Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Cognitive Brain Research 20 (2004) 26–36
COGNITIVE BRAIN RESEARCH
www.elsevier.com/locate/cogbrainres

Research report

Frequency discrimination at different frequency levels as indexed by electrophysiological and behavioral measures

Nikolai Novitski[a,b,c], Mari Tervaniemi[a,c,*], Minna Huotilainen[a,c], Risto Näätänen[a,c]

[a] *Cognitive Brain Research Unit, Department of Psychology, University of Helsinki, PO Box 9, FIN-00014, Helsinki, Finland*
[b] *Neuroscience Unit, Institute of Biomedicine/Physiology, University of Helsinki, Helsinki, Finland*
[c] *Helsinki Brain Research Centre, Helsinki, Finland*

Accepted 18 December 2003

Abstract

The present study systematically compared the neural and behavioral accuracy of discriminating a frequency change (“deviant”) in a repetitive tone (“standard”) across a frequency range of 250–4000 Hz. The sound structure (pure sinusoidal vs. harmonically rich tones) and the magnitude of frequency change (2.5%, 5%, 10%, 20%) were also varied. The accuracy of neural frequency-change detector was determined by comparing the auditory event-related potentials (ERP) elicited by deviant and standard stimuli in the absence of attention. In a separate behavioral task, subjects were to indicate when they noticed a frequency change. The ranges of the across-subject means of ERP parameters across the conditions were: the mismatch negativity (MMN) amplitude -0.9 to -4.9 μV, latency 125–218 ms, the P3a amplitude 0.3–3.2 μV, latency 239–304 ms. The ERP latency was shortest for the standard-stimulus frequency from 1000 to 2000 Hz suggesting that automatic frequency discrimination was the most accurate in that range. The ERP latencies and amplitudes correlated with the hit rate (HR) and reaction time (RT), with highest correlation found between the MMN amplitude and the HR ($r = 0.8$). The harmonical tones elicited MMN and P3a with shorter latencies and larger amplitudes, than did pure sinusoidal tones in all frequency bands. The results may have implication to pitch-perception theories.

Theme: Sensory systems
Topic: Auditory systems: central physiology

Keywords: Frequency discrimination; Complex sounds; Human; Auditory event-related potential (ERP); Mismatch negativity (MMN); P3a; Automatic processing

1. Introduction

Accurate pitch perception across a wide frequency range is a prerequisite for music and speech perception, for instance, in melodic, harmonic, and prosodic processing. In all natural pitched sounds, the sound spectrum consists of a time-varying pattern of multiple harmonic partials across a large frequency spectrum. However, to the date, the overwhelming majority of the experiments on pitch perception in psychoacoustics and auditory neuroscience used sinusoidal tones consisting of one harmonic partial (fundamental) only. For instance, Wier et al. [26] compared the frequency discrimination accuracy of sinusoidal tones at 8 frequency ranges from 200 to 8000 Hz by presenting 500-ms sounds once a second against the background of a low-level broadband noise. More recently, Sek and Moore [21] employed three different psychoacoustic methods (difference limen for single tones, in paired sounds, and for frequency-modulation) at six frequency ranges from 250 to 8000 Hz. Both experiments showed that frequency discrimination is the most accurate up to 2000 Hz, with the accuracy thereafter deteriorating at a rate depending on the method used. Very recently, Kishon-Rabin et al. [10] compared the frequency discrimination of 300-ms sinusoidal tones at three frequency ranges 250, 1000, and 1500 Hz with each other by using 2- and 3-interval forced-choice methods. They found that the higher the frequency was, the more accurately subjects

---

* Corresponding author. Cognitive Brain Research Unit, Department of Psychology, University of Helsinki, PO Box 9, FIN-00014, Helsinki, Finland. Tel.: +358-9-1912-9456; fax: +358-9-1912-9450.
E-mail address: mari.tervaniemi@helsinki.fi (M. Tervaniemi).

0926-6410/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.cogbrainres.2003.12.011

detected the frequency differences. In addition, their musically trained subjects outperformed the musically untrained subjects.

In auditory cognitive neuroscience, neural dynamics underlying pitch discrimination have recently been intensively investigated by recording the mismatch negativity (MMN) and P3a components of the auditory event-related potential (ERP). The MMN is elicited when the incoming stimulus with "deviant" parameter(s) is discrepant with the neural representation of the repetitive "standard" stimulus (for reviews, see Refs. [12,13,15,19]). Importantly, the subjects' conscious attention to sounds is not necessary for the MMN generation system to be activated. Thus, the MMN provides indirect index of the neural representation of any sound event, without being contaminated by, for instance, demands set by the behavioral task or fluctuation in subject's attention to, or motivation of detecting, the sound change.

Recently, it was shown that the frequency-change discrimination of tones, as reflected by the MMN, with a rich spectral structure is facilitated when compared with that of pure sinusoidal tones [21]. The MMN amplitude was larger and latency shorter in spectrally rich tones consisting of three harmonic partials than in pure tones, both with a 500 Hz fundamental. Consistent with this, the subject's behavioral responses in detecting the frequency change in a separate condition were more accurate for spectrally rich tones than for sinusoidal tones. Further, three harmonic partials were sufficient for this facilitation to occur [24]. This was established by comparing the MMN elicited by a frequency change of 2.5% in one-, three-, and five-partial harmonical sounds in separate blocks. The frequency-MMN amplitude was larger with three- and five-partial tones when compared with that with a one-partial tone, with no significant difference in MMN elicitation between three and five partial tones. This suggests that, at least with sounds with a fundamental frequency of 500 Hz, relatively few harmonic partials result in the maximal pitch discrimination accuracy.

The MMN is often followed by the P3a, which is a fronto-centrally maximal positive deflection peaking at 250–350 ms. It reflects an involuntary attention switch towards the deviant or novel sounds in inter- and intramodal ignore and attend conditions (for reviews, see Refs. [5,6]). For instance, the P3a is evoked in an experimental paradigm in which subjects are performing a visual categorization task between odd and even numbers while they are simultaneously presented with a sound sequence including a standard and several deviant frequencies [27]. In such an auditory-visual paradigm, the majority of the deviant sounds (5–80%) elicited the P3a. Correspondingly, while subjects are performing an auditory categorization task with short and long sounds (200 ms vs. 400 ms), the frequency deviant (task-irrelevant) sounds elicit the P3a component [20]. In both studies, the P3a elicitation was associated with deteriorated performance in the primary task, indicating that attention had indeed switched away from the task-relevant stimuli.

Thus, as reviewed above, it is well established that the P3a reflects a (behaviorally determined) attention switch in involuntary auditory-visual and auditory-auditory paradigms. However, it is not known how accurately the P3a latency and/or amplitude parameters and the behavioral measures (hit rate, reaction time) correspond with each other.

The present study will systematically compare the neural and behavioral accuracy of frequency discrimination across different frequency ranges (250–4000 Hz). The spectral complexity of the sound structure and the magnitude of frequency change were also varied. The neural accuracy of frequency discrimination was determined by means of ERP recordings during a parallel task unrelated to sounds. The behavioral pitch-discrimination accuracy was measured in a separate condition in which subjects were to indicate when they detected a frequency change in a pair of sounds. The different paradigms in active and ignore conditions were chosen to provide a better comparability of the present results with the ERP data to one hand and to the psychophysical frequency discrimination data to the other hand.

2. Methods

2.1. EEG experiment

2.1.1. Subjects, procedure, and stimulation

During the EEG recordings, subjects (10 healthy volunteers; 6 males; aged 19–40 years; 9 right-handed; all with normal hearing) were instructed to watch a silenced movie (with subtitles) of their own choice and not to pay attention to sound stimuli.

The sound sequences included standard tones ($p=0.76$) and eight different kinds of frequency changes: decreases and increases of 2.5%, 5%, 10%, and 20% in magnitude ($p=0.03$ each). These magnitudes of deviance cover the range from just-noticeable frequency change (eliciting an MMN with a small amplitude) until close to the ceiling-level response (the MMN somewhat overlapping with the N1).

The tones were, in separate blocks, either pure sinusoidal tones or consisted of a fundamental and its two lowest harmonics, with all these three partials with equiloud intensity (harmonical tones).

The standard tone (and the fundamental of the harmonical tones) was, in separate blocks, 250, 500, 1000, 2000, or 4000 Hz in frequency. The stimulus blocks were pseudorandom so that each deviant tone was preceded by at least one standard tone. This procedure avoids the attenuation of the MMN amplitude to a deviant tone because of the formation of a memory trace for the deviant frequency.

Stimuli of 100 ms in duration (including 10 ms rise and fall times) were binaurally presented via headphones at an intensity of 50 dB above the individually determined hearing threshold (separately performed at each frequency range with both sound types). The constant stimulus onset asynchrony (SOA) was 400 ms. Each of the 10 conditions (both sinusoidal and harmonical tones, each at 250, 500, 1000, 2000, and 4000 Hz standard-stimulus frequency levels) were administered in two blocks of 1700 sounds presented in a randomized order in two sessions. The sessions were recorded in separate days and lasted 4 h including attaching the electrodes and their removal.

2.1.2. EEG recording and analysis

The EEG (passband 0.1–40 Hz, sampling frequency 500 Hz) was recorded in an acoustically dampened and electrically shielded room. The EEG was recorded with Ag–AgCl electrodes from scalp locations as follows: Fpz, Fz, Cz and Pz (10–20 system), left and right mastoids (LM and RM, respectively), and L1, L2, R1, and R2 along the coronal arch equidistantly connecting the Fz with LM (L1, L2) and RM (R1, R2). In addition, horizontal eye movements were monitored with bipolar electrodes attached to the right and left outer canthi and vertical eye movements with electrodes attached to the upper and lower outer canthi of the right eye. EEG was referenced to the nose.

The EEG was divided into 500-ms epochs (− 100–400 ms) and separately averaged for each stimulus type except for the standards following deviants. To improve the signal-to-noise ratio, the deviant-stimulus ERPs for the decreases and increases of frequency of the same magnitude were pooled together. All epochs including voltage changes exceeding 100 μV were automatically rejected. Frequencies higher than 12 Hz and lower than 1 Hz were filtered out to stabilize the amplitude and latency quantification [19]. Difference waveforms (deviant-tone ERP minus standard-tone ERP) were calculated.

2.1.3. ERP quantification

The MMN was quantified by determining the MMN peak amplitude and latency from the Fz grand-average difference waves separately for each deviant as the most negative peak between 100 and 300 ms in the individual difference waves.

The statistical significance of the MMN was tested with one-tailed t-tests by comparing the MMN amplitude at Fz and at the right mastoid lead with zero. Previous studies showed that, when a nose reference is used, the MMN has its maximum amplitudes at Fz (the most negative value) and at the mastoid leads (the most positive value) (see, e.g.,) [1]. In further analyses, the Fz values were used after rereferencing the data against the average of the mastoid-lead amplitudes. This procedure was used to project the strength and timing of the MMN from several electrodes with opposite polarities (fronto-centrally negative values, mastoidally positive values) to one topographical point [19].

The P3a was quantified by determining the P3a peak latency from the Cz grand-average difference waves separately for each deviant as the most positive peak between 200 and 400 ms. The statistical significance of the P3a was tested with one-tailed *t*-tests by comparing the P3a amplitude at Cz with zero.

2.1.4. Statistical analyses

Four-way ANOVAs were used to determine whether the component type (MMN or P3a), sound structure (levels: Sinusoidal/Harmonical), magnitude of deviance (levels: 2.5%, 5%, 10%, and 20%), and the standard-stimulus frequency (250, 500, 1000, 2000, 4000 Hz) affected the ERP amplitude and latency at the Fz electrode (for the MMN) or Cz (for P3a). The Greenhouse-Geisser correction was applied whenever the factor contained more than two levels. In the latter case, the Tukey HSD test was applied as a post-hoc comparison. In the case of significant interaction, the data were divided accordingly to obtain the main effects of each factor on each level of the other factor [18]. Bonferroni correction was applied for the main effects to compensate for the repetitive measures.

2.2. Behavioral experiment

In the behavioral session, the subjects who had participated in the EEG recordings were instructed to indicate whether the tones delivered in pairs differed in pitch. The pairs had 300 ms silent within-pair interval and were presented with a 1900-ms between-pair interval. Altogether 144 pairs were presented in one block, with 16 pairs of identical standard stimuli and 32 pairs with the standard stimuli paired with the deviant stimulus of each level (16 upward and 16 downward deviants). The stimuli in the behavioral session were the same as in the ignore sessions. There were 10 experimental blocks in the behavioral session, one per each condition (five frequency bands for both harmonical sinusoidal tones), resulting in experimental time of 1 h. The session also included a 10-min break in the middle.

Subjects were instructed to press one button with the index finger of the dominant hand if the tones in the pair were identical and another button with the middle finger if they were different. Subjects were not informed either of the nature of the sounds or of their proportion. No practice was given before the actual experiment. The behavioral session was always administered after the EEG recording sessions to avoid the carry-over effects of attention on the ERPs.

The response “different” for a standard-deviant pair was accepted as a hit if it was given before the onset of the next stimulus pair. The response “different” after the presentation of a pair of identical standard stimuli was counted as a false alarm.

The hit rate (HR) was separately calculated for each deviant pair as a proportion of the hits to the total number of the pairs containing this deviant type within a block. D' sensitivity measure was calculated according to the formula:

$d' = Z(\mathrm{HR_p}) - Z(\mathrm{FA})$, in which $\mathrm{HR_p}$ is a pooled average hit rate of four deviance levels within a block.

The HR and RT were analyzed with three-way ANOVA (factors: sound structure, magnitude of deviance, and standard-stimulus frequency). The principles of the statistical analysis were the same as for the ERP parameters (see above). D' was analyzed with two-way ANOVA (factors: sound structure, standard-stimulus frequency).

2.2.1. Correlation between ERP and behavioral indices

The correspondence between the present ERP (MMN and P3a amplitude and latency) and behavioral (HR, RT)

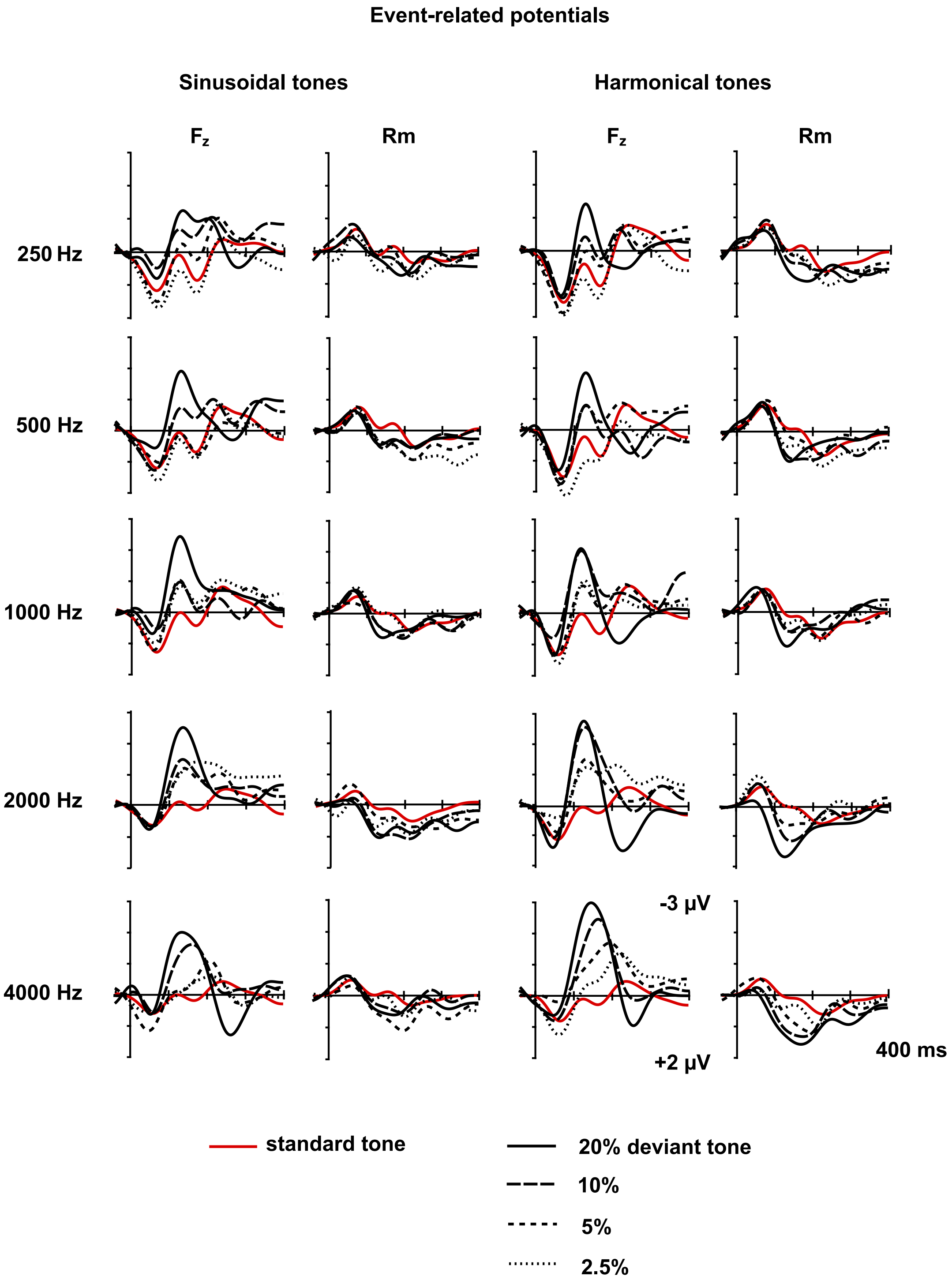

Fig. 1. Grand-average ERPs elicited by standard tones (grey line) and deviant tones at the Fz and Rm electrodes with the sinusoidal (left panel) and harmonical tones (right panel) with different standard-stimulus frequency levels and with different magnitudes of deviance as indicated below the figure.

indices was determined by calculating Pearson's product-moment correlation coefficients.

3. Results

Frequency changes elicited the MMN as indexed by the negative displacement of the ERPs for deviant tones at the Fz electrode and by the accompanying positivity at the corresponding latency at the right mastoid for all standard-stimulus frequencies (250–4000 Hz) (Fig. 1). The MMN amplitude was significantly different from zero ($p<0.05$) as measured either at Fz (sinusoidal tones: 250 Hz, 5% deviance) or at mastoids (sinusoidal tones: 250 Hz 2.5% deviance, 500 Hz 5% deviance, 4000 Hz 5% deviance; harmonical tones: 500 Hz, 2.5% deviance) or at both Fz and mastoids (the rest of conditions). The nose-referenced MMN amplitude ranged from −0.6 to −3.6 μV (standard error SE 0.2–0.7 μV) at Fz and from 0.5 to 2.1 μV (SE 0.2–0.4 μV) at mastoids. The mastoid-referenced MMN at Fz ranged from −0.9 to −4.9 μV (SE 0.2–0.6 μV) (Fig. 2). The MMN latency varied in the 125–218 ms range (SE 3.5–15.3 ms). Fig. 3 (upper row) shows the mastoid-referenced MMN amplitudes and latencies.

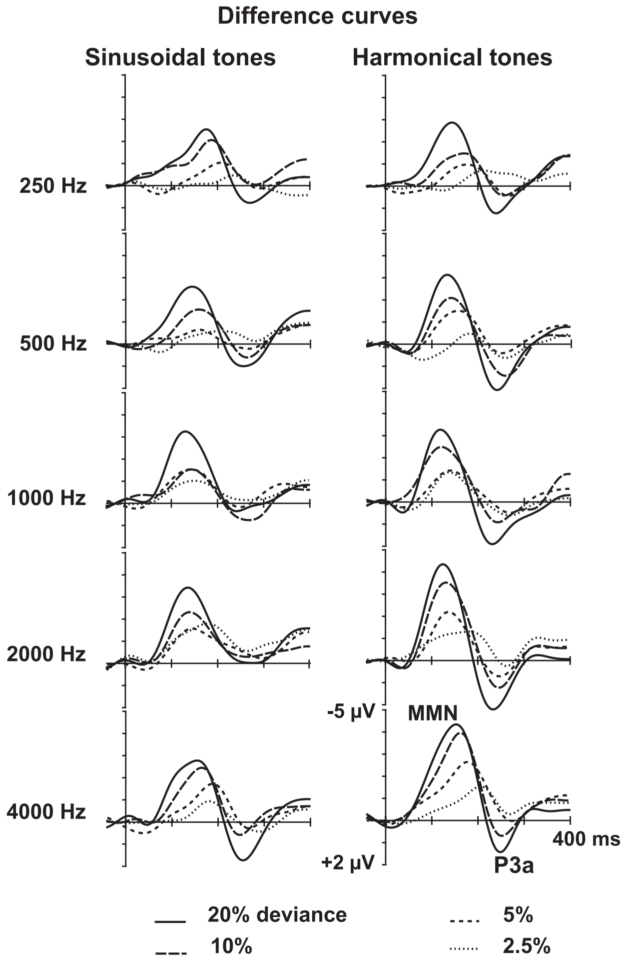

Fig. 2. The difference curves (deviant minus standard tone ERPs) at the Fz electrode with the sinusoidal (left panel) and harmonical tones (right panel) at different standard-stimulus frequency levels and with different magnitudes of deviance as indicated below the figure. These data, used in statistical analysis, were rereferenced to the mastoid-lead data.

In the majority of conditions, the MMN was followed by the P3a. The P3a amplitude ranged from 0.3 to 3.2 μV (SE 0.2–0.7 μV) and the P3a latency range was 239–304 ms (SE 5–17.6 ms). Fig. 3 (middle row) shows the P3a amplitudes and latencies. The values which are significantly different from zero ($p<0.05$) are plotted as filled symbols and not significant ones with empty symbols.

3.1. Standard-stimulus frequency

There were significant main effects of standard-stimulus frequency on the ERP latency [$F(4,36)=11.7$, $p<0.001$] and amplitude [$F(4,36)=8.3$, $p<0.001$] and on the HR [$F(4,36)=9.0$, $p<0.001$]. These findings will be separately described in the following.

In general, the ERP latencies were shorter at 1000 and 2000 Hz than at 250 and 4000 Hz standard-stimulus frequencies, and at 1000 Hz than at 500 Hz standard-stimulus frequency ($p<0.05$). Moreover, there was an interaction between the standard-stimulus frequency and the ERP component on the ERP latency [$F(4,36)=5.2$, $p<0.01$]. This was caused by the MMN latency being shorter at the 1000 and 2000 Hz than at 250 and 4000 Hz standard-stimulus frequencies as well as at 1000 Hz than at 500 Hz standard-stimulus frequency [simple main effect $F(4,36)=12.3$, $p<0.001$; $p<0.05$ in all comparisons], whereas the P3a latency was shorter at 1000 Hz than at 250 and 4000 Hz standard-stimulus frequencies [simple main effect $\mathrm{F}(4,36)=6.6$, $p<0.01$; $p<0.05$ in all comparisons].

In addition, there was an interaction between the standard-stimulus frequency and the deviance on latency [$F(12,108)=2.9$, $p<0.05$]. The main effects of standard-stimulus frequency were found at 2.5% [$F(4,36)=8.6$, $p<0.01$] and 5% [$F(4.36)=8.8$, $p<0.001$] deviance levels. Both MMN and P3a latency at the 2.5% deviance level was longer at 250, 500, and 4000 Hz than at 1000 Hz standard-stimulus frequencies ($p<0.05$ in all comparisons). Correspondingly, the MMN and P3a latency at the 5% deviance level was longer at 250 and 4000 Hz than at 2000 and 1000 Hz, and at 4000 than 500 Hz standard-stimulus frequency ($p<0.05$ in all comparisons).

The ERP amplitudes were smaller at the 250, 500, and 1000 Hz in comparison with 2000 and 4000 Hz standard-stimulus frequencies ($p<0.05$ in all comparisons). An interaction between the standard-stimulus frequency and the ERP component on the ERP amplitude was revealed [$F(4,36)=3.6$, $p<0.05$]. The main effect of the standard-stimulus frequency on the MMN amplitude was found [$F(4,36)=13.1$, $p<0.001$; $p<0.001$), being caused by a

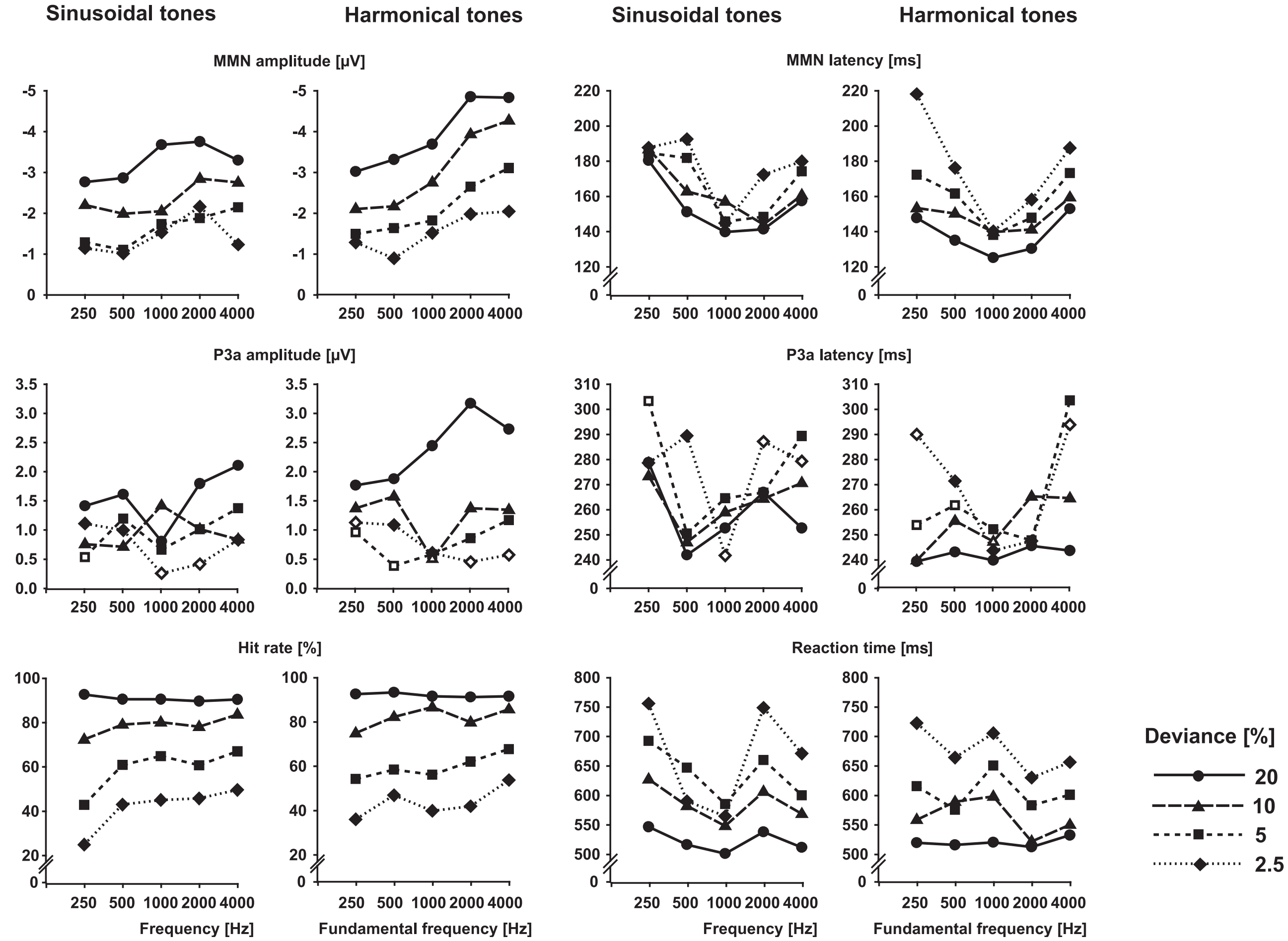

Fig. 3. The MMN amplitude and latency (top panel), the P3a amplitude and latency (middle panel), and the behavioral indices (HR and RT, bottom panel) as a function of frequency (Sinusoidal tones) and fundamental frequency (Harmonical tones) of the standard tone. For the MMN and P3a amplitudes the filled symbols mark values which are different from zero ($p < 0.05$) in contrast to empty symbols indicating non-significant values.

smaller MMN amplitude at 250 and 500 Hz than at 2000 and 4000 Hz standard-stimulus frequencies ($p < 0.05$ in all comparisons).

The HR was less accurate at the 250 Hz standard-stimulus frequency than at the 500 and 4000 Hz standard-stimulus frequencies ($p < 0.05$ in all comparisons). There was also an interaction between the standard-stimulus frequency and deviance [$F(12,108) = 2.9$, $p < 0.05$]. The simple main effect of standard-stimulus frequency was found for 2.5% [$F(4,36) = 5.8$, $p < 0.01$] and 5% [$F(4,36) = 7.24$, $p < 0.01$] deviances. The HR at the 2.5% deviance level was smaller at the 250 Hz than at the 500, 2000, and 4000 Hz standard-stimulus frequencies ($p < 0.05$ in all comparisons). The HR at the 5% deviance was smaller at the 250 Hz than 500, 1000, 2000 and 4000 Hz standard-stimulus frequencies ($p < 0.05$ in all comparisons).

An interaction between the effects of the sound type and the standard-stimulus frequency on the RT was also found [$F(4,36) = 4.3$, $p < 0.05$]. There was no simple main effect of the standard-stimulus frequency, however.

3.2. Magnitude of deviance

There was a significant main effect of deviance on the ERP latency [$F(3,27) = 13.4$, $p < 0.001$] and amplitude [$F(3,27) = 72.3$, $p < 0.001$] and on the HR [$F(3,27) = 13.8$, $p < 0.001$] and RT [$F(3,27) = 20.05$, $p < 0.01$]. These findings will be separately described in the following.

The latency of the ERPs was longer for the 2.5% than for the 20% deviance and for the 5% longer than for the 10% and 20% deviations ($p < 0.05$ in all comparisons). The effect of deviance on the ERP latency interacted with the standard-stimulus frequency [$F(12, 108) = 2.9$, $p < 0.05$]. The main effects of deviance were found for the 250 [$F(3,27) = 10.0$, $p < 0.001$], 500 [$F(3,27) = 10.6$, $p < 0.01$] and 4000 Hz [$F(3,27) = 12.3$, $p < 0.001$] standard-stimulus frequencies. Both MMN and P3a latency within the 250 Hz standard-stimulus frequency were longer at 2.5% than at 10% and 20% and at 5% it was longer than at 20% deviance level ($p < 0.05$ in all comparisons). The MMN and P3a latency for 500 Hz standard-stimulus frequency level was longer for the 5% and 2.5% than for the 20% and for the 2.5% than

10% deviance ($p<0.05$ in all comparisons). Finally, the ERP latency within the 4000 Hz standard-stimulus frequency was longer for the 2.5% and 5% than for the 10% and 20% deviance ($p<0.05$ in all comparisons).

The amplitude of the ERPs was smaller for the 2.5% and 5% than for the 10% and 20% deviance and for the 10% than 20 % deviance ($p<0.05$ in all comparisons). There was an interaction between the effect of the deviance and the effect of the ERP component on the ERP amplitude [$F(3,27)=3.5$, $p<0.05$]. The simple main effect of deviance was present for both the MMN [$F(3,27)=56.3$, $p<0.05$] and P3a [$F(3,27)=10.9$, $p<0.05$] amplitudes.

The MMN amplitude was smaller for the 2.5% and 5% than for the 10% and 20% deviances and for the 10% than 20% deviance ($p<0.05$ in all comparisons). The P3a amplitude was larger for the 20% than 2.5%, 5% and 10% deviances ($p<0.05$ in all comparisons).

There was also an interaction between the effect of deviance and that of the sound type on the ERP amplitude [$F(3,27)=5.8$, $p<0.01$]. The simple main effects of deviance were present in both harmonical [$F(3,27)=56.5$, $p<0.001$] and sinusoidal [$F(3.27)=37.3$, $p<0.001$] conditions ($p<0.05$). The ERP amplitude for the sinusoidal tones was smaller for the 2.5% than 10% and for the 2.5%, 5%, and 10% than for the 20% deviance ($p<0.05$ in all comparisons). The ERP amplitude for the harmonical tones was smaller for the 2.5% and 5% than for the 10% and 20% deviances, and for the 10% than 20% deviance ($p<0.05$).

The HR was less accurate at the 2.5% than 10% and 20% as well as at 5% than at 20% deviance level ($p<0.05$). The interaction between the effects of deviance and the effect of standard-stimulus frequency on HR was obtained [$F(12,108)=2.9$, $p<0.05$]. The main effect of deviance was found at the 250 [$F(3,27)=19.9$, $p<0.001$], 500 [$F(3,27)=10.6$, $p<0.01$], and 1000 Hz [$F(3,27)=12.3$, $p<0.01$] standard-stimulus frequencies. The HR for the 250 Hz standard-stimulus frequency was less accurate at the 2.5% and 5% than at the 10% and 20% deviance levels ($p<0.05$ in all comparisons). The HR within the 500 and 1000 Hz standard-stimulus frequencies was less accurate for the 2.5% than 10% and 20%, and for the 5% than 20% deviance ($p<0.05$ in all comparisons).

The RT was longer for the 2.5% than 10% and 20%, and for the 5% than 20% deviance ($p<0.05$ in all comparisons).

3.3. Sound structure

The harmonical sounds elicited ERPs with a shorter latency [$F(1,9)=14.5$, $p<0.01$] and a larger amplitude [$F(1,9)=28.1$, $p<0.001$] than did the sinusoidal sounds (Fig. 3). Additionally, there was an interaction of the effects of sound type and deviance on the ERP amplitude [$F(3,27)=5.8$, $p<0.01$]. The main effect of the sound structure was found for the 20% deviance [$F(1,9)=28.7$, $p<0.001$].

An interaction between the effects of the sound type and standard-stimulus frequency on the RT was also obtained [$F(4,36)=4.3$, $p<0.05$], resulting from longer RT for the sinusoidal than for the harmonical sounds for the 2000 Hz standard-stimulus frequency [$F(1,9)=9.6$, $p<0.05$].

The discrimination of the harmonical tones resulted in a higher d' sensitivity than did the sinusoidal ones [$F(1,9)=8.4$, $p<0.05$], but this effect was present only with the 250 and 500 Hz standard-stimulus frequencies (Fig. 4). Thus, an interaction between the effects of the sound type and standard-stimulus frequency was found [$F(4,36)=4.4$, $p<0.01$].

3.4. MMN vs. P3a

The latency of the P3a was longer of course than that of the MMN [$F(1,9)=879.5$, $p<0.001$]. Further, the P3a amplitude was smaller than that of the MMN [$F(1,9)=24.1$, $p<0.001$] (Figs. 2 and 3).

There was an interaction between the effects of the component and the standard-stimulus frequency on ERP latency [$F(4,36)=5.2$, $p<0.01$]. Simple main effects of the component were found for the 250 [$F(1,9)=224.0$, $p<0.001$], 500 [$F(1,9)=305.7$, $p<0.001$], 1000 [$F(1,9)=461.5$, $p<0.001$], 2000 [$F(1,9)=501.7$, $p<0.001$] and 4000 Hz [$F(1,9)=606.4$, $p<0.001$] standard-stimulus frequencies.

There was also an interaction between the ERP component and the standard-stimulus frequency on the ERP amplitude [$F(4,36)=3.6$, $p<0.05$]. The MMN amplitude was larger than that of the P3a for the higher standard-stimulus frequencies of 1000 [$F(1,9)=25.8$, $p<0.01$], 2000 [$F(1,9)=34.6$, $p<0.001$] and 4000 Hz [$F(1,9)=29.2$, $p<0.001$].

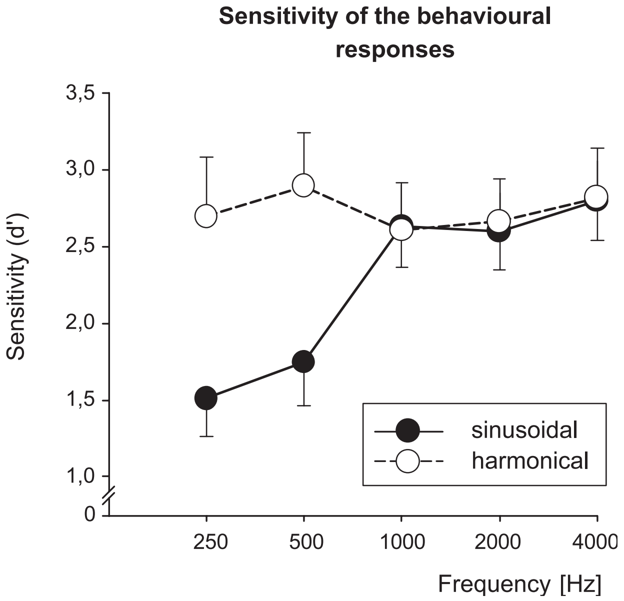

Fig. 4. The sensitivity (d') of the behavioral responses for harmonical and sinusoidal tones.

Moreover, an interaction between the effects of component and deviance on the ERP amplitude was found [$F(3, 27) = 3.5$, $p < 0.05$]. The MMN amplitude was larger than that of the P3a for the 10% deviance [$F(1,9) = 44.5$, $p < 0.001$].

3.5. Correlation between the ERP indices and behavioral performance

Fig. 5 displays the mean MMN and P3a amplitudes together with the corresponding behavioral results as well as the correlation coefficients. The MMN and P3a amplitudes significantly correlated with the HR and RT measures (the P3a latency vs. the HR and RT: $p < 0.01$; in all other comparisons $p < 0.001$).

The correlation coefficients reaching the statistical significance at each standard-stimulus frequency are displayed in Table 1. To summarize, the MMN amplitude and the HR correlated significantly at all standard-stimulus frequencies. For the other measures, no significant correlation was obtained for the 1000 Hz. In addition, the MMN amplitude and the RT did not correlate at the 250 Hz. The P3a amplitude and latency significantly correlated with HR and RT at the 4000 Hz range. In addition, the correlation

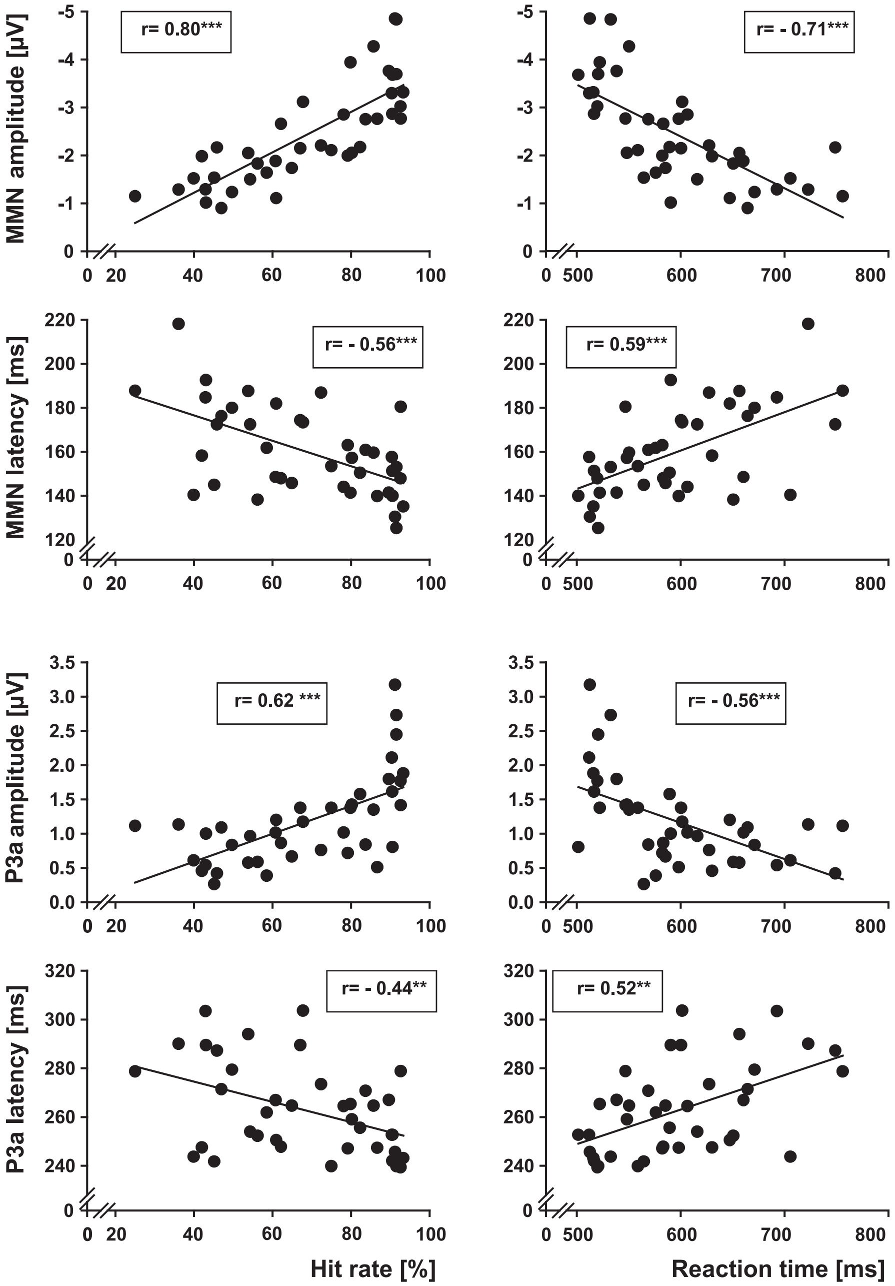

Fig. 5. The correspondence between the behavioral measures (hit rate, HR, left panel; reaction time, RT, right panel) and the MMN amplitude (top panel) and latency (second panel), the P3a amplitude (third panel) and latency (bottom).

Table 1
Correlation coefficients between the MMN parameters and the hit rate (HR) and reaction time (RT) separately at each standard-stimulus frequency

Standard stimulus	Amplitude		Latency	
	HR	RT	HR	RT
MMN				
250 Hz	− 0.97***	0.91**	− 0.65	0.77*
500 Hz	− 0.94***	0.89**	− 0.90**	0.73*
1000 Hz	− 0.87**	0.69	− 0.22	0.04
2000 Hz	− 0.88**	0.84**	− 0.89**	0.91**
4000 Hz	− 0.85**	0.82*	− 0.97***	0.94***
P3a				
250 Hz	0.58	− 0.62	− 0.57	0.71*
500 Hz	0.62	− 0.41	− 0.88**	0.48
1000 Hz	0.59	− 0.48	0.14	− 0.09
2000 Hz	0.83*	− 0.74*	− 0.18	0.60
4000 Hz	0.73*	− 0.78*	− 0.77*	0.74*

* $p<0.05$.
** $p<0.01$.
*** $p<0.001$.

between the P3a amplitude and the HR and RT was significant with the 2000 Hz as well as between the P3a latency and the HR at the 500 Hz and the P3a latency and the RT at the 250 Hz.

4. Discussion

The present study was conducted to compare the neural and behavioral accuracy of frequency discrimination across different frequency levels (250–4000 Hz). The sound structure (pure sinusoidal vs. harmonically rich tone) and the magnitude of frequency change (2.5–20%) were also manipulated. The accuracy of frequency discrimination was measured by recording the ERPs during a parallel task unrelated to the sounds and by using behavioral task, in which subjects were to indicate when they noticed a frequency change in sound pairs. The data obtained demonstrated that the ERP (MMN and P3a) amplitude and latency were differently affected by standard-stimulus frequency. The ERP amplitude increased from 250 to 4000 Hz, whereas the ERP latency had minimum at 1000 and 2000 Hz and was prolonged at higher and lower frequencies. Further, larger pitch changes caused an increase in the ERP amplitude and a shortening of the ERP latency, as well the increase in the HR and the shortening of the RT. The harmonical tones elicited MMNs and P3s with larger amplitude and shorter latencies than pure sinusoidal tones, supporting the results obtained with a single standard-stimulus frequency level in an earlier study. Finally, correlation across conditions between ERP and behavioral parameters has been observed.

The ERP (MMN and P3a) latency data indicate that the latencies were shortest for the 1000 and 2000 Hz standard-stimulus frequencies. This result confirms and expands the previous behavioral findings, according to which at comparable frequency levels frequency discrimination is the most accurate up to 2000 Hz, and thereafter, differentially deteriorates as a function of the method used [21,26]. In contrast, the MMN (but not P3a) amplitude was systematically increased with frequency. The higher amplitude of the MMN with a high frequency might reflect a change in the spatial configuration of the source, known to vary with frequency [1], rather than a frequency-related change in the magnitude of activation. At the same time, the MMN latency is not sensitive to the dipole orientation. The alternative explanation for the discrepancy might be that the pitch-discrimination mechanisms are different at higher and lower frequencies (see below). The HR of the behavioral discrimination in our study was smallest at the 250 Hz (especially with minor deviations) and more or less constant at higher standard-stimulus frequencies. Although these data may suggest some ceiling effect, the general standard-stimulus frequency effect is similar for both ERP and behavioral data.

The accuracy of pitch discrimination in present experiment was lower than in the psychophysical studies by Wier et al. [26] and Sek and Moore [21]. The subjects in our study were naïve to the procedure, not trained, got no feedback of their performance, and were not instructed on the type of sounds or their ratio in the test sequence. The test sounds were of short duration and were presented via headphones. In addition, we used same-different procedure instead of two/three interval forced-choice task, which is more common in the difference-limen measurements. All these factors impoverish performance. Thus, the test was more difficult than in many other studies and therefore the ceiling effect was mostly avoided. The design of the behavioral test was a compromise between the comparability with our ERP data on one hand and with the conventional psychophysical procedures on the other hand.

Additionally, the characteristics of the sound onset and offset, known to be important for ERP elicitation [3], might have affected the present results differentially at the different frequency levels. Namely, at higher frequencies, the sound cycles are shorter than at the lower frequencies and may thus result in different amounts of spectral splatter with the same rise-fall time. However, defining the rise-fall time as a fixed number of cycles rather than as a fixed time period would add variability in the total duration of the stimuli and its temporal characteristics. With a minimum of five cycles for onset-offset (like in the present study), the width of the splatter was negligibly small. Therefore the uniform rise-fall time for all the frequencies used did not bias the results.

The MMN in our study was quantified in a traditional way as a negative deflection on a difference wave between the deviants and the standards of the same oddball sequence (Fig. 1). This way of comparison may lead to the enhancement of the early part of the mismatch negativity (MMN) by the N1 due to the stronger refractoriness of the response to standards than to deviants. The alternative way is to make subtraction between the deviant and the same sound, when presented as a standard in a separate block. Recently, to overcome this problem, Jacobsen and Schröger [8] presented standards in a special control sequence with the same probability as the deviants in the experimental sequence. However, the addition of control conditions would have substantially increased the present experimental time which already consisted of three sessions. Also, the MMN latency of in our data was somewhat longer than the N1 latency at the corresponding frequencies [17].

The present design also permits comparisons between two subsequent ERP components, which are elicited as the response to the pitch change, the MMN and P3a. While MMN was proved to be the reflection of automatic memory-based change detection [5,13], P3a is believed to correlate with involuntary attention shift in response to the change [5,13]. In the present data, the MMN was significantly different from zero at all deviance levels and its amplitude started to increase from 5% deviance level. In contrast, the P3a was not always significant at the lowest deviance levels and its amplitude increased only from 10% to 20% deviance level. The frequency-dependent course of the P3a latency is similar to that of the MMN, while the frequency-dependent courses of their amplitudes were different. The MMN amplitude increased with frequency, whereas the P3a amplitude remained stable (Fig. 3). Finally, the correlation coefficients between the P3a parameters and behavioral indices (HR and RT) were smaller than for the corresponding parameters of the MMN (Fig. 5). This supports the view of a threshold-like mechanism behind the P3a elicitation, in other words, that for P3a elicitation, a relatively large deviance is necessary.

Changes in the harmonical tones elicited ERP with longer latency and larger amplitude than those in sinusoidal tones (Fig. 2). Since no sound type-frequency interaction was found we can conclude that this finding is equally valid for all standard-stimulus frequency levels from 250 to 4000 Hz. Thus, the earlier finding of the harmonical sounds being superior to sinusoidal only at 500 Hz [23,24] is confirmed and extended to a wide spectral band. Interestingly, the difference between one-harmonic (sinusoidal) and three-harmonic tones was found for the standard-stimulus fundamental frequency as high as 4000 Hz, even though the highest harmonic (12 000 Hz) of it is already close to the limits of human hearing. The robustness of the sound-structure effects at the higher frequencies supports the hypothesis that the familiarity (due to life-long experience) of complex sounds in contrast to the pure tones accounts for more accurate discrimination. It has been shown before that the addition of two partials to the pure tone is enough to cause the augmentation of MMN amplitude [24]. The addition of one partial is probably not enough because of the octave relationship between the first overtone and the fundamental, which may reduce the perceptual contrast. For example, the octave effect on the MMN has been shown in an fMRI study, in which a 50% of frequency deviance caused more activation than a 100% deviance [14].

Furthermore, the present results also extend the previous ones [23,24] by showing that the P3a latency is shortened with spectrally rich sounds when compared with pure sinusoidal tones. This might be due to the larger frequency deviance employed in the present (2.5–20%) than in the previous study (2.5–10%) [24]. In other words, as previously shown, the P3a is elicited by an attention-catching deviance [4]. In previous studies comparing pitch discrimination in sinusoidal vs. spectrally rich sounds, the 10% frequency was not large enough for the P3a elicitation. Thus, the present study is the first one to report such a relationship between the sound structure (sinusoidal vs. spectrally rich) and the P3a parameters.

The present within-subject design also allowed reliable correlative analyses between the ERP and behavioral indices of the pitch discrimination accuracy. It was found that the MMN and P3a amplitudes as well as latencies highly significantly correlated with the subject's behavioral discrimination accuracy as indicated by the HR and RT at each standard-stimulus frequency level (Fig. 5).

Thus, the correlation coefficients obtained between the MMN and P3a amplitudes and the subjects' HR and RT corroborate and also extend the previous results obtained with sinusoidal tones at a very limited standard-stimulus frequency [2,9,25]. This is an encouraging result when educational applications of the ERP recordings are considered. For instance, in exploring the limits of expertise of a population with specific, e.g., musical, talents (e.g., Ref. [22]), using stimulation with acoustic properties relevant to the ability under interest is of essential importance. Furthermore, since the present data indicated that the facilitation caused by the spectrally rich sound structure relative to the processing of pure tones is equally present for behavioral and (passively recorded) ERP measures. Thus, if active paradigms are not suitable for practical or methodological reasons (e.g., in infants) then the MMN recordings might offer a probe to the integrity and limits of neural auditory discriminative functions.

In addition, our data might have implications to the pitch perception theory. Currently it is generally accepted that both the "rate" and "place" mechanisms are employed in encoding the frequency of the sounds in the auditory system [11]. The rate mechanism is used in the coding of lower frequency sounds and the place mechanism is dominant in the perception of the higher-frequency sounds. However, there is no agreement with regard to the borderline between these registers. For pure tones, it might lay at around 5 kHz [11]. For frequency modulated sounds with the rate of 10 Hz and higher, the place mechanism was suggested [11]. For periodical tonal bursts the upper limit of the rate mechanism is 200–300 Hz, while the lower limit of the place mechanism is around 2000 Hz, leaving an area in between representing a mixture of the two mechanisms, with 600 Hz being an equilibrium point [16]. For periodical clicks, the border between the phase-locking and place mechanism is between 100 and 200 Hz. Moore's [11] model of the pitch perception of complex tones integrates both mechanisms in a multi-level processing scheme. The data on the phase-locking ability of the auditory nerve in animals (squirrel monkeys) support the 5 kHz limit. However, the properties of the neuronal network in audition may vary between the species [7]. Therefore, the data on the properties of the human neural frequency discrimination in vivo, provided by the MMN measurement, are of great value for the pitch-perception theory.

Although our present data cannot directly address the rate vs. time mechanism controversy, they show the difference between the frequency bands in the neural pitch discrimination. The change-related ERPs recorded at lower frequencies (250 and 500 Hz) are clearly different from those recorded at higher frequencies (2000 and 4000 Hz). The most prominent is the discrepancy in the way how the MMN amplitudes and latencies change as a function of frequency. From 250 to 1000 Hz the amplitude of the MMN is rising and its latency shortening with the increase of frequency. However, on the further increase of frequency the amplitude continues to grow, while the latency starts to lengthen (Fig. 3). The changing point at 1000 Hz may indicate a transition between two different neuronal mechanisms of pitch discrimination, presumably the rate and place mechanisms. Thus, our data indirectly support the view of a border between the rate and time mechanisms being at 500–1000 Hz (cf. Ref. [16]).

Acknowledgements

The study was supported by the Academy of Finland (projects 53294 and 78265), Centre of International Mobility (Finland), Deutsche Akademische Austauschdienst (Germany), and Marie Curie Individual fellowship (European Commission).

References

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

[2] E. Amenedo, C. Escera, The accuracy of sound duration representation in the human brain determines the accuracy of behavioural perception, Eur. J. Neurosci. 12 (2000) 2570–2574.

[3] S. Biermann, P. Heil, Parallels between timing of onset responses of single neurons in cat and of evoked magnetic fields in human auditory cortex, J. Neurophysiol. 84 (2000) 2426–2439.

[4] C. Escera, K. Alho, I. Winkler, R. Näätänen, Neural mechanisms of involuntary attention to acoustic novelty and change, J. Cogn. Neurosci. 10 (1998) 590–604.

[5] C. Escera, K. Alho, E. Schröger, I. Winkler, Involuntary attention and distractibility as evaluated with event-related brain potentials, Audiol. Neurootol. 5 (2000) 151–166.

[6] D. Friedman, Y.M. Cycowicz, H. Gaeta, The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty, Neurosci. Biobehav. Rev. 25 (2001) 355–373.

[7] R.S. Heffner, G. Koay, H.E. Heffner, Audiograms of five species of rodents: implications for the evolution of hearing and the perception of pitch, Hear. Res. 157 (2001) 138–152.

[8] T. Jacobsen, E. Schröger, Is there pre-attentive memory-based comparison of pitch? Psychophysiology 38 (2001) 723–727.

[9] M. Jaramillo, P. Paavilainen, R. Naatanen, Mismatch negativity and behavioural discrimination in humans as a function of the magnitude of change in sound duration, Neurosci. Lett. 290 (2000) 101–104.

[10] L. Kishon-Rabin, O. Amir, Y. Vexler, Y. Zaltz, Pitch discrimination: are professional musicians better than non-musicians? J. Basic. Clin. Physiol. Pharmacol. 12 (2001) 125–143.

[11] B.C.J. Moore, An Introduction to the Psychology of Hearing, Academic Press, London, 1997, p. 373.

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

[13] R. Näätänen, I. Winkler, The concept of auditory stimulus representation in cognitive neuroscience, Psychol. Bull. 125 (1999) 826–859.

[14] B. Opitz, T. Rinne, A. Mecklinger, D.Y. von Cramon, E. Schröger, Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results, NeuroImage 15 (2002) 167–174.

[15] T.W. Picton, C. Alain, L. Otten, W. Ritter, A. Achim, Mismatch negativity: different water in the same river, Audiol. Neurootol. 5 (2000) 111–139.

[16] J. Pierce, Introduction to pitch perception, in: P.R. Cook (Ed.), Music, Cognition, and Computerized Sounds, MIT Press, Cambridge, MA, 1999, pp. 57–70.

[17] T.P. Roberts, D. Poeppel, Latency of auditory evoked M100 as a function of tone frequency, NeuroReport 7 (1996) 1138–1140.

[18] M.J. Roberts, R. Russo, A Student's Guide to Analysis of Variance, Routledge, London, 1999. 265 pp.

[19] E. Schröger, On the detection of auditory deviations: a pre-attentive activation model, Psychophysiology 34 (1997) 245–257.

[20] E. Schröger, M.H. Giard, C. Wolff, Auditory distraction: event-related potential and behavioral indices, Clin. Neurophysiol. 111 (2000) 1450–1460.

[21] A. Sek, B.C. Moore, Frequency discrimination as a function of frequency, measured in several ways, J. Acoust. Soc. Am. 97 (1995) 2479–2486.

[22] M. Tervaniemi, Musical sound processing in the human brain. Evidence from electric and magnetic recordings, Ann. NY Acad. Sci. 930 (2001) 259–272.

[23] M. Tervaniemi, T. Ilvonen, J. Sinkkonen, A. Kujala, K. Alho, M. Huotilainen, R. Näätänen, Harmonic partials facilitate pitch discrimination in humans: electrophysiological and behavioral evidence, Neurosci. Lett. 279 (2000) 29–32.

[24] M. Tervaniemi, E. Schröger, M. Saher, R. Näätänen, Effects of spectral complexity and sound duration on automatic complex-sound pitch processing in humans—a mismatch negativity study, Neurosci. Lett. 290 (2000) 66–70.

[25] H. Tiitinen, P. May, K. Reinikainen, R. Näätänen, Attentive novelty detection in humans is governed by pre-attentive sensory memory, Nature 372 (1994) 90–92.

[26] C.C. Wier, W. Jesteadt, D.M. Green, Frequency discrimination as a function of frequency and sensation level, J. Acoust. Soc. Am. 61 (1977) 178–184.

[27] E. Yago, M.J. Corral, C. Escera, Activation of brain mechanisms of attention switching as a function of auditory frequency change, NeuroReport 12 (2001) 4093–4097.