received 8 February 2000; accepted 7 March 2000

To investigate a part of the structure of the memory trace, auditory event-related potentials (ERPs) were recorded from reading subjects while they were presented with two different stimulus-series simultaneously. A clear mismatch negativity (MMN) was obtained from each series, when the stimulus sequence consisted of a high-frequency series and a low-frequency series. Moreover, the MMN showed independent elicitation within each series. However, if the frequency range of one series overlapped with that of the other series, the amplitude of the MMN was prominently reduced, suggesting that the two processing functions indexed by MMN coexisted simultaneously in the preattentive acoustic system and were produced by the respective grouping of high-frequency tones and low-frequency tones. *NeuroReport* 11:1597–1601 © 2000 Lippincott Williams & Wilkins.

Key words: Auditory event-related potentials; Memory trace; Mismatch negativity; Parallel processing; Sound grouping

INTRODUCTION

It is clear in general that we are surrounded by various kinds of auditory stimuli and stimulus-sequences which must be integrated and segregated by the auditory information processing system [1]. Therefore, multiple information processing must exist simultaneously in the human brain. The auditory cortex is a complex of parallel and hierarchical auditory fields whose integrated functioning and representation of the acoustic environment is yet to be understood [2]. Mismatch negativity (MMN), a negative component of the event-related potential (ERP), which indexes preattentive acoustic processing [3,4] and is generated in supratemporal auditory cortex [5,6], has partly revealed that these integrative and segregative functions exist in the processing of auditory information [7–12]. However, there are only a few studies which have investigated the global structure of the memory trace minutely.

We can partly infer the structure of the auditory processing system from perceptual phenomena. For example, auditory stream segregation [1] is a perceptual phenomenon that is the general process of auditory scene analysis in which links are formed between parts of the sensory data. In this phenomenon, the auditory system is grouping tones that are similar to one another in preference to grouping tones that follow one another immediately. The perception is that the sets of high and low tones split into separate streams of sound. This phenomenon is an example of the sequential type of grouping. In a recent study, Sussman *et al.* [12] used MMN analysis in order to investigate the streaming effect, an auditory effect created in the laboratory by principles of stream segregation, and indicated that the segregation associated with the streaming effect occurs preattentively. According to them, MMN analysis reveals that when the streaming effect occurs for a sequence of high and low tones, the preattentive processing system arranges the mixture of sounds arriving at the ears into meaningful segments or groupings.

On the other hand, there are some cases where the MMN system shows dissociation from the perceptual phenomenon, for example, as in the case of the octave illusion [13–15]. The octave illusion is experienced when two tones with frequencies separated by a full octave are presented simultaneously to the opposite ears, so that the low and the high tone continuously reverse their positions between the two ears. Under such conditions, subjects tend to perceive only one tone at a time instead of two tones. Right-handed subjects mostly reported the high tone in the right ear when this tone was actually presented to the right ear and the low tone was presented to the left ear, but they reported the low tone in the left ear when it was presented to the right ear and the high tone was presented to the left ear. However, as Ross *et al.* pointed out [16], this illusion is encoded by the MMN system in terms of their physical (spectral) properties rather than their perceptual (illusory) properties.

There is, therefore, uncertainty concerning the organization of sensory input in the early stage of acoustic processing when the perceptual streaming effect does not occur for a sequence of high and low tones because of the slow pace. We hypothesized that the MMN system might be subject to various complicated processing constraints in order to achieve efficient processing of complex information, even if the perceptual phenomenon does not occur. The purpose of the present study is to investigate a portion of the structure of the memory trace formed from an acoustic input consisting of two different stimulus series, when the streaming effect does not occur.

MATERIALS AND METHODS

Eighteen healthy female subjects (range 20–23 years) participated in the experiment. The data from three of the participants were excluded due to excessive artifacts. Subjects gave informed consent after the testing procedure was explained to them. Subjects were seated in a comfortable chair in an electrically shielded and acoustically attenuated room. Each subject was instructed to read a self-selected book and to ignore the auditory stimuli which were delivered monaurally via earphones to the left ear. There were five conditions including three experimental conditions and two control conditions (the alone-conditions). Procedurally, 225 deviants were collected for each stimulus type for each condition. The order of the runs for each condition was randomized. In the experimental conditions, stimulus block consisted of two kinds of stimulus-series, as shown in Fig. 1. In the 15/15 alternating condition, one series was composed of 750 Hz deviant tones (15%) and 500 Hz standard tones (85%). The other series was composed of 2000 Hz deviant tones (15%) and 1500 Hz standard tones (85%). Deviant tones occurred randomly in each series. The constant stimulus onset asynchrony (SOA) of each series was 1000 ms. The time-lag between two series was 500 ms. Therefore, in stimulus sequence which actually impinged on the ear, high-frequency (2000 or 1500 Hz) and low-frequency (750 or 500 Hz) tones were alternated at a constant SOA of 500 ms. In the 15/50 alternating condition, stimulus condition was the same as in the 15/15 alternating condition, except that the probability of the deviants of the high-frequency series was increased from 15% to 50%. In the overlapping condition, the stimulus block consisted of two kinds of stimulus-series, but the overlapped frequency-range of two stimulus-series collapsed the alternating pattern. One series was composed of 750 Hz deviant tones (15%) and 500 Hz standard tones (85%). The other series was composed of 400 Hz tones (50%) and 900 Hz tones (50%). For two control conditions (alone conditions), low tones only (low tone alone condition) or high tones only (high tone alone condition) were used. The order of the tones in alone conditions was kept the same as that of the each series in the 15/15 alternating condition. Each tone lasted 50 ms (rise and fall times 5 ms) and the intensity of the stimuli was 80 dB SPL.

The EEG was recorded with Ag–AgCl electrodes from 10 channels. The electrode positions at the midline were Fpz, Fz, Cz and Pz. Six lateral electrodes were placed equidistantly on the coronal line connecting the mastoids through Fz. The electrodes over the left hemisphere were designated L1, L2 and LM (left mastoid), and those over the right hemisphere as R1, R2 and RM (right mastoid). The electro-oculogram (EOG) was recorded from electrodes placed above the right eye and below the left eye. The reference electrode was attached to the nose.

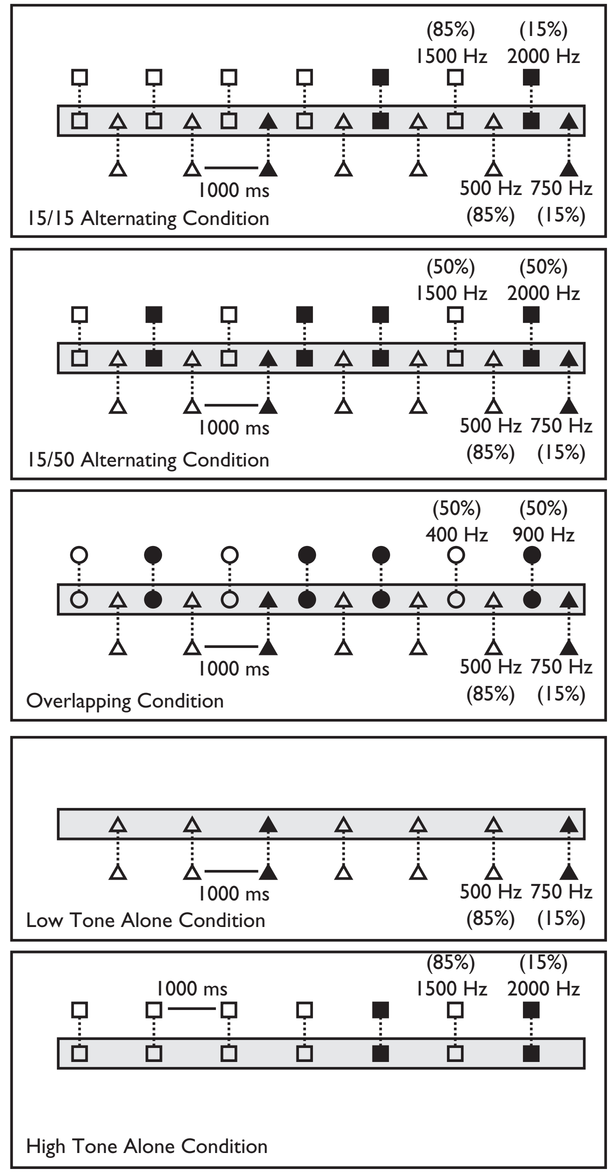

Fig. 1. Schematic illustration of stimulus sequence. Stimulus sequence which actually impinged on the ear (the gray portion) consisted of two kinds of stimulus-series in two alternating conditions and an overlapping condition.

The EEG was amplified with frequency limits of 0.16–120 Hz. The analysis period was 512 ms (sampling rate 1000 Hz), including a 50 ms pre-stimulus baseline. The EEG epochs contaminated by extracerebral artifacts (amplitude change exceeding $\pm 120\,\mu V$) were automatically rejected. After averaging, frequencies < 2 Hz were digitally filtered out from the ERPs (FFT-filter).

ERPs were averaged separately for each series and stimulus type. The MMN was measured from the deviant−standard difference waves. The peak latencies and amplitudes of the MMNs, that is the coordinates of the peak, were measured at Fz during the time window 70–200 ms from stimulus onset. These data were treated to paired *t*-tests and a multivariate analyses of variance (MANOVA, using multivariate test statistic Wilks's lambda) with repeated measures to assess the morphological difference of MMNs.

RESULTS

Subject report: At the end of the recording session, subjects were asked about their subjective experience of hearing the tones occurring in the three experimental conditions. All subjects reported hearing the relatively slow-paced tones. Alternating high- and low-frequency tones at the rate of 2 stimuli/s could not create a streaming effect, as Bregman previously mentioned [1]. The relatively slow-paced tones were reported as alternating high and low pitches in alternating conditions. None of the subjects experienced streaming, but after they were instructed to count the deviant tones in low-frequency series, almost all subjects could detect the deviant stimuli in low-frequency series (i.e. 750 Hz deviant tones). In the overlapping condition, however, in which the frequency ranges of two stimulus series overlapped each other, nobody could detect 750 Hz deviant tones.

ERPs: Figure 2 presents the across-subjects averages ($n = 15$) of the standard−deviant difference waves at five typical recording sites. The black bars represent time spans of significant differences between the amplitudes of ERPs obtained with the standards and the deviants ($p < 0.05$; two-tailed paired t-test). In alternating conditions as well as alone conditions, prominent negativity was elicited to all kinds of deviant stimuli occurring from about 70 to 200 ms relative to stimulus onset. The scalp distribution by EEG monitoring showed a frontocentral maximum and polarity inversion at the mastoids. Because of this polarity inversion, these negativities were identified as MMNs.

Figure 3 shows the superimposed difference waves obtained by the 750 Hz deviants in two alternating conditions. The righthand column shows p values corresponding to each recording site. As for the 750/500 Hz series in the alternating conditions, the MMNs obtained with the deviants did not change, even though the probability of the deviants of the other intervened series (i.e. 2000/1500 Hz series) increased. At all 10 recording sites there was no

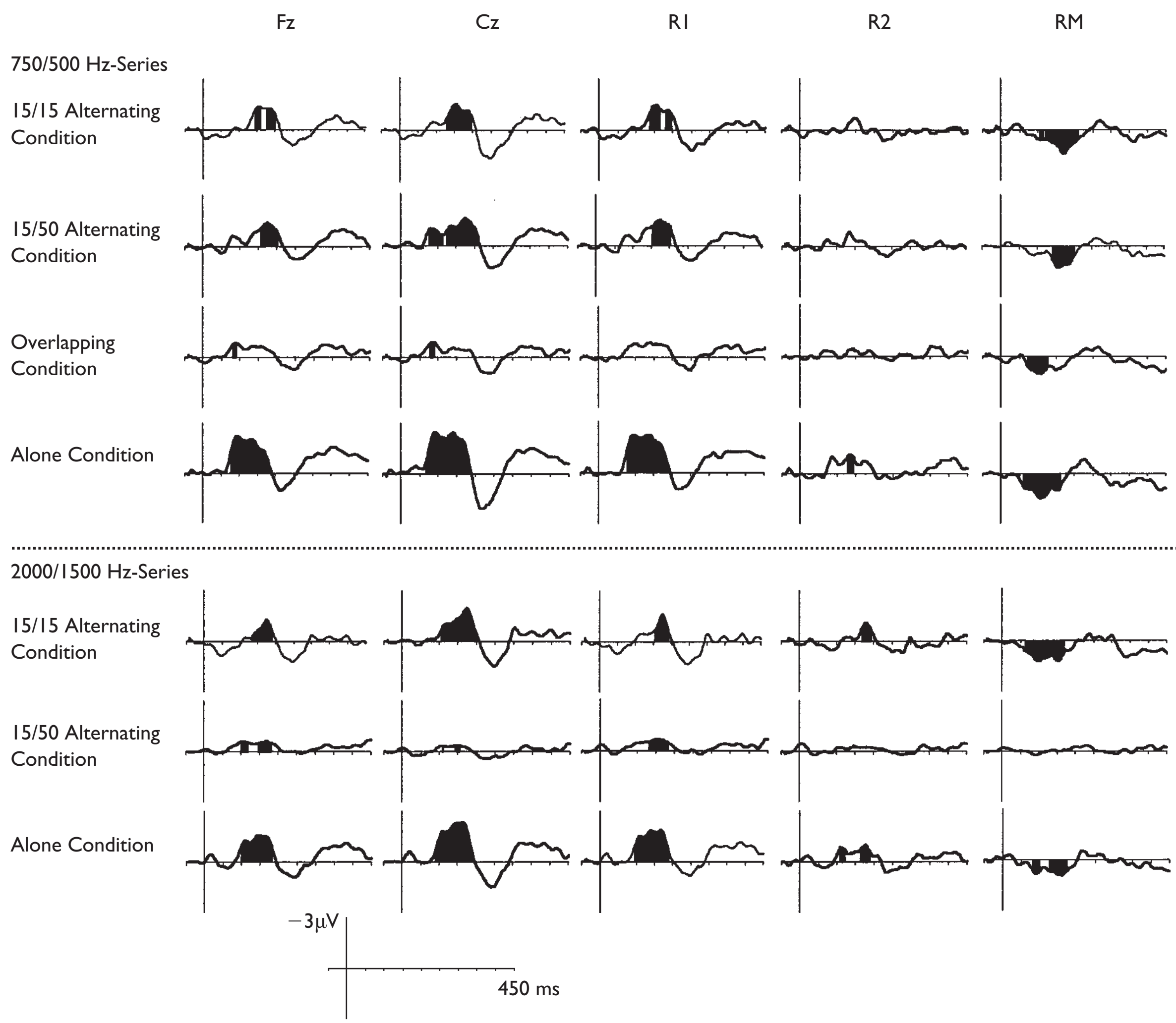

Fig. 2. Grand-average deviant−standard difference waves ($n = 15$) obtained with the deviants in each series and condition. The black bars represent time spans of significant differences between the amplitudes of ERPs obtained with the deviants and the standards ($p < 0.05$; two-tailed paired t-test).

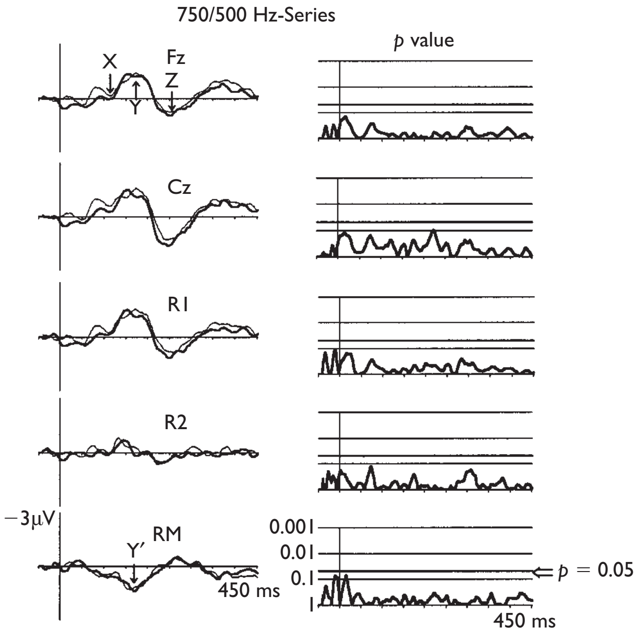

Fig. 3. Left column: Comparison of deviant−standard difference waves obtained with the 750 Hz deviants in two alternating conditions. Thick lines show the ERPs elicited in the 15/15 alternating condition. Thin lines show the ERPs elicited in the 15/50 alternating condition. The arrows indicate the onset-point (X), the peak-point (Y,Y′) and the offset-point (Z) of the MMNs. Right column: *p* values corresponding to each recording site (two-tailed paired *t*-test). No statistically significant point is seen.

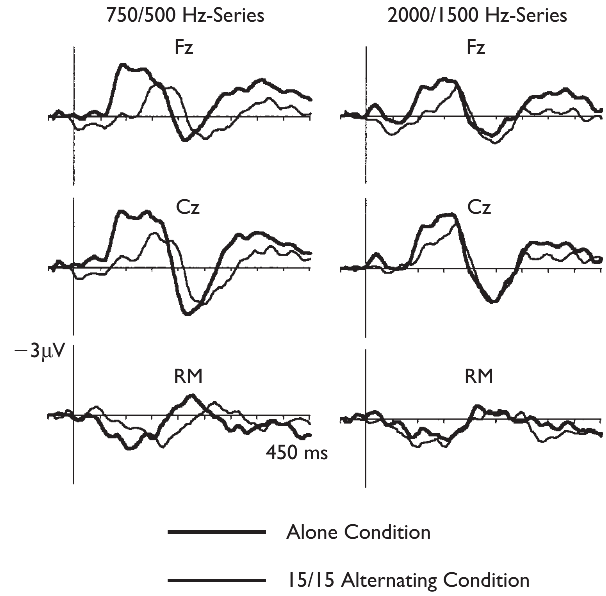

Fig. 4. Superimposed deviant−standard difference waves obtained in the 15/15 alternating condition and alone conditions. Thick lines show the ERPs elicited from alone conditions. Thin lines show the ERPs elicited from the 15/15 alternating condition.

time span of significant differences between the amplitudes of the MMNs obtained from the 750 Hz deviants in the 15/15 alternating condition and those in the 15/50 alternating condition. As can be seen from Fig. 3, the onset-points (X), the peak-points (Y) and the offset-points (Z) of the MMNs at Fz, and the peak-points (Y′) of the reversed MMNs at RM in two conditions were identical to each other, respectively. A repeated-measures MANOVA with one within subject factor (15/15 alternating condition and 15/50 alternating condition) showed no statistical significance for condition for the coordinates of X-, Y-, Z- and Y′-point (X: Wilks's $\lambda = 0.983$; $F(2,13) = 0.111$, $p = 0.896$; Y: Wilks's $\lambda = 0.896$; $F(2,13) = 0.757$, $p = 0.489$; Z: Wilks's $\lambda = 0.966$; $F(2,13) = 0.229$, $p = 0.799$; Y′: Wilks's $\lambda = 0.982$; $F(2,13) = 0.121$, $p = 0.887$). Therefore, the MMNs obtained from the 750 Hz deviants in two alternating conditions revealed almost identical time courses (onset, peak, offset) and morphologies. On the other hand, the MMNs obtained with the deviants of 2000/1500 Hz series changed significantly in accordance with the change of the probability, as shown in Fig. 2. For the 15/50 alternating condition, the amplitude of the MMN elicited from the 2000/1500 Hz series decreased prominently. In addition, its scalp distribution did not show a polarity inversion at the mastoids.

Compared with the MMNs obtained from 750 Hz deviants in the alternating conditions, the MMN elicited in the overlapping condition, which was obtained from 750 Hz deviant tones with the intervention of 900/400 Hz series, decreased prominently (Fig. 2).

Figure 4 shows superimposed deviant−standard difference waves obtained with deviants in the 15/15 alternating condition and those in alone conditions. The peaks of the MMNs at Fz in the 15/15 alternating condition were different from those in alone conditions. A repeated-measures MANOVA with one within-subject factor (alternating condition and alone condition) showed statistically significant effects for condition for the coordinates of the peaks of the MMNs (750/500 Hz series: Wilks's $\lambda = 0.327$; $F(2,13) = 13.368$, $p = 0.001$; 2000/1500 Hz series: Wilks's $\lambda = 0.500$; $F(2,13) = 6.504$, $p = 0.011$). As for the 750/500 Hz series, the peak latency of the MMN obtained with the deviants of the 15/15 alternating condition was significantly longer than that of the alone condition ($p < 0.001$; two-tailed paired *t*-test; Table 1).

DISCUSSION

The logic underlying the design of these studies is that when the stimulus sequence consists of two different stimulus series, the independent elicitation of the MMNs within each series means the coexistence of the two independent processing functions. The present results showed a definite negativity to deviant stimuli in one

Table 1. Peak amplitudes and peak latencies of MMNs at Fz (mean ± s.d.).

Series and condition	Peak amplitude (μV)	Peak latency (ms)
2000/1500 Hz series		
Alone condition	−2.50 ± 1.18	146.4 ± 31.6
Alternating condition	−1.71 ± 1.32	158.8 ± 28.4
750/500 Hz series		
Alone condition	−2.97 ± 2.00	128.3 ± 31.8
Alternating condition	−2.25 ± 1.74	180.5 ± 25.5

stimulus series even if the other stimulus series intervened, when the stimulus sequence consisted of high-frequency series (2000/1500 Hz series) and low-frequency series (750/500 Hz series). This component was identified as the MMN because its scalp distribution showed a fronto-central maximum and polarity inversion at the mastoids [17]. Moreover, the MMN obtained with the deviants of the 750/500 Hz series did not change at all, even though the MMN obtained with the deviants of the 2000/1500 Hz series changed because of the probability effect [18]. As shown in Fig. 2, it seemed that there was a small negativity obtained with the 2000 Hz tones in the 50% condition (15/50 alternating condition), i.e., when there are no deviants. This negativity may be due to the difference between the late exogenous ERP component elicited by 2000 Hz tones and that elicited by 1500 Hz tones. These findings suggest that the processing within the low-frequency series was independent of that within the high-frequency series. Our results could indicate the simultaneous existence of at least two processing functions in the human auditory cortex. In contrast to this, all subjects reported hearing alternating high and low pitches in alternating conditions. None of the subjects had a perceptual experience of two coherent streams (streaming). Therefore, the preattentive auditory system shows dissociation from the perceptual phenomenon in our experiment.

However, if the frequency range of one series overlaps with other series, the preattentive auditory system may not produce the independent processing functions. It appears that the input of the 900/400 Hz series embedded in the 750/500 Hz series interfered with the processing of the 750/500 Hz series. One explanation for this interference may be that the simultaneous existence of the two processing functions in preattentive acoustic systems was produced by grouping tones that were similar to one another in frequency rather than grouping tones that were presented at steady rhythm. In other words, the frequency range overlap prevented grouping.

In the present study, the MMNs obtained with deviants in alone conditions were very different from the MMNs obtained with the deviants in alternating conditions. This distinction between the two conditions, possibly contributing to the difference of the MMNs, was that an MMN was elicited within a single group of sound in alone conditions and an MMN was elicited within the context of two groups of sound in alternating conditions, as Sussman *et al.* mentioned [12]. In addition, the detection of deviants during alternating conditions also seemed to require a higher level of concentration than the detection of deviants during alone conditions when subjects counted the deviant tones.

CONCLUSION

When alternating high- and low-frequency tones at the rate of 2 stimuli/s were presented to subjects, clear MMNs were obtained from the high-frequency series and the low-frequency series respectively. These two MMNs were independent of each other. However, when the tones presented to the subjects were not alternating, that is when the frequency range of one series overlapped with the other series, the amplitude of the MMNs was prominently reduced. These results suggest that the simultaneous existence of the two processing functions in preattentive acoustic system is produced by grouping tones that are similar to one another in frequency.

REFERENCES

1. Bregman AS. *Auditory Scene Analysis. The Perceptual Organization of Sounds.* London: MIT Press, 1990.
2. Schreiner CE. *Curr Opin Neurobiol* **2**, 516–521 (1992).
3. Näätänen R. *Attention and Brain Function*. Hillsdale, NJ: Erlbaum, 1992.
4. Näätänen R and Winkler I. *Psychol Bull* **125**, 826–859 (1999).
5. Kropotov JD, Näätänen R, Sevostianov AV *et al. Psychophysiology* **32**, 418–422 (1995).
6. Alho K. *Ear Hearing* **16**, 38–50 (1995).
7. Ritter W, Deacon D, Gomes H *et al. Ear Hearing* **16**, 52–67 (1995).
8. Sussman E, Ritter W and Vaughan HG Jr. *Brain Res* **789**, 130–138 (1998).
9. Yabe H, Tervaniemi M, Reinikainen K *et al. Neuroreport* **8**, 1971–1974 (1997).
10. Yabe H, Tervaniemi M, Sinkonen J *et al. Psychophysiology* **35**, 615–619 (1998).
11. Winkler I, Czigler I, Jaramillo M *et al. Neuroreport* **9**, 495–499 (1998).
12. Sussman E, Ritter W and Vaughan HG, Jr. *Psychophysiology* **36**, 22–34 (1999).
13. Deutsch D. *Nature* **251**, 307–309 (1974).
14. Deutsch D. *Sci Am* **233**, 92–104 (1975).
15. Deutsch D. *J Acoust Soc Am* **63**, 184–186 (1978).
16. Ross J, Tervaniemi M and Näätänen R. *Neuroreport* **8**, 303–306 (1996).
17. Alho K, Paavilainen P, Reinikainen K *et al. Psychophysiology* **23**, 613–623 (1986).
18. Näätänen R, Sams M, Järvilehto T *et al.* Probability of deviant stimulus and event-related brain potentials. In: Sinz R. and Rosenzweig R, eds. *Psychophysiology 1980.* Amsterdam: VEB Gustav Fischer Verlag Jena (GDR) and Elsevier Biomedical Press: 1980; pp. 397–405.

Acknowledgements: We thank Risto Näätänen, University of Helsinki, for his valuable comments on earlier versions of this article. This study was supported by a Grant-in-Aid (09670972) for Scientific Research from the Japanese Ministry of Education, Science and Culture.