Neuroscience Letters 279 (2000) 29–32

Neuroscience Letters

www.elsevier.com/locate/neulet

Harmonic partials facilitate pitch discrimination in humans: electrophysiological and behavioral evidence

M. Tervaniemi[a,*], T. Ilvonen[a], J. Sinkkonen[a, b], A. Kujala[a], K. Alho[a], M. Huotilainen[a], R. Näätänen[a]

[a]Cognitive Brain Research Unit, Department of Psychology, University of Helsinki, Helsinki, Finland
[b]BioMag Laboratory, Medical Engineering Center, Helsinki University Central Hospital, Helsinki, Finland

Received 8 November 1999; accepted 23 November 1999

Abstract

The effect of the spectral tone structure on pre-attentive and attentive pitch discrimination was investigated. The mismatch negativity (MMN) component was recorded from reading subjects to pitch changes of identical magnitude in pure tones with only one sinusoidal frequency component and in spectrally rich tones with two additional harmonic partials. In a separate condition, subjects were asked to indicate detection of pitch change by a button press. The MMN was elicited with a larger amplitude and shorter latency by change in spectrally rich tones than by change in pure tones. Furthermore, the subjects' behavioral responses were more accurate for spectrally rich tones than for sinusoidal tones. Together these data indicate that pre-attentive and attentive pitch discrimination is facilitated with spectrally rich sounds in comparison to pure sinusoidal tones.

Keywords: Frequency discrimination; Complex sounds; Human; Auditory event-related potentials; Mismatch negativity; Automatic processing

Accurate processing of pitch is essential for speech and music perception: only after encoding the pitch of a sound, its relation to the preceding and following sounds can be determined. Preliminary evidence suggests that the pitch of spectrally rich sounds is easier to discriminate than that of sinusoidal tones consisting of only one frequency component [5,14,15]. These results, however, are not commonly cited to while discussing the perception of complex-tone pitch. This ignorance might result from missing evidence about the relative importance of, for instance, the magnitude of frequency change and number of spectral components in this facilitation. In addition, the existing behavioral evidence leaves it open whether increased amount of spectral components facilitates pitch discrimination despite the focus of subjects' conscious attention.

Brain mechanisms automatically activated by a sound change can be indexed by recording the mismatch negativity (MMN) component of the event-related potentials (ERPs) while the subject performs a task unrelated to the sounds. The MMN is elicited when a cortical memory trace representing the regularity of 'standard' stimuli is in discrepancy with the neural code of the incoming stimulus with 'deviant' parameter(s) [8]. MMN is elicited by any discriminable sound change in sinusoidal tones, such as stimulus frequency, intensity, duration, or in the locus of sound origin [4]. In addition, the MMN is generated by a change in spectrally complex stimuli such as phonemes [1,7,9] or chords [3,18]. The direction of subjects' attention does not affect the amplitude of MMN elicited by a frequency change in a sinusoidal tone [6,10] or by a phoneme change in consonant-vowel syllables [2].

The MMN latency and amplitude reflect the magnitude of the physical deviation between the deviant and standard stimulus, for instance, the frequency-MMN being earlier in latency and larger in amplitude with large frequency deviance [12]. Furthermore, the MMN amplitude and latency reflect perceptual accuracy, as determined by hit rates and reaction times in a behavioral experiment employing the same stimuli as the MMN recordings (e.g. [1,7,21]). This correlation between the MMN parameters and behavioral responses imply that preattentive neural functions determine the accuracy of the further, attentive, processes [11].

Spectrally rich sound structure has tentatively been

* Corresponding author. Cognitive Brain Research Unit, Department of Psychology, P.O. Box 13, University of Helsinki, 00014 Helsinki, Finland. Tel.: +358-9-1912-3408; fax: +358-9-1912-2924.
E-mail address: tervanie@helsinki.fi (M. Tervaniemi)

0304-3940/00/$ - see front matter
PII: S0304-3940(99)00941-6

shown to facilitate frequency discrimination even at the pre-attentive level by enhancing the MMN amplitude and shortening the MMN latency [17]. Further indirect evidence for facilitated pitch discrimination by harmonic partials was provided by a study in which an MMN was elicited by only 3% frequency change present in one out of 10 harmonic partials of a complex sound [20]. In behavioral studies this frequency change causes 1-Hz change of the perceived pitch in 155-Hz fundamental frequency, equaling a pitch change of 0.65% [4]. MMN studies employing pure sinusoidal tones have indicated that larger frequency differences are necessary for reliable MMN elicitation [12,21].

This study was conducted to systematically investigate whether spectrally rich tone structure facilitates pre-attentive and attentive pitch discrimination. To this end, ERPs elicited by stimuli with standard and deviant frequencies were recorded from healthy adult subjects while they were reading. The stimulation employed both pure sinusoidal and spectrally rich sounds in separate blocks. In addition, behavioral responses to the same stimuli were measured during a separate pitch-discrimination task.

Twelve healthy right-handed subjects with normal hearing were employed, none of them having any previous experience with EEG recordings. One subject was omitted from further analysis because of too noisy EEG. The remaining 11 subjects (three males) ranged in age from 18 to 29 years (mean 22 years).

The stimulation included pure tones and spectrally rich sounds (Table 1). The pure tones were sinusoidal tones of the fundamental frequency, which was 500 Hz in the standard stimulus. The spectrally rich stimuli consisted of three frequency components: in the standard stimulus, the frequencies were 500, 1000 and 1500 Hz. The second and third components were 3 and 6 dB lower in intensity, respectively, than the first component. Stimuli were presented binaurally via headphones at an intensity of 50 dB above individually determined hearing threshold. Tone duration was 75 ms, including 5 ms rise and fall times. The constant stimulus onset asynchrony (SOA) was 300 ms.

Each block consisted of either pure or spectrally rich tones including standard tones and six different kinds of deviant tones. The deviant tones were either higher or lower (±2.5, ±5, ±10%) in frequency than the standard tone, the frequency change occurring in spectrally rich tones in all frequency components (Table 1). The stimulus sequences were random except that each deviant tone was preceded by at least one standard tone.

Table 1
The frequencies of pure sinusoidal and spectrally rich tones[a]

	P	Deviance (%)	Pure tones (Hz)	Spectrally rich tones (Hz)
Standard	0.76		500	500 + 1000 + 1500
Deviant	0.04	+2.5	513	513 + 1026 + 1539
Deviant	0.04	−2.5	488	488 + 976 + 1464
Deviant	0.04	+5.0	525	525 + 1050 + 1575
Deviant	0.04	−5.0	476	476 + 952 + 1428
Deviant	0.04	+10.0	550	550 + 1100 + 1650
Deviant	0.04	−10.0	454	454 + 908 + 1362

[a] In spectrally rich tones, the intensities of the second and third harmonic partials were 3 and 6 dB softer than the first (fundamental) frequency. The probability of each stimulus type in the MMN experiment is denoted by *P*.

In the ignore condition, subjects were instructed to read a book of their own choice and to pay no attention to stimuli. They were presented with four blocks of 2000 stimuli. Two of the blocks consisted of pure tones and two of spectrally rich tones. The blocks were presented in randomized order. Each block included standard tones ($P = 0.76$) and six different kinds of deviant tones ($P = 0.04$ each).

In the discrimination condition (without ERP recording), subjects were instructed to listen to the stimuli and to press a response key as fast and accurately as possible whenever they heard a deviant tone. This condition consisted of six blocks of 1000 stimuli, with three of them consisting of pure and three consisting of spectrally rich tones. The blocks were presented in randomized order. Each block included standards ($P = 0.88$) and six types of deviants ($P = 0.02$ for each). The probabilities of deviant tones were lower in the discrimination condition than in the ignore condition to make it easier to determine to which deviant tone the button was pressed. The discrimination condition was always administered after the ignore condition to avoid any carry-over effects of attention to the sounds.

EEG recordings were performed in an acoustically and electrically shielded room. The EEG was recorded (pass band 0.1–40 Hz, sampling frequency 250 Hz) in the ignore condition by using a 32-channel electrode cap. In addition, horizontal eye movements were monitored with an electrode attached to the right outer canthus and vertical eye movements with the Fpz electrode. The EEG and EOG were referenced to the nose during the recordings and re-referenced to the right mastoid for the subsequent analysis.

The EEG was epoched (−50…300 ms including −50 ms pre-stimulus baseline) and averaged separately for each deviant stimulus and for standards immediately preceding deviants, for standards following deviants, and for the remaining standards. All epochs including voltage changes exceeding ± 100 μV were automatically rejected. Frequencies higher than 30 Hz were filtered out.

Difference waveforms (deviant-tone ERP minus standard-tone ERP) were calculated by using standards immediately preceding each deviant. To improve the signal-to-noise ratio, difference waves for identical magnitudes of deviance (+2.5 and −2.5%, +5 and −5%, +10 and −10%) were pooled together. The statistical significance of the MMN was tested by first determining the MMN peak latency from the Fz grand-average difference waves separately for each pure sinusoidal and spectrally rich deviant (±2.5, ±5, ±10%) as the most negative peak between 100–300 ms. The MMN amplitude was thereafter measured from individual difference waves as the average amplitude

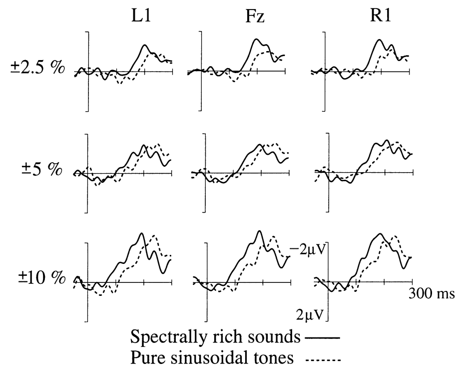

Fig. 1. Grand-average deviant-standard difference waves for sinusoidal tones (dashed line) and for spectrally rich tones (continuous line). The electrodes were located above the left (L1) and right (R1) fronto-temporal hemispheres and at the midline (Fz). The magnitude of frequency change is indicated on the left. The *x*-axis displays time (starting 50 ms before and ending 300 ms after stimulus onset) and the *y*-axis the amplitude of response (± 2 μV, negativity upwards).

calculated over the 20-ms time window centered at this grand-average peak latency.

For subsequent comparisons, the MMN latency and amplitude were quantified from the individual difference waves. For these analyses, the data were first high-pass filtered with 1-Hz cutoff. The amplitude and latency were then estimated from low-pass filtered data with 10-Hz cutoff at the time of the most negative peak between 100 and 300 ms at Fz electrode. This procedure was applied to decrease the data variability [19].

The significance of MMN was tested with two-tailed *t*-tests by comparing the mean MMN amplitude to zero. Two-way ANOVAs were used to determine whether the sound structure (levels: pure/spectrally rich) and the magnitude of deviance (levels: ± 2.5, ± 5, $\pm 10\%$) affected the MMN amplitude and latency. In addition, two-way ANOVAs were used to determine whether the sound structure (levels: pure/spectrally rich) and the magnitude of deviance (levels ± 2.5, ± 5, $\pm 10\%$) affected hit rates and d's. The statistical results are informed after applying Greenhouse–Geisser correction when appropriate (however, the original degrees of freedom are reported).

In the discrimination condition, a behavioral response was accepted as a hit if it occurred within 300–1200 ms interval after the onset of the deviant stimulus. All the other responses were regarded as false alarms. Thereafter, the hit rates and d' were calculated.

In the ignore condition, all deviant tones elicited a MMN with a fronto-central maximum and slightly larger amplitude over the right than over the left hemisphere (Fig. 1). The MMN peaked between 190–235 ms and was significant for all deviant tones ($t(10)$ between 4.7 and 12.3, $P < 0.001$ at Fz electrode).

The MMN latency was shorter when tones were spectrally rich than when they were pure sinusoidal tones ($F(1, 10) = 8.9$, $P < 0.014$) (Fig. 2; left). There was no significant effect of the magnitude of the deviance, neither was there significant interaction between magnitude of deviance and sound structure.

The MMN amplitude was larger when tones were spectrally rich when they were pure sinusoidal tones ($F(1, 10) = 11.7$, $P < 0.007$) (Fig. 2; middle). In addition, MMN amplitude was enhanced by increased frequency deviance ($F(2, 20) = 13.8$, $P < 0.0002$). There was no significant interaction between the magnitude of frequency deviance and the sound structure.

The behavioral performance of the subjects as indexed by the hit rates are in line the MMN data described above (Fig. 2; right): discrimination performance was more accurate with spectrally rich than with pure sinusoidal tones ($F(1, 10) = 6.2$, $P < 0.03$). In addition, the greater was the frequency deviance, the higher was the hit rate ($F(2, 20) = 20.6$, $P < 0.0001$). There was no significant interaction between the magnitude of deviance and the sound structure. Due to a relatively small number of false alarms, the d' results corroborate with those displayed by the hit rates. The d' was higher with spectrally rich tones than with sinusoidal tones ($F(1, 10) = 11.4$, $P < 0.007$). In addition, the greater was the magnitude of the frequency deviance, the higher was the d' ($F(2, 20) = 31.8$, $P < 0.0002$). There was no significant interaction between the magnitude of deviance and the sound structure.

The present MMN and behavioral data thus indicate that pitch discrimination is facilitated in the human auditory

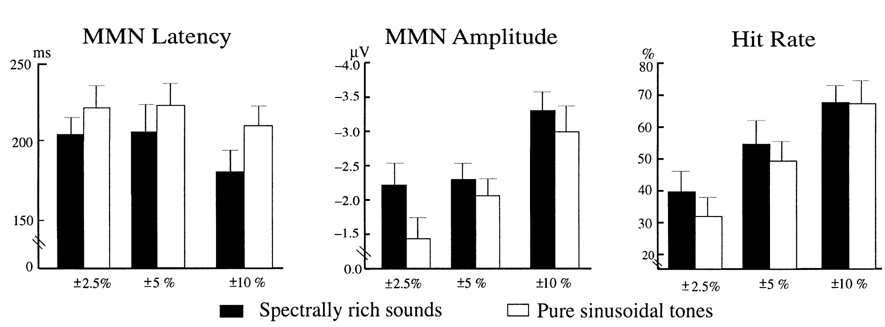

Fig. 2. The mean MMN latency (left), the mean MMN amplitude (middle), and the mean hit rate (right) (the half range of the standard error of mean being presented on the top of each bar).

system when the tones are spectrally rich compared with pure sinusoidal tones. The MMN latency was shorter when tones were spectrally rich than when they were pure sinusoidal tones. In contrast, there was no significant effect of the magnitude of the frequency deviance nor interaction between the magnitude of the frequency deviance and the sound structure on MMN latency (Fig. 2; left). This suggests that with these relatively small frequency changes, the MMN latency was modulated by the sound structure rather than by the magnitude of sound change. In contrast, the MMN amplitude was enhanced by both the rich sound structure and the increased frequency deviance (however, without an interaction between these factors) (Fig. 2; middle). This pattern of data was reflected in the behavioral performance as well: the hit rate was more accurate with spectrally rich than with pure sinusoidal tones and with larger than smaller deviances (Fig. 2; right). Thus, the MMN latency reflected the sound structure and was saturated with regard to the magnitude of the sound change, whereas the MMN amplitude (and the subject's perceptual discrimination) was sensitive to both manipulations. Interestingly, the lacking interaction between the sound structure and the magnitude of sound change in any of the present MMN or behavioral measures suggests that the facilitation caused by the rich sound structure is not influenced by the magnitude of frequency change.

The present facilitation in pitch discrimination with spectrally rich sounds might be explained by information increase: spectrally rich tones carry more information than sinusoidal tones, in both spectral [16] and temporal terms [13]. In addition, this facilitation could be accounted for by the difference in the stimulus familiarity. Since our auditory system has extensive experience with spectrally rich stimulation like speech and music, it appears plausible that it reacts more vigorously to changes in spectrally complex than simple (ecologically unrepresentative) information.

In conclusion, the present ERP and behavioral data indicate that pitch discrimination is facilitated with spectrally rich sounds in comparison with pure sinusoidal tones. This facilitation was found at the behavioral level when subjects were engaged in a discrimination task related to the sounds and at the physiological level when they were reading a book during the auditory stimulation.

The study was supported by EU contract BMH4-CT96–0819 (COBRAIN) and by the Academy of Finland. The authors thank Professor E. Schröger and Dr. P. Cariani for their comments on the earlier version of the manuscript.

[1] Aaltonen, O., Eerola, O., Lang, A.H., Uusipaikka, E. and Tuomainen, J., Automatic discrimination of phonetically relevant and irrelevant vowel parameters as reflected by mismatch negativity. J. Acoust. Soc. Am., 96 (1994) 1489–1493.

[2] Alho, K., Connolly, J.F., Cheour, M., Lehtokoski, A., Huotilainen, M., Virtanen, J., Aulanko, R. and Ilmoniemi, R.J., Hemispheric lateralization in preattentive processing of speech sounds. Neurosci. Lett., 258 (1998) 9–12.

[3] Alho, K., Tervaniemi, M., Huotilainen, M., Lavikainen, J., Tiitinen, H., Ilmoniemi, R.J., Knuutila, J. and Näätänen, R., Processing of complex sounds in the human auditory cortex as revealed by magnetic brain responses. Psychophysiology, 33 (1996) 369–375.

[4] Darwin, C. 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.

[5] Flanagan, J.L. and Saslow, M.G., Pitch discrimination for synthetic vowels. J. Acoust. Soc. Am., 30 (1958) 435–442.

[6] Kathmann, N., Frodl-Bauch, T. and Hegerl, U., Stability of the mismatch negativity under different stimulus and attention conditions. Clin. Neurophysiol., 110 (1999) 317–323.

[7] Kraus, N., McGee, T.J., Carrell, T.D., Zecker, S.G., Nicol, T.G. and Koch, D.B., Auditory neurophysiologic responses and discrimination deficits in children with learning problems. Science, 273 (1996) 971–973.

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

[9] Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Ilvonen, A., Vainio, M., Alku, P., Ilmoniemi, R.J., Luuk, A., Allik, J., Sinkkonen, J. and Alho, K., Language specific phoneme representations revealed by the electric and magnetic brain responses. Nature, 385 (1997) 432–434.

[10] Näätänen, R., Paavilainen, P., Tiitinen, H., Jiang, D. and Alho, K., Attention and mismatch negativity. Psychophysiology, 30 (1993) 436–450.

[11] Novak, G.P., Ritter, W., Vaughan Jr., H.G. and Witznitzer, M.L., Differentiation of negative event-related potentials in an auditory discrimination task. Electroenceph. clin. Neurophysiol., 75 (1990) 255–275.

[12] Sams, M., Paavilainen, P., Alho, K. and Näätänen, R., Event-related potentials to infrequent changes in synthesized phonetic stimuli. Electroenceph. clin. Neurophysiol., 62 (1985) 437–448.

[13] Schouten, J.F., The residue revisited. In R. Plomp and G.F. Smoorenburg (Eds.), Frequency Analysis and Periodicity Detection in Hearing, Sijthoff, Leiden, 1970.

[14] Sidtis, J., On the nature of the cortical function underlying right hemisphere auditory perception. Neuropsychology, 18 (1980) 321–330.

[15] Sundberg, J., The Science of Musical Sounds, Academic Press, San Diego, CA, 1991.

[16] Terhardt, R.J., Pitch, consonance, and harmony. J. Acoust. Soc. Am., 55 (1974) 1061–1969.

[17] Tervaniemi, M., Alho, K., Paavilainen, P., Sams, M. and Näätänen, R., Absolute pitch and event-related brain potentials. Music Percept., 10 (1993) 305–316.

[18] Tervaniemi, M., Kujala, A., Alho, K., Virtanen, J., Ilmoniemi, R.J. and Näätänen, R., Functional specialization of the human auditory cortex in processing phonetic and musical sounds. NeuroImage, 9 (1999) 330–336.

[19] Tervaniemi, M., Lehtokoski, A., Sinkkonen, J., Virtanen, J., Ilmoniemi, R.J. and Näätänen, R., Test-retest reliability of mismatch negativity (MMN) for duration, frequency, and intensity changes. Clin. Neurophysiol., 10 (1999) 1388–1393.

[20] Tervaniemi, M., Schröger, E. and Näätänen, R., Pre-attentive processing of spectrally complex sounds with asynchronous onsets: an event-related potentials study in human subjects. Neurosci. Lett., 227 (1997) 197–200.

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