ELSEVIER

Neuroscience Letters 227 (1997) 197–200

NEUROSCIENCE LETTERS

Pre-attentive processing of spectrally complex sounds with asynchronous onsets: an event-related potential study with human subjects

M. Tervaniemi[a,*], E. Schröger[b], R. Näätänen[a]

[a]Cognitive Brain Research Unit, Department of Psychology, P.O. Box 13, FIN-00014 University of Helsinki, Helsinki, Finland
[b]Institute for Psychology, University of Leipzig, Seeburgstrasse 14–20, D-04103 Leipzig, Germany

Received 18 April 1997; revised version received 5 May 1997; accepted 5 May 1997

Abstract

Neuronal mechanisms involved in the processing of complex sounds with asynchronous onsets were studied in reading subjects. The sound onset asynchrony (SOA) between the leading partial and the remaining complex tone was varied between 0 and 360 ms. Infrequently occurring deviant sounds (in which one out of 10 harmonics was different in pitch relative to the frequently occurring standard sound) elicited the mismatch negativity (MMN), a change-specific cortical event-related potential (ERP) component. This indicates that the pitch of standard stimuli had been pre-attentively coded by sensory-memory traces. Moreover, when the complex-tone onset fell within temporal integration window initiated by the leading-partial onset, the deviants elicited the N2b component. This indexes that involuntary attention switch towards the sound change occurred. In summary, the present results support the existence of pre-perceptual integration mechanism of 100–200 ms duration and emphasize its importance in switching attention towards the stimulus change. © 1997 Elsevier Science Ireland Ltd.

Keywords: Auditory information; Pre-perceptual integration; Sensory memory; Pitch perception; Complex tones; Mismatch negativity; N2b; N1

In our everyday life, we are surrounded by a multitude of parallel or sequential sounds which pre-attentively are either integrated to form a unitary auditory percept or segregated to separate objects having different subjective sources [2]. The limits for the temporal integration of auditory information may be determined by the shorter phase of sensory memory (<250 ms) [3]. Recently, it was also shown that when task-irrelevant and task-relevant tones were separated by a 200 ms stimulus-onset asynchrony (SOA) the subjects' behavioral discrimination performance was deteriorated. In contrast, with longer SOAs (up to 560 ms) this did not occur [18]. This suggests that even involuntary attention switch mechanisms may be sensitized during the integration process.

Neural processing of auditory information, including mechanisms of pre-perceptual integration and segregation as well as sensory memory and attention switches, can be probed by recording event-related potentials (ERPs), especially the mismatch negativity (MMN) component [13] which reflects sensory-memory functions [12,19]. The MMN is evoked by infrequent 'deviant' stimuli which differ from the majority of the 'standard' stimuli in any sound parameter even when a subject concentrates on, e.g. reading [11]. MMN implies an incongruency between the sensory input and the stimulus parameters encoded by the cortical [1,6] sensory-memory traces [11]. In addition to simple stimulus features, sensory-memory traces also encode, e.g. the sound order [15]. Furthermore, two separate tones seem to be integrated into a unitary memory trace if presented during temporal window of integration [11]. When two paired tones were separated with an SOA shorter than 200 ms an omission of the second tone elicited an MMN which did not occur with longer SOAs ([20], see also [4,22]). In addition, if an easy primary task and a wide deviation in stimuli is employed, the MMN may be overlapped by N2b component which is suggested to reflect involuntary switch of attention [11,16].

Some sensory-memory functions and the temporal window of integration seem to be reflected by the N1 component as well [8,17]. The anterior subcomponent of the magnetic N1 (N100 mA) peaking at 140 ms was enhanced when the paired tones were presented with SOAs shorter than 300 ms [7]. The authors suggested that the N1 enhance-

---

* Corresponding author. Tel.: +358 9 19123408; fax: +358 9 19122924; e-mail: mari.tervaniemi@helsinki.fi

0304-3940/97/$17.00
PII S0304-3940(97)00346-7

ment reflects the phase when sensory input still is in active stage during integration process.

The present study addressed the neural auditory information processing within and across the temporal window of integration. ERPs were recorded from reading subjects while they were presented with harmonically complex stimuli (consisting of 10 partials) of high-probability standard and low-probability deviant pitch. The deviant stimulus was produced by slightly mistuning the fourth partial from its natural frequency present in standard stimuli. In separate blocks, this mistuned partial commenced at 0, 30, 90, 180, or 360 ms in advance of the remaining complex tone (see Fig. 1). Previously it has been shown [5,10] that if the mistuned frequency precedes the remaining complex tone by 320 ms the tone is segregated into two separate entities whereas with shorter SOAs all partials are integrated together. By combining the ERP methodology [11] with the paradigm developed in behavioral research [5,10] we wished to determine whether the N1 and/or MMN components reflect the presence of the mistuned frequency among a harmonically complex sound. Moreover, it was investigated whether involuntary attention switches occur when SOAs fall within or across the temporal window of integration.

The stimuli [5] were produced by NeuroStim PC software and presented binaurally via headphones. The harmonically complex part of the sound was constructed from 10 sinusoidal partials based on the 155 Hz frequency at equal 45 dB (SPL) intensities (Fig. 1). The complex tone-duration was 100 ms including 10 ms rise and fall times. In standard stimuli, the leading partial was 620 Hz, belonging to the harmonic series of the 155 Hz fundamental. In deviant stimuli, it was replaced by a 600 or 640 Hz tone ($P = 0.075$ for each; ±3.2% mistuning from the natural harmonic frequency). In different blocks, the standard or deviant frequency partial preceded the 100 ms complex tone either by 0, 30, 90, 180, or 360 ms, the offset of all harmonics being simultaneous (see Fig. 1). The intensity of the leading partial was 45 dB (SPL).

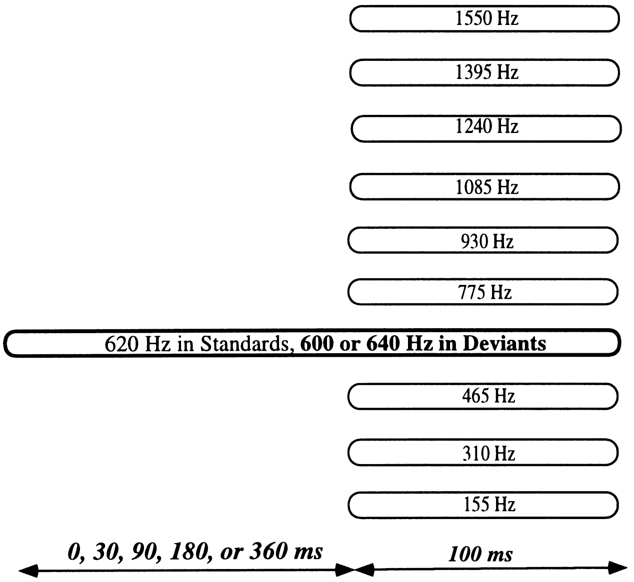

Fig. 1. Schematic illustration of the stimuli. The alternative durations of the leading partial preceding the complex-tone part are marked on the lower left and the frequencies employed in harmonic complex are marked on the right. The duration of the leading partial corresponds to the term SOA used in the text.

The interval between the onsets of two consecutive stimuli was 900 ms, which resulted in different ISIs between complex-tone offset and the leading-partial onset (from 800 ms with the 0 ms SOA to 440 ms with the 360 ms SOA) but in equal offset-to-onset ISIs between the successive complex tones.

Ten paid subjects (mean age 22.6 years, range 18–30 years; six females) were employed. During the experiment the subject sat in an acoustically and electrically shielded chamber reading a self-selected book. The EEG was recorded (0.1–100 Hz) by continuous sampling (digitization frequency 500 Hz) with Ag-AgCl electrodes at 10 scalp locations: Fpz, Fz, Cz, Pz at the midline and 6 equidistant electrodes along the coronal line connecting the mastoids through Fz. The horizontal EOG was recorded from the outer canthus of the right eye. The reference electrode was attached to the nose. The amplified and digitized EEG was stored on a computer disc for off-line averaging. Epochs with EEG change exceeding 150 μV at any EEG or EOG electrode were omitted from averaging. The analysis period was 1000 ms including a 100 ms prestimulus baseline. Data were digitally filtered (bandpass 1–30 Hz) by an FFT-filter. Grand-average ERPs were calculated in each condition for standards and deviants (ERPs to both deviant tones being pooled). The difference waves (the ERP to the standard tone subtracted from the corresponding deviant-tone ERP) were calculated.

The N1 amplitude was defined as the mean amplitude of during the 20 ms time window centered at the most negative peak of the grand-average ERPs between 70 and 150 ms for all SOAs separately for standard and deviant ERPs at Cz. The MMN peak latency was measured from the most negative peak in the grand-average difference waves at Fz/Cz between 150 and 250 ms after stimulus onset or leading-partial onset for all SOAs and after the complex-tone onset for the SOAs of 90, 180, and 360 ms. The MMN amplitudes were measured from the difference waves as the mean amplitude during the 50 ms time window centered at the peak latency.

With the 90, 180, and 360 ms SOAs, a significant N1 was elicited after complex tone onset, peaking at 124–132 ms ($t(9) = 4.1$–6.5, $P < 0.01$; one-tailed t-tests comparing N1 amplitude to zero). These N1 amplitudes significantly differed from each other ($F(2,9) = 9.7$, $P < 0.01$; one-way ANOVA with N1 amplitudes as its levels). N1 was smaller in amplitude with the 360 ms SOA than with the 90 or 180 ms SOAs (Student-Newman–Keuls, $P < 0.01$). The N1 amplitudes for standard and deviant tones did not differ from each other after any stimulus or complex-tone onsets (paired *t*-tests between corresponding standard and deviant N1s, *P*-values 0.18–0.87).

Moreover, a fronto-centrally maximal MMN to the deviant stimulus was elicited at 0, 30, 180, and 360 ms SOAs by the stimulus/leading-partial onset, and at the 360 ms SOA also by the complex-tone onset. The light grey area in Fig. 2 indicates the electrodes and conditions in which responses were statistically significant ($P < 0.05$ in one-tailed *t*-test, difference wave compared against zero). The MMN peaked between 196 and 224 ms and was maximal fronto-centrally in these conditions ($F(2,18) = 4.49–8.15$, $P < 0.05$; one-way ANOVA with MMN amplitudes along the midline as its levels). In addition, a significant polarity reversal was found at mastoid leads.

Most importantly, the leading partial and the complex tone at the 90 ms SOA and the complex tone at the 180 ms SOA triggered a negativity without the fronto-central amplitude maximum and the mastoid polarity reversal (the dark grey areas in Fig. 2; the third and fourth rows; $P < 0.05$ in one-tailed *t*-test, difference wave compared against zero). Instead, there either were no amplitude differences between the midline electrodes or, alternatively, even a negativity with parietal maximum was displayed (the complex tone with the 90 ms SOA $F(2,18) = 0.81$, $P < 0.44$; the leading partial with the 90 ms SOA $F(2,18) = 0.59$, $P < 0.52$; the complex tone with the 180 ms SOA $F(2,18) = 4.09$, $P < 0.05$; one-way ANOVA with the difference wave amplitudes along the midline as its levels). This negativity probably (at least partly) represent the N2b component.

The present data indicate, first of all, that the exogenous N1 to the complex-tone onset was remarkably enhanced when the SOA fell within the ranges of the postulated temporal window of integration [3,9,11] (Fig. 2). This suggests that the leading-partial onset primed the neural system, started the integration period, and facilitated the reactions

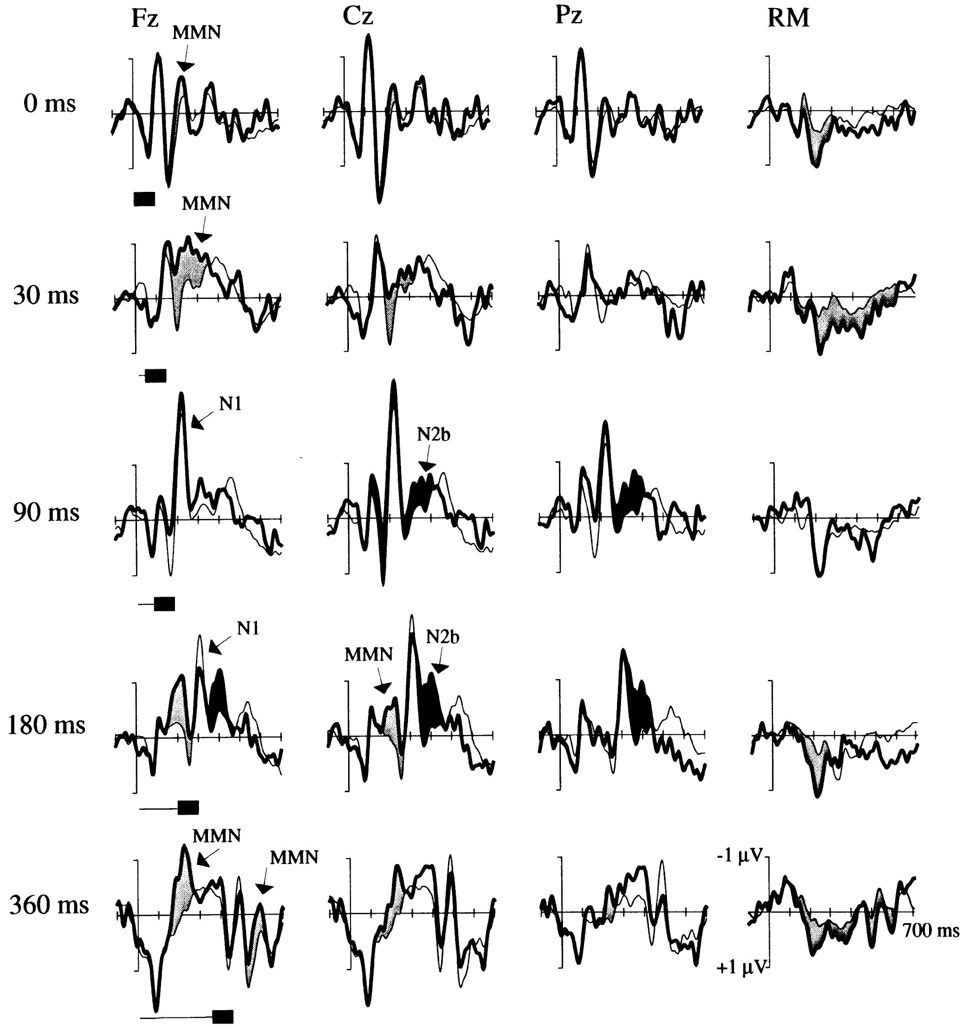

Fig. 2. ERPs at different SOAs as indicated on the left, recorded at the midline and right mastoid electrodes. The thin line represents the standard-tone response (including 620 Hz partial), the thick line the deviant tone response (including mistuned 600 or 640 Hz partial). The shadings display significant responses ($P < 0.05$; one-tailed *t*-test). The light grey area represents the frontally maximal mismatch negativity, present in 0, 30, and 360 ms SOA conditions. The dark grey area represents more broadly distributed N2b, present in 90 and 180 ms SOA conditions. Stimulus onset is indicated under the Fz responses, and the amplitude and time scales are superimposed on the ERP at the right bottom corner of the figure.

for the remaining complex part of the stimulus [7]. However, since at no SOA did the N1 amplitude differ between standard and deviant leading partials or between standard and deviant complex tones, the N1 reflects general sensitization during the integration period rather than stimulus quality [14,17].

Second, the change-specific MMN component was elicited by a ±3% change in one out of 10 harmonic partials of a complex sound, with the other nine harmonics being unchanged. Since subjects were not paying attention to auditory stimulation, this shows that even such minimal changes in complex tones may be detected on a pre-attentive level of central auditory processing: in the original pitch-adjustment experiments [5,10], the perceived pitch shift was only about 1 Hz for the 155 Hz fundamental (0.65%). When compared with previous evidence with sinusoidal tones which showed that somewhat larger (2–10%) frequency deviations are necessary for reliable MMN elicitation (e.g. [21]) the present data suggests that the auditory system reacts highly sensitively to deviances presented among harmonically complex stimuli.

Third, with the 90 and 180 ms SOAs, the deviant tones elicited more broadly-distributed negativity which probably consists of overlapping MMN and N2b components [11,16]. The presence of N2b suggests that an attention switch towards the small change in the harmonic complex occurred. That this negativity is related to deviance detection against a memory representation was confirmed by the control experiment. There 10 reading subjects were presented with natural and mistuned 180 ms SOA stimuli equiprobably. There, mistuning per se was not sufficient for triggering the response. This attention-switch interpretation is consistent with the general sensitization within the temporal window of integration as indexed by the N1. If the neural system is in an increased mode of sensitization, changes are more likely to elicit an attention switch. Intriguingly, it was elicited in the present data with 90 and 180 ms SOAs but no longer with 360 ms SOA. This suggests that with 90 and 180 ms SOAs, the central auditory system is more sensitive in switching attention towards the stimulus changes.

In summary, the present results indicate that even very small pitch deviances among harmonically complex sounds are preattentively coded by sensory memory. Furthermore, when presented among tones with SOAs falling critically within temporal window of integration these changes lead to attention switch even during the performance of a parallel task, unrelated to stimulation.

We thank Professor T.W. Picton, Dr. C. Darwin, and Dr. I. Winkler for their helpful comments during the preparation of the manuscript and S. Maury, B.A., and S. Kurjenluoma, B.A., for their help in conducting the experiments. The study was supported by the Academy of Finland, Deutscher Akademischer Austauschdienst, Jenny and Antti Wihuri Foundation, and the Finnish Psychological Society.

[1] Alho, K., Cerebral generators of mismatch negativity and its magnetic counterpart (MMNm) elicited by sound changes, Ear Hearing, 16 (1995) 38–51.

[2] Bregman, A.S., Auditory Scene Analysis, MIT Press, Cambridge, MA, 1990.

[3] Cowan, N., Attention and Memory. An Integrated Framework, Oxford University Press, Oxford, 1995.

[4] Csépe, V., Pantev, C., Hoke, M., Ross, B. and Hampson, S., Mismatch field to tone pairs: neuromagnetic evidence for temporal integration at the sensory level, Electroenceph. clin. Neurophysiol., 104 (1997) 1–9.

[5] Darwin, C.J. and Ciocca, V., Grouping in pitch perception: effects of onset asynchrony and ear of presentation of a mistuned component, J. Acoust. Soc. Am., 91 (1992) 3381–3390.

[6] Hari, R., Hämäläinen, M., Ilmoniemi, R., Kaukoranta, E., Reinikainen, K., Salminen, J., Alho, K., Näätänen, R. and Sams, M., Responses of the primary auditory cortex to pitch changes in a sequence of tone pips: neuromagnetic evidence in man, Neurosci. Lett., 50 (1984) 127–132.

[7] Loveless, N., Levänen, S., Jousmäki, V., Sams, M. and Hari, R., Temporal integration in auditory sensory memory: neuromagnetic evidence, Electroenceph. clin. Neurophysiol., 100 (1996) 220–228.

[8] Lu, Z.-L., Williamson, S.J. and Kaufman, L., Behavioral lifetime of human auditory sensory memory predicted by physiological measures, Science, 258 (1992) 1668–1670.

[9] Moore, B.C.J., An Introduction to the Psychology of Hearing, Academic Press, London, 1989.

[10] Moore, B.C.J., Glasberg, B.R. and Peters, R.W., Relative dominance of individual partials in determining the pitch of complex tones, J. Acoust. Soc. Am., 77 (1985) 1853–1860.

[11] Näätänen, R., Attention and Brain Function, Lawrence Erlbaum, Hillsdale, NJ, 1992.

[12] Näätänen, R. and Alho, K., Higher-order processes in auditory change detection, Trends Cogn. Sci., 1 (1997) 44–45.

[13] Näätänen, R., Gaillard, A.W.K. and Mäntysalo, S., Early selective-attention effect on evoked potential reinterpreted, Acta Psychol., 42 (1978) 313–329.

[14] Näätänen, R. and Picton, T.W., The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure, Psychophysiology, 24 (1987) 375–425.

[15] Nordby, H., Roth, W.T. and Pfefferbaum, A., Event-related potentials to time-deviant and pitch-deviant tones, Psychophysiology, 25 (1988) 249–261.

[16] Novak, G., Ritter, W. and Vaughan, Jr., H.G., The chronometry of attention-modulated processing and automatic mismatch detection, Psychophysiology, 29 (1990) 412–430.

[17] Picton, T.W., Campbell, K.B., Baribeau-Braun, J. and Proulx, G.B., The neurophysiology of human attention. A tutorial review. In J. Requin (Ed.), Attention and Performance VII, Erlbaum, Hillsdale, NJ, 1978, pp. 429–467.

[18] Schröger, E., A neural mechanism for involuntary attention shifts to changes in auditory stimulation, J. Cognit. Neurosci., 8 (1996) 527–539.

[19] Schröger, E., Higher-order processes in auditory change detection: response from Schröger, Trends Cogn. Sci., 1 (1997) 45–46.

[20] 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., 38 (1994) 157–167.

[21] Tiitinen, H., May, P., Reinikainen, K. and Näätänen, R., Attentive novelty detection in humans is governed by preattentive sensory memory, Nature, 370 (1994) 90–92.

[22] Yabe, H., Tervaniemi, M., Reinikainen, K. and Näätänen, R. Temporal window of integration revealed by MMN to sound omission, NeuroReport, (1997) in press.