Does familiarity facilitate the cortical processing of music sounds?

Georg Neuloh[1,2,CA] and Gabriel Curio[1]

[1]Neurophysics Group, Department of Neurology and Clinical Neurophysiology, Campus Benjamin Franklin, Charité-University Medicine Berlin, Hindenburgdamm 30, 12200 Berlin; [2]Department of Neurosurgery, Friedrich-Wilhelm-Universität, Sigmund-Freud-Strasse 25, 53105 Bonn, Germany

[CA]Corresponding Author: neuloh@ukb.uni-bonn.de

Received 5 September 2004; accepted 23 September 2004

Automatic cortical sound discrimination, as indexed by the mismatch negativity (MMN) component of the auditory evoked potential, is facilitated for familiar speech sounds (phonemes). In musicians as compared to non-musicians, an enhanced MMN has been observed for complex non-speech sounds. Here, musically trained subjects were presented with sequences of either familiar (tonal) or structurally matched unfamiliar (atonal) triad chords, both with either fixed or randomly transposed chord root pitch. The MMN elicited by deviant chords did not differ for familiar and unfamiliar triad sequences, and was undiminished even to unfamiliar deviant sounds which were consciously undetectable. Only subsequent attention-related components indicated facilitated cognitive processing of familiar sounds, corresponding to higher behavioral detection scores. *NeuroReport* 15:2471–2475 © 2004 Lippincott Williams & Wilkins.

Key words: Mismatch negativity; Music; Sensory discrimination; Tonal triads

INTRODUCTION

Cortical processing of acoustic stimuli can be probed non-invasively by auditory evoked potentials (AEP). The mismatch negativity (MMN) is an AEP component which indicates that in a stimulus series a new auditory event deviates from a sensory memory trace representing invariant features of preceding acoustic standard stimuli [1]. This memory trace encodes not only simple stimulus properties such as pitch, loudness, or duration, but also higher-order attributes of complex stimuli, such as spectral composition, frequency ratios, dynamic or pitch contour of sequential tone patterns, and combinations of such features [2–8]. The MMN process is automatic as it operates without attention directed toward the auditory stimulus stream [9]. In turn, it can trigger an involuntary shift of attention towards the deviant event; this shift is reflected by subsequent evoked potential components, such as the N2b-P3a complex and the P3 component [10]. Typically, these components indicate conscious novelty processing.

The mismatch paradigm permits to investigate the discriminative processing of domain-specific complex stimuli, such as speech (phonemes) and music sounds [5,11–14]. MMN results obtained for speech sounds revealed that sensory processing is facilitated specifically for phonemes from the listener's native language [3]. In the music domain, the MMN to non-speech sound mismatch has been found enhanced when comparing musicians and non-musicians [7,8,15,16]. It is an interesting further question to what extent facilitation by long-term familiarity is a general feature of auditory mismatch processing, e.g., whether it holds also for musical sounds, in particular for highly familiar triad chords of the Western tonal system which can be considered as units akin to phonemes of a Western listener's native musical language. Furthermore, it is not clear whether a discrimination performance better for familiar as compared to unfamiliar complex stimuli would be preceded in general by a higher mismatch brain response.

Here, we tested the hypothesis that long-term familiarity of complex non-speech sounds facilitates their sensory mismatch processing. Evoked potentials were recorded from musically trained subjects who were presented with sound mismatches in sequences of either highly familiar or unfamiliar but structurally matched music chords. A behavioral detection task allowed to correlate discrimination performance and mismatch responses.

MATERIALS AND METHODS

Twenty-two subjects (nine females, 13 males; median age 28 years; 21 right-handed) were paid to participate in this study. All subjects had extensive musical training on up to four musical instruments, beginning at a median age of 7 years, for a median period of 16.5 years, and listened to Western tonal music for a median of 15 h/week. Stimuli were synthesized with the Turbosynth software (Digidesign, Daly City, CA, USA), generated at 20 kHz sampling rate, and replayed binaurally from a DAT-tape via headphones at ~70 dB SPL. Each stimulus consisted of three simultaneous tones (triads), each composed of 16 harmonics with exponentially decaying amplitudes, to ease sensory and cognitive discrimination [4]. The resulting chords had a rich, organ-like timbre. Stimulus duration was 1000 ms including 25 ms rise and fall times. A silent period of 1000 ms separated every two stimuli to permit aware assessment of the sound category. Oddball chord sequences were designed as detailed below, with a probability of $p \sim 0.2$ for deviant chords and 3–5 standard chords intervening between any two deviants. Each sequence consisted of 433 stimuli and was presented in three parts of 5 min each to counteract fatigue. In experiment 2, subjects performed a two-alternative forced-choice detection task, and discriminative performance was calculated by Grier's non-parametric unbiased sensitivity index [17].

EEG activity was recorded using a 0.1–70 Hz analog bandpass and digitized at 250 Hz sampling rate from Ag/AgCl electrodes at the Fz, Cz, Pz, T3, T4 positions of the international 10-20 system, from both mastoids (Ml, Mr) and from frontolateral positions at 1/3 of the distance between Fz and the mastoids (Fl and Fr). The reference electrode was placed at the tip of the nose. Impedances were kept below 5 kΩ. The baseline was defined by the mean amplitude within 50 ms before stimulus onset. Electro-oculogram readings from the upper nasal *vs* the lower temporal right orbital rim provided the artifact rejection criterion ($\pm 70\,\mu V$ within -500 to $+1500$ ms relative to stimulus onset). EEG responses were digitally filtered off-line (30 Hz low-pass, FIR). Sub-averages were calculated for standards and deviants, in experiment 2 grouped also according to discrimination performance criteria (Fig. 2), and compared in a multifactorial design as detailed in the results section. Amplitudes were calculated as mean values within 50 ms time windows centered around subcomponent peak latencies in the grand average waveforms. To enhance the signal-to-noise ratio for the analysis of the MMN subcomponent, which reverses its polarity at recording sites inferior to the supratemporal plane [1], re-referenced amplitudes were calculated by subtracting the averaged values at the mastoids from those at the non-mastoid leads (e.g., Fz-(M1 + M2)/2). Analyses for later subcomponents (N2b, P3), which do not show corresponding phase reversal characteristics, were performed without re-referencing.

Experiment 1: Eight subjects participated in experiment 1, which consisted of two blocks. In the first (tonal) block the standard stimuli were C-major triad chords (C-E-G, C=261 Hz, 4 + 3 semitone intervals, i.e., minor over major third) and the deviants were C-minor triad chords (C-E^b-G, 3 + 4 semitones, i.e., major over minor third). In the second (atonal) block, C-major and C-minor triads were replaced with structurally matched chords which, instead of major and minor thirds, were built of tritone and perfect fourth intervals, i.e., B^b-E-A (6 + 5 semitones, i.e., perfect fourth over tritone) replacing standard C-major chords, and B^b-E^b-A (5 + 6 semitones, i.e., tritone over perfect fourth) replacing deviant C-minor triads. Thus, the upper and lower tones of tonal as well as atonal triads remained unchanged between standards and deviants within a block, whereas a change from E to E^b in the middle triad tone marked the occurrence of a deviant sound in both blocks. In contrast to the ubiquitous major/minor triads, the atonal chords do not have a particular harmonic function in traditional Western tonal music and do rarely occur in this context. Although they may be employed more frequently in the extended tonality of 20th century classic and modern jazz music, they remain highly unusual to the traditionally trained musician. Tonal and atonal blocks were presented in randomized order. Subjects were instructed to listen attentively; no specific task was to be performed.

Experiment 2: Twenty subjects participated in experiment 2 where the chord root notes were varied in pseudo-random order between $C^{\#}$ and D′ on a well-tempered chromatic scale (C=261 Hz). Accordingly, tonal chord sequences consisted of major triads as standards and minor triads as deviants, and atonal chord sequences consisted of tritone/fourth triads as standards and fourth/tritone triads as deviants. In addition, the triad sequences were arranged such that no frequency memory traces from tone repetitions between consecutive triad chords could emerge. Thus, there was no simple frequency mismatch as in experiment 1, but a complex frequency ratio difference between deviants and standards. Subjects were instructed to discriminate standards and deviants in a forced-choice finger tapping task within the 1 s silent interval following the stimulus.

RESULTS

Experiment 1: Figure 1a,b depicts standard, deviant, and difference (deviant minus standard) waveforms at Fz for the tonal block and the atonal block. A typical MMN was elicited in both blocks, i.e., irrespective of tonality. A significant MMN phase reversal was found between Fz and mastoids in the 150–200 ms poststimulus time window, with negative amplitudes at Fz (p=0.002 in block 1, p=0.003 in block 2) and positive amplitudes at the mastoids ($p<0.001$ at M1 and p=0.001 at M2 in block 1, p=0.002 at M1 and p=0.021 at M2 in block 2). A three-way ANOVA with the factors familiarity (familiar tonal chords in block 1 *vs* unfamiliar atonal chords in block 2), deviance (deviants *vs* standards), and distribution (recording electrode positions) showed a significant deviance main effect (F(1,7)= 109.343, $p<0.001$) and a trend for a distribution × deviance interaction (F(7,1)=139.8, p=0.065), but did not reveal any familiarity main effect or interaction. A separate ANOVA including only frontolateral leads to test for a possible laterality effect did not show any significant right/left difference. Thus, a typically distributed MMN was found which did not differ between familiar and unfamiliar sound sequences, and did not show any significant left-right difference.

The difference curves in Fig. 1a,b display an N2b-like negativity (peaking within the 250–300 ms window after the MMN) only in the tonal block 1, with no corresponding negativity in the atonal block 2. A two-way left-right × deviance ANOVA (including only the frontolateral recording sites with the expectedly highest N2b amplitude), revealed a nearly significant deviance effect in this time window for block 1 (F(1,7)=5.431, p=0.053), but no right/left main effect or interaction.

Following the N2b in block 1, and the MMN in block 2, a positive deflection peaking within 350–400 ms poststimulus is visible in the difference curves of both blocks, corresponding to a P3a component. A three-way familiarity × deviance × distribution ANOVA revealed only a trend for a deviance effect (F(1,7)=5.223, p=0.056); however, a significant familiarity × deviance interaction was found (F(1,7)= 8.598, p=0.022). Therefore, two-way deviance × distribution

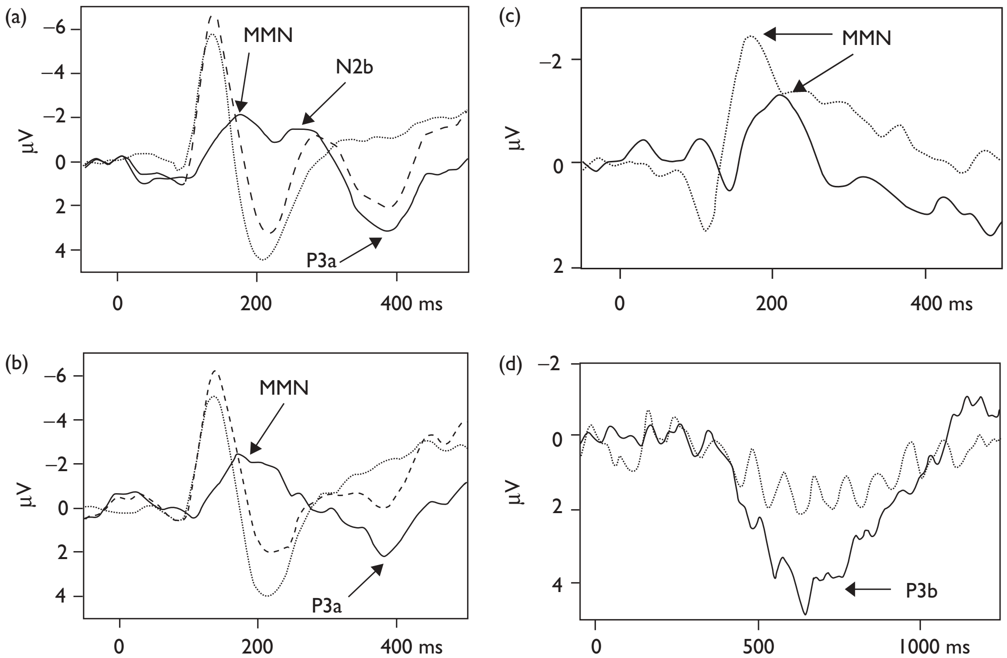

Fig. 1. Grand-average standard (.......), deviant (– – – – –), and difference (deviant-standard —) waveforms of the tonal (**a**) and atonal (**b**) conditions of experiment 1. While the MMN is of comparable amplitude and latency in both blocks, there is a distinct N2b only in the tonal condition, followed by a larger P3a in this block. Difference waveforms of the tonal (—) and atonal (.......) block of experiment 2 at Fz (**c**) and Pz (**d**) show a seemingly larger, and earlier MMN component in the atonal condition, but a significant P3b component only in the tonal condition.

ANOVAs were calculated for each block separately, which showed a significant deviance main effect restricted to the tonal block 1 (F(1,7)=6.619, p=0.037), whereas the corresponding positivity did not reach significance in block 2. Taken together, deviant triads elicited a distinct N2b/P3a complex only in the familiar chord sequence.

Experiment 2: Figure 1c depicts difference (deviant minus standard) waveforms at Fz for both the tonal and atonal block. A typical MMN with reversed polarity between Fz and the mastoids in the corresponding 175–225 ms time window was elicited in both blocks, with a slightly earlier peak latency in block 1. A three-way familiarity × deviance × distribution ANOVA (calculated for re-referenced values in the 175–225 ms window) showed a significant deviance main effect (F(1,19)=6.457, p=0.02) and a deviance × distribution interaction (F(7,13)=8.358, p=0.001), but no effect involving familiarity, despite the seemingly larger MMN amplitude in block 2. Corresponding to experiment 1, a separate ANOVA including only frontolateral leads showed no significant laterality effect. Thus, a typically distributed MMN was elicited in both blocks which did not differ significantly for familiar and unfamiliar chords.

In contrast to the MMN, a typical large, parietally maximal P3 positivity occurred only in block 1, where the majority of subjects scored high in the detection paradigm (Fig. 1d), whereas only a slight positive deflection was elicited in block 2. Correspondingly, a three-way ANOVA for the time-window 500–1000 ms showed a significant familiarity × deviance interaction (F(1,19)=4.498, p=0.047) along with a deviance × distribution interaction (F(7,13)= 6.74, p=0.002).

A scatter diagram of Grier's non-parametric unbiased sensitivity index [17] for the forced-choice detection paradigm in both blocks (Fig. 2a) shows that majorities of subjects scored high for tonal (continuous ellipse) and low (dotted ellipse) for atonal sounds. Figure 2b displays the re-referenced difference waveform at Fz from block 2 (unfamiliar chords) for all subjects, and superimposed subaverages for the good and poor performers in the discrimination task of this block, revealing a similar MMN in all groups. The corresponding differences waveforms of block 2 at Pz are superimposed in Fig. 2c, expectedly displaying a P3 component only for the small group of good performers. A separate two-way deviance × distribution ANOVA was performed with only the correctly detected standards and the deviants falsely classified as standards (missed deviants), which still showed a deviance main effect (F(1,19)=6.249, p=0.022), as well as a deviance × distribution interaction (F(7,13)=2.87, p=0.048). Even with the analysis restricted to the misses in the majority subgroup of 15 subjects with a close-to-chance performance in the detection task in order to eliminate negative response bias (truly undetectable misses, Fig. 2d), there was still both, a significant deviance main effect (F(1,14)=12.919, p=0.003) and a deviance × distribution interaction (F(7,8)=3.854, p=0.039). Thus, a typical MMN was elicited by missed, and even undetectable, deviants in the detection task.

DISCUSSION

The present study found a typical sensory mismatch response in sequences of tonal and atonal music chords: In experiment 1 the MMN reflects both an absolute frequency mismatch and a frequency ratio mismatch, whereas in

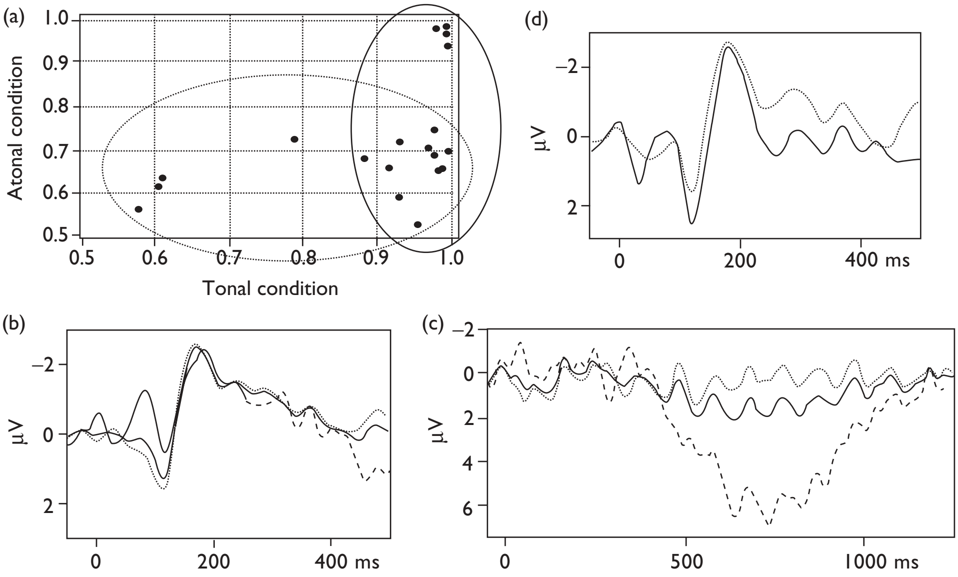

Fig. 2. Grier's A' non-parametric detection scores [17] for both tonal and atonal conditions are plotted in (**a**), majorities of subjects scoring high for tonal (continuous ellipse), and low (dotted ellipse) for atonal sounds. Re-referenced grand-average difference curves at Fz from the atonal block (**b**) with a similar MMN for all subjects (—), good (– – – –), and poor (…….) performers in the detection task. Difference curves at Pz for the same groups (**c**), with a significant P3b deflection only for the group with a high detection score. Re-referenced difference curves at Fz only for missed deviants and correctly detected standards in the atonal block (**d**), with an undiminished MMN for all subjects (—) and particularly for the poor performers (…….) in the detection task.

experiment 2 only a frequency ratio mismatch was available to generate the MMN response, notably with an amplitude lower than in experiment 1. In contrast to the hypothesis put forward in the introduction, the MMN amplitude was not enhanced for triads drawn from familiar musical categories. Only the subsequent cognitive components indicated a facilitated novelty response for familiar sounds as opposed to structurally matched, but unfamiliar sounds; this was reflected also by the behavioral discrimination performance. Conspicuously, an undiminished sensory mismatch response was generated also by missed deviant atonal chords, even in subjects who were in general unable to consciously discriminate these highly unfamiliar triads.

In previous studies it was consistently shown that the MMN depends on musical proficiency [7,8,15,16], and intact musical competence [18]. The present results are not in conflict with these findings, which seem to reflect a general improvement of auditory sensory discrimination in musicians and musically competent subjects, and this effect is not restricted to the domain of tonal music [7,8,15,16,19]. Here, the comparison was not between musicians and non-musicians, but between specific familiar tonal and unfamiliar atonal sounds in musically trained subjects (although subjects could be grouped according to their discriminative performance as depicted in Fig. 2).

The fact that no MMN amplitude facilitation was found for highly familiar music sounds is, however, in contrast to previous results in the language domain. Phoneme discrimination and categorization as reflected by the MMN depend on long-term familiarity with speech sounds, i.e., they are facilitated through language-specific phonetic representations which develop during the first few months of life, or during acquisition of a foreign language (for review see [11]). The exposure to speech very early in life may contribute to this language-specificity of sensory sound processing [3,20], as compared to the discrimination of music sounds in the present group of subjects who had their musical training at a later age.

The lack of correlation between discrimination efficiency and novelty-detection related AEP components on the one hand, and the MMN on the other hand in experiment 2 (Fig. 2), where an identical MMN was elicited in all performance groups, and even by undetectable deviant chords, is in contrast to previous studies where a positive correlation of MMN amplitude and detection performance has been found [6,21,22]. However, the improved detection performance in these studies was typically related to an improved sensory discrimination of small physical differences, e.g., with hearing recovery and increased MMN in cochlear implant users [22], rather than due to an improved analytical discrimination of more familiar complex sounds, as it occurred with tonal chords in the present study.

From the perspective of music processing, it is a remarkable finding that an undiminished MMN was elicited for unfamiliar triads even by deviances which were undetectable. Such consciously unnoticed neuronal information about changes in chord category (e.g., in harmonic modulation) might be effective in subconscious musical influences on mood modulation, both in tonal and non-tonal music. The mechanism for gating to conscious perception can be studied when operating at the output pathways of the auditory cortex mismatch detector, e.g., with the induction of cognitive components related to rule-based processing of Western music [23].

The apparent lack of a familiarity effect on sensory chord processing in the present study might be due to a weak, rather than an absent amplitude facilitation for tonal triads, hidden by a low signal-to-noise ratio. However, this would still be in contrast with findings in sensory speech processing where the facilitating effect of phoneme familiarity typically leads to a clearly enhanced MMN amplitude [11]. Likewise, an amplitude ceiling effect is unlikely, because the MMN peak amplitudes of around 2 μV in the present data were not exceedingly high. It is also possible that some roughness or dissonance in the atonal chords might have enhanced the MMN amplitude in the unfamiliar chord sequences [24,25], which would provide an explanation for the (insignificantly) higher amplitude of the MMN in the atonal block of experiment 2. However, a dominant roughness effect, which might have even paradoxically masked some amplitude facilitation by sound familiarity, is unlikely since it should lead to an enhanced primary cortical response (N1), which was not observed (Fig. 1); likewise, it should be paralleled by facilitation of the subsequent cognitive components and improved discrimination scores, which is contrary to what was found in the present study.

CONCLUSION

The results of this study do not show a facilitation of preattentive sensory discrimination of familiar music chords as reflected by the MMN; only later AEP components related to novelty-detection were facilitated, corresponding to an improved behavioral discrimination performance. This contrasts with the previously described facilitation of the sensory discrimination response for familiar speech sounds, and emphasizes a leading role for speech with regard to long-term plasticity in preattentive sensory processing. The undiminished MMN even for undetectable atonal chord deviances indicates that sensory discrimination is fully effective below the level of conscious analytical categorization of complex music sounds, which was shown to reflect long-term familiarity in the present study.

REFERENCES

1. Naatanen R, Gaillard AW and Mantysalo S. Early selective-attention effect on evoked potential reinterpreted. *Acta Psychol* 1978; **42**:313–329.
2. Alho K and Sinervo N. Preattentive processing of complex sounds in the human brain. *Neurosci Lett* 1997; **233**:33–36.
3. Naatanen R, Lehtokoski A, Lennes M, Cheour M, Huotilainen M, Iivonen A *et al*. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 1997; **385**:432–434.
4. Tervaniemi M, Schroger E, Saher M and Naatanen R. Effects of spectral complexity and sound duration on automatic complex-sound pitch processing in humans – a mismatch negativity study. *Neurosci Lett* 2000; **290**:66–70.
5. Tervaniemi M and Huotilainen M. The promises of change-related brain potentials in cognitive neuroscience of music. *Ann NY Acad Sci* 2003; **999**:29–39.
6. Paavilainen P, Simola J, Jaramillo M, Naatanen R and Winkler I. Preattentive extraction of abstract feature conjunctions from auditory stimulation as reflected by the mismatch negativity (MMN). *Psychophysiology* 2001; **38**:359–365.
7. Fujioka T, Trainor LJ, Ross B, Kakigi R and Pantev C. Musical training enhances automatic encoding of melodic contour and interval structure. *J Cogn Neurosci* 2004; **16**:1010–1021.
8. van Zuijen TL, Sussman E, Winkler I, Naatanen R and Tervaniemi M. Grouping of sequential sounds – an event-related potential study comparing musicians and nonmusicians. *J Cogn Neurosci* 2004; **16**:331–338.
9. Naatanen R. *Attention and Brain Function*. Hillsdale: Lawrence Erlbaum; 1992.
10. Friedman D, Cycowicz YM and Gaeta H. The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neurosci Biobehav Rev* 2001; **25**:355–373.
11. Naatanen R. The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology* 2001; **38**:1–21.
12. Pantev C, Ross B, Fujioka T, Trainor LJ, Schulte M and Schulze M. Music and learning-induced cortical plasticity. *Ann NY Acad Sci* 2003; **999**: 438–450.
13. Besson M and Schon D. Comparison between language and music. *Ann NY Acad Sci* 2001; **930**:232–258.
14. Munte TF, Nager W, Beiss T, Schroeder C and Altenmuller E. Specialization of the specialized: electrophysiological investigations in professional musicians. *Ann NY Acad Sci* 2003; **999**:131–139.
15. Nager W, Kohlmetz C, Altenmuller E, Rodriguez-Fornells A and Munte TF. The fate of sounds in conductors' brains: an ERP study. *Brain Res Cogn Brain Res* 2003; **17**:83–93.
16. Hashimoto T, Hirata Y and Kuriki S. Auditory cortex responds in 100 ms to incongruity of melody. *Neuroreport* 2000; **11**:2799–2801.
17. Grier JB. Nonparametric indexes for sensitivity and bias: computing formulas. *Psychol Bull* 1971; **75**:424–429.
18. Kohlmetz C, Altenmuller E, Schuppert M, Wieringa BM and Munte TF. Deficit in automatic sound-change detection may underlie some music perception deficits after acute hemispheric stroke. *Neuropsychologia* 2001; **39**:1121–1124.
19. Tervaniemi M, Ilvonen T, Karma K, Alho K and Naatanen R. The musical brain: brain waves reveal the neurophysiological basis of musicality in human subjects. *Neurosci Lett* 1997; **226**:1–4.
20. Cheour M, Martynova O, Naatanen R, Erkkola R, Sillanpaa M, Kero P *et al*. Speech sounds learned by sleeping newborns. *Nature* 2002; **415**: 599–600.
21. Novitski N, Tervaniemi M, Huotilainen M and Naatanen R. Frequency discrimination at different frequency levels as indexed by electrophysiological and behavioral measures. *Brain Res Cogn Brain Res* 2004; **20**:26–36.
22. Lonka E, Kujala T, Lehtokoski A, Johansson R, Rimmanen S, Alho K and Naatanen R. Mismatch negativity brain response as an index of speech perception recovery in cochlear-implant recipients. *Audiol Neurootol* 2004; **9**:160–162.
23. Koelsch S, Schroger E and Tervaniemi M. Superior pre-attentive auditory processing in musicians. *Neuroreport* 1999; **10**:1309–1313.
24. De Baene W, Vandierendonck A, Leman M, Widmann A and Tervaniemi M. Exploration of roughness by means of the mismatch negativity paradigm. *Ann NY Acad Sci* 2003; **999**:170–172.
25. Itoh K, Suwazono S and Nakada T. Cortical processing of musical consonance: an evoked potential study. *Neuroreport* 2003; **14**:2303–2306.

Acknowledgements: The authors thank Professor Dr. A. Friederici, head of the Max Planck Institute of Cognitive Neuroscience in Leipzig/Germany, and her group, for support during data acquisition and evaluation, and for numerous helpful discussions.