Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Cognitive Brain Research 23 (2005) 270–276

COGNITIVE BRAIN RESEARCH

www.elsevier.com/locate/cogbrainres

Research report

Auditory organization of sound sequences by a temporal or numerical regularity—a mismatch negativity study comparing musicians and non-musicians

Titia L. van Zuijen[a,d,*], Elyse Sussman[b], István Winkler[c], Risto Näätänen[a,d], Mari Tervaniemi[a,d]

[a]*Cognitive Brain Research Unit, Department of Psychology, University of Helsinki, P.O. Box 9 (Siltavuorenpenger 20 C), FIN-00014 Helsinki, Finland*
[b]*Department of Neuroscience and Department of Otolaryngology, Albert Einstein College of Medicine, New York, USA*
[c]*Institute for Psychology, Hungarian Academy of Sciences, Budapest, Hungary*
[d]*Helsinki Brain Research Center, University of Helsinki, Finland*

Accepted 26 October 2004
Available online 23 November 2004

Abstract

The human auditory system can encode regularities in the acoustic environment without the requirement of attention. We investigated whether the auditory system of musicians is more sensitive than that of non-musicians in encoding complex regularities. We presented tone sequences containing either a temporal or a numerical regularity. The sequence with the temporal regularity could be divided into segments of a constant duration while the segments contained a varying number of tones. The sequence with the numerical regularity, on the other hand, could be divided into segments containing a constant number of tones while the segments varied in duration. Auditory encoding of the regularity was determined by measuring whether an occasional segment lengthening, either in time or by number elicited the mismatch negativity (MMN). In both musicians and non-musicians, an MMN was elicited when the temporal regularity was violated. In contrast, only in musicians an MMN was elicited to violations of the numerical regularity. The results show that temporal processing is of general importance in audition since at an involuntary auditory processing stage a complex temporal regularity can be encoded irrespective of musical expertise. Furthermore, the auditory system of professional musicians can encode a numerical regularity without attention being required reflecting the functional importance of beat tracking in the perceptual organization of music.

Theme: Neural basis of behavior
Topic: Learning and memory: systems and functions

Keywords: Mismatch negativity; Regularity detection; Musician; Temporal processing; Processing of numerosity

---

1. Introduction

The auditory system needs to organize the sound input in order to extract meaning from the acoustic environment. An important aspect of auditory organization is the integration of sound over time—a process that benefits from regularity and repetition. A regular context compared to an irregular context improves the detection of small (just notable) differences in interval duration [6]. It is also easier to detect silent gaps in a stimulus sequence with a regularly varying temporal organization than to detect gabs in a stimulus sequence with an irregular varying temporal organization [11]. Musicians have various pronounced skills in auditory perception that correlate with their expertise in music. For instance, musicians can detect smaller pitch differences than non-musicians, are more accurate in judging pitch intervals [5,26] and are better in structuring rhythms [7]. In this study, we investigate whether the auditory system of musicians is more sensitive

---

* Corresponding author. Cognitive Brain Research Unit, Department of Psychology, University of Helsinki, P.O. Box 9 (Siltavuorenpenger 20 C), FIN-00014 Helsinki, Finland. Fax: +358 9 191 29450.
E-mail address: titia@cbru.helsinki.fi (T.L. van Zuijen).

0926-6410/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.cogbrainres.2004.10.007

than that of non-musicians in encoding a musically relevant regularity which, in order to be detected, requires sequential integration.

The mismatch negativity (MMN) of the auditory event-related potentials (ERPs) is elicited when the auditory system detects occasional changes in an otherwise repetitive or regular sequence [13,14,18,19]. The MMN is generated involuntarily also to task-irrelevant sounds, originates mainly from the auditory cortex [1,2] and peaks at about 120–220 ms after deviant onset. By investigating which regularity violations can elicit an MMN, we can assess which regularities the auditory system can extract from the sound material and encode in a memory template serving the acoustic-change detection. It has been shown that in addition to repetition of individual sounds, also repetitive relationships between sounds can be encoded in the memory system underlying MMN elicitation. For example, MMN has been elicited to violations of tone alternation [15], occasional changes in sequential tone patterns [23] and to changes in tonal relations of a short melody [25,27]. The MMN can therefore be used to assess how the auditory system organizes information over time [3,22–24,29].

The superior auditory processing skills of musicians [4,12,17] are reflected in the MMN elicitation. Chords that deviate from repetitive standard chords by only 0.75% in pitch elicit an MMN in professional violinists, whereas for MMN to be elicited in musical novices, a much larger pitch deviation is required [8]. Musicians have also a longer time window for integrating sounds as indicated by MMN elicitation [20]. In addition, in van Zuijen et al. [29], we showed that the auditory system of musicians organizes sounds differently from that of non-musicians. Professionally trained classical musicians showed auditory grouping of a tone sequence into four-tone groups according to a more complex Gestalt-grouping rule than non-musicians. We found evidence of auditory grouping according to pitch-similarity in both musicians and non-musicians but grouping according to a good-continuation-of-pitch was only observed in musicians. Grouping of the sequence was indicated by MMN elicitation to an occasional fifth tone that extended the length of the standard-tone groups. This deviating fifth tone extended both the duration of the standard-tone groups and added one more tone to the tone groups. The deviant could therefore have been detected by encoding a temporal and/or a numerical regularity.

In the current study, we tested whether the auditory system can encode a temporal regularity *or* a numerical regularity and whether musicians and non-musicians differ herein. Therefore, we constructed stimuli in which temporal and numerical aspects of the sound sequences varied independently of each other. We are aware that this creates an artificial situation from a musical point of view because time and beat tracking are mostly codependent in music. Tone sequences consisted of segments with multiple consecutive tones of the same pitch. The start of a new segment was indicated by the tones changing to another pitch (see Fig. 1). In the 'time' condition, the duration of a segment was constant (750 ms) while the number of tones that occurred within that interval varied from segment to segment. Occasionally, a deviant violated the temporal regularity by extending the duration of the standard segments beyond 750 ms. In the 'number' condition, the number of tones contained in a segment was constant (four) while the duration of the segments varied. Occasional deviants violated the numerical regularity by adding one more tone to the segments. MMN elicitation by the time or number violations would indicate that the regularity has been extracted from the tone sequence and encoded in the auditory memory underlying MMN elicitation.

If musical expertise gives rise to more sensitive auditory processing we expect that musicians would elicit an MMN in both the 'time' and the 'number' condition but not the non-musicians. On the other hand, if musical expertise leads to honing of specific auditory skills, then we expect that musicians differ from non-musicians in one condition only. Keeping track of the number of beats in a measure is a specifically important aspect of music perception and performance. This could be reflected in a higher sensitivity for encoding numerical regularity in musicians than in non-musicians. If both types of regularity processing are of a more universal nature, we expect that there would be no difference in MMN elicitation between the two groups.

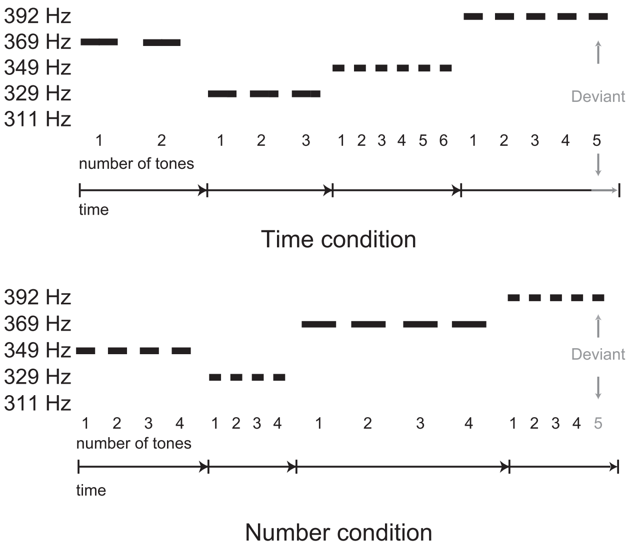

Fig. 1. Schematic illustration of the paradigm. The upper panel illustrates the 'time' condition, in which a variable number of tones of the same pitch followed each other for a fixed period of time (750 ms), followed by a segment of tones of another pitch (roving on five pitch levels). Deviant tones extended the duration of the standard segments beyond 750 ms. The lower panel displays the 'number' condition, in which four tones of one pitch are followed by four tones of another pitch (also roving on five pitch levels). These four-tone segments varied from each other in duration. Deviant tones extended the standard segments by one tone.

2. Methods

2.1. Subjects

Thirteen classically trained musicians (mean 24.7 years, three males) and 15 non-musicians (mean 21.8 years, two males) participated in the experiment. Subjects were classified as musicians if they had reached, as a minimum, the level of acceptance in the Sibelius Music Academy (Helsinki, Finland). On average, the musicians had begun musical training at the age of 6, had been playing at a professional level for 7 years at the time of the experiment and were practicing 4 h a day. One musician reported having absolute pitch. The group of musicians comprised seven pianists, two flutists, two organists, one bassoonist, and one guitarist.

Subjects classified as non-musicians had never played an instrument at a professional level and never received formal musical training. Most of them had received music lessons in school. All subjects were paid for their participation and were naïve with respect to the paradigm.

2.2. Stimuli

Each condition consisted of a certain number of tones of the same pitch (a so-called segment) followed by a certain number of tones of another pitch. The segments varied on five pitch levels, ranging from 311.1 to 392 Hz in semitone steps on the musical scale. Deviants occurred in 10% of the segments.

In the 'time' condition (see the upper panel of Fig. 1), the onset-to-onset time of the segments was always 750 ms whereas the number of tones within that interval varied between two and six. The tones used to construct this sequence varied in duration from 60 to 200 ms and the inter-tone-intervals (ITI) from 58.3 to195 ms (in combinations resulting in segment durations of 750 ms). Within each segment, the individual tone duration and ITI were always the same. For example, a segment could contain two tones of 200 ms each followed by an ITI of 175 ms or a segment could contain five tones of 72 ms each followed by an ITI of 78 ms. The deviant tone extended the segment and lengthened the segment duration beyond 750 ms (e.g., a deviant tone added to the five-tone example extended the segment duration to 900 ms: 6×72 ms tone duration+6×78 ms ITI=900 ms). Only the standard segments with two to five tones could become a deviant so that the total number of tones in a deviating segment was never more than six and was thus not also a number-of-tone violation.

In the 'number' condition (see lower panel of Fig. 1), there were always four tones of the same pitch followed by four tones of another pitch. The duration of the segments varied from 610 to 890 ms. The duration of the individual tones of a tone segment varied from 65 to 100 ms and the ITI from 87.5 to 122.5 ms. Within a segment, the individual tone duration and ITI were always the same. The deviant was an additional tone violating the number of tones in the standard segments but never extended the duration of the segments beyond 890 ms to avoid a temporal violation.

2.3. Procedure

Subjects were comfortably seated in a reclining chair in an acoustically attenuated and electrically shielded room. They were instructed to watch a subtitled movie without sound and to disregard the auditory stimuli. The tones were presented binaurally through headphones at 40 dB above hearing threshold. Six blocks of stimuli, three of each condition, were presented in a counter-balanced order. Stimulus blocks were 10 min and contained 80 deviants. After the experiment, subjects were asked whether they had difficulty ignoring the stimuli. None of the subjects reported any difficulty in disregarding the stimuli.

2.4. Data acquisition

A nose-referenced MMN is typically negative at frontal electrodes, shows diminishing amplitudes towards posterior sites and usually, but not necessarily, reverts polarity at the mastoids. In order to determine the presence of MMN, we recorded the electroencephalogram (EEG) from the following nine electrode sites: three midline positions Fz, Cz, and Pz, at the left and right mastoids (Lm and Rm) and at sites along the coronal chain at one third (L1 and R1) and two thirds (L2 and R2) between Fz and the mastoids, on each side of the head. The reference electrode was placed on the tip of the nose. The horizontal electrooculogram (EOG) was monitored using a bipolar configuration with electrodes placed lateral to the outer canthi of the eyes. Vertical EOG was recorded with electrodes placed above and below the left eye. Signals were amplified between 0.5 and 30 Hz and digitized (Synamps amplifiers) with a 250-Hz sampling frequency.

2.5. Data analysis

In order to remove slow drifts from the EEG and EOG signals, they were digitally filtered offline with a 2-Hz high-pass Butterworth filter (attenuation 12 dB per octave). ERP epochs started 50 ms before and ended 375 ms after stimulus onset. Baseline correction was applied on single trials after which trials that contained electrical activity exceeding ±75 μV at any electrode were rejected. The ERPs that were denoted as 'standard' were calculated from the responses to the last tone of the standard segments (but only from those standard segments that could be extended to a deviant, see Section 2.2). The ERPs that were denoted 'deviant' were averages of the responses to the tones extending the segment beyond the length of the standard segments either in time or by number. On average, the physical properties of the tones eliciting the standard and deviant ERPs were the same.

In order to determine the presence of MMN, the latency at Fz of the negative peak from the group-averaged difference waveform (deviant minus standard) between 120 and 200 ms was determined for each subject group and condition. The mean amplitude in a 32-ms window centered on the peak latency of the group average was calculated for the standard and deviant waveforms for the individual subjects at the different electrodes. For each condition and subject group, one-sided t-tests were conducted in order to determine whether the deviant responses were more negative than the responses to the standards sounds at Fz and the average of L1 an R1. To compare subject groups and conditions, a mixed-model two-way ANOVA of the deviant-minus-standard difference amplitude was performed with factors of Expertise (musicians and non-musicians), Condition (time and number) and Electrode (Fz, L1 and R1). Planned post hoc comparisons were also performed.

After visual inspection of the ERPs elicited by the musicians, a second ANOVA (factor Condition (time and number) and Electrode (Fz, Cz, L1, L2, R1, R2, Pz, Lm and Rm) was performed to test whether the MMN differed in scalp distribution between the 'time' and the 'number' condition.

3. Results

3.1. Time condition

In Fig. 2, the group-averaged ERPs recorded in the 'time' condition are displayed for both musicians and non-musicians. In both subject groups, a significant MMN was elicited. This can be seen from the frontally more negative deflections peaking at 178 ms post-stimulus at Fz in musicians ($t(12)=2.7$, $p<0.009$; $t(12)=3.4$, $p<0.0025$, Fz and the average of L1 and R1, respectively) and also at 178 ms at Fz in non-musicians ($t(14)=2.0$, $p<0.03$; $t(12)=1.7$, $p<0.05$, Fz and the average of L1 and R1, respectively) in the deviant waveform compared to the standard waveform. These results indicate that the auditory system of both musicians and non-musicians have extracted and encoded the presented temporal regularity.

3.2. Number condition

In Fig. 3, the group-averaged ERPs recorded in the 'number' condition are displayed for both musicians and non-musicians. In the musicians, an MMN was elicited as indicated by a frontal negative difference between the standard and the deviant ($t(12)=2.6$, $p<0.01$; $t(12)=2.6$, $p<0.01$, Fz and the average of L1 and R1, respectively) peaking on Fz at 146 ms. For the non-musicians, there was no indication of an MMN ($t(14)=0.4$, $p<0.36$; $t(12)=0.6$, $p<0.29$). These results indicate that only the auditory system of musicians could extract and encode the numerical regularity presented in this condition.

3.3. Differences between level of expertise and condition

A main effect of Expertise was found ($F(1,26)=5.2$, $p<0.03$) with the musicians having larger deviant-minus-standard differences than the non-musicians. This effect was mainly caused by a significant MMN in the musicians compared to no MMN in the non-musicians for the 'number' condition (planned post hoc comparison:

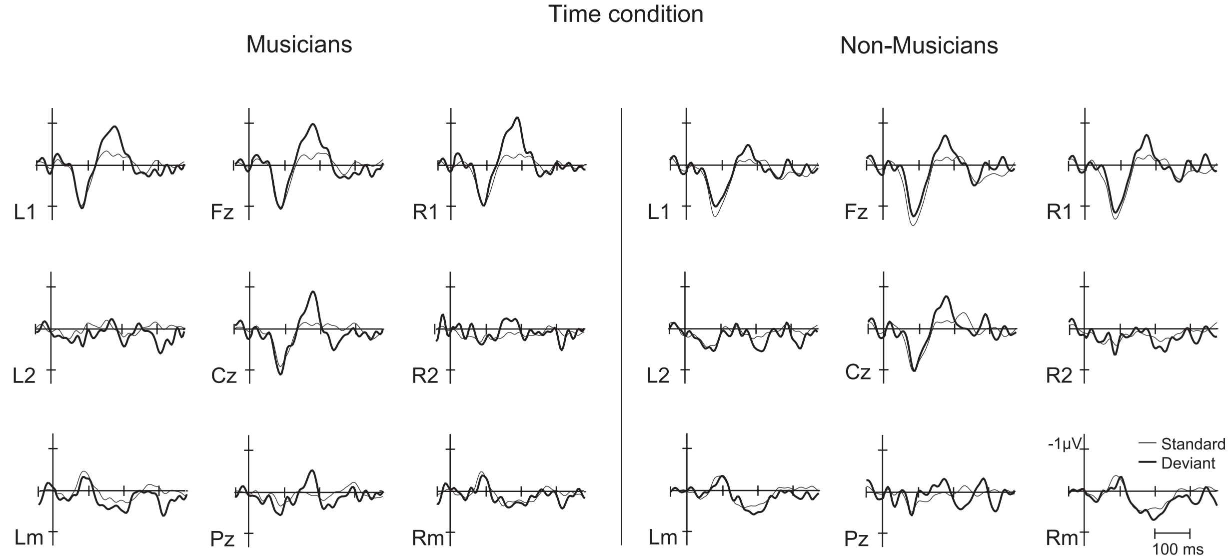

Fig. 2. ERPs elicited in the 'time' condition. Group-averaged ERPs are shown for the standard (thin line) and deviant (thick line) tones at all recording channels, for the musicians (left) and for the non-musicians (right).

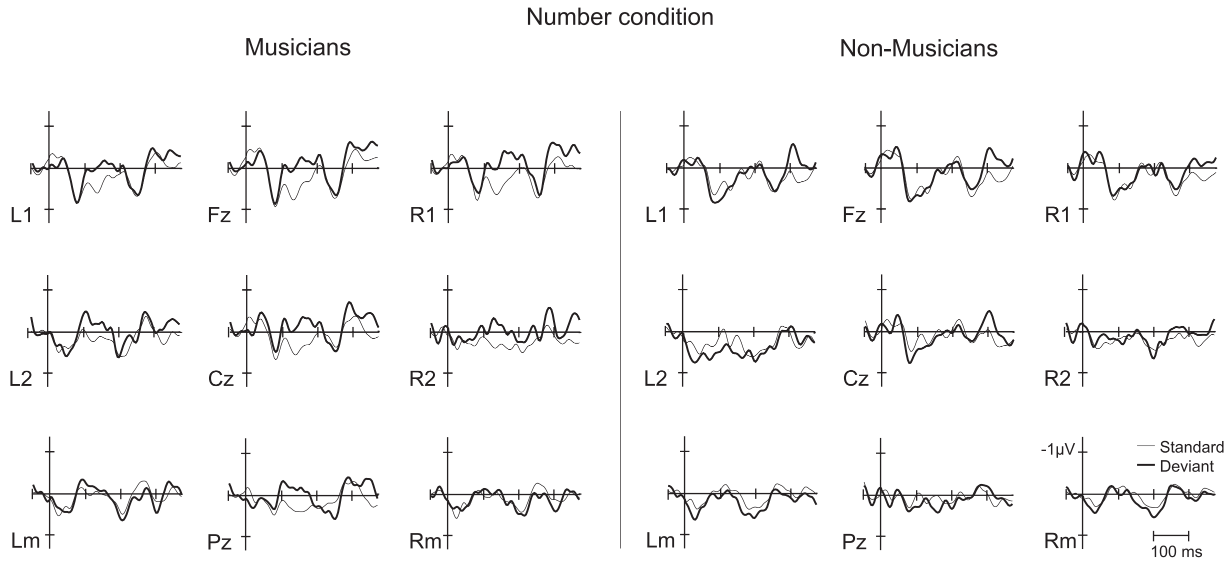

Fig. 3. ERPs elicited in the 'number' condition. Group-averaged ERPs are shown for the standard (thin line) and deviant (thick line) tones at all recording channels, for the musicians (left) and for the non-musicians (right).

F(1,26)=4.3, p<0.05). There was no main effect of condition and no interaction between Expertise and Condition. Taken together, there is an overall enhanced response for musicians compared to non-musicians partially caused by the significantly larger response of the musicians in the 'number' condition (Fig. 4).

3.4. Scalp distributions

The MMN response elicited by the musicians in the 'time' condition is frontally negative with declining amplitude towards the posterior electrode sites and appears with a reverted polarity at the mastoids. This scalp distribution is the one most commonly observed for the MMN. In the 'number' condition, the MMN amplitude looks more equal on the three midline sites and we do not see a clear polarity reversal at the mastoids. This scalp distribution resembles the MMN elicited to violations of abstract regularities [16]. However, the ANOVA testing for differences in scalp distribution did not yield a significant interaction between Electrode and Condition (*F*(8,96)=0.5, p<0.9; after amplitude normalization *F*(8,96)=0.8, p<0.6, [9]) only a main effect of Electrode (*F*(8,96)=11.0, p<0.0001; Greenhouse Geisser adjusted). We therefore cannot conclude that the scalp distributions differ between the two conditions.

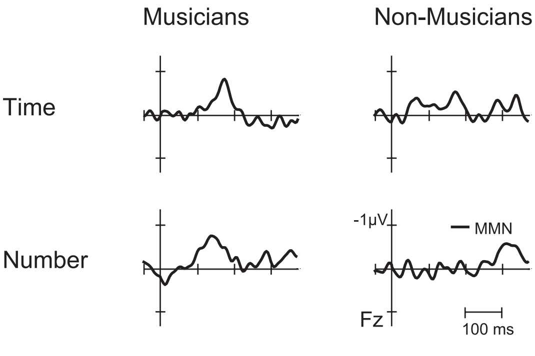

Fig. 4. MMN difference waves. Deviant-minus-standard waveforms on Fz, for the musicians (left) and for the non-musicians (right) for both conditions.

4. Discussion

We investigated whether the auditory system of musicians is more sensitive than that of non-musicians in encoding a temporal or numerical regularity in a complex tone sequence that was not made relevant by a task. We found that for encoding this particular temporal regularity musical expertise was not a necessary prerequisite. This was demonstrated by the MMN being elicited in both groups of subjects in the 'time' condition. This is consistent with the notion that temporal parameters of sound are generally relevant to auditory processing, e.g., to the analysis of speech [21,28] and not only important to the perception of music.

The encoding of the numerical regularity depended on musical expertise because an MMN was elicited in musicians in the 'number' condition only. Musicians can organize a sound sequence according to a number regularity even though the tone sequence was not made relevant by any task instruction. The musicians were not actively counting to four or five but the auditory system could nevertheless encode numerosity [10] as it could distinguish the segments containing four tones from the segments containing five tones. This type of auditory specialization is in line with the specific auditory processing demands of professional musicians. Keeping track of the number of beats in a measure is an integral part of music perception and performance. The perceptual organization of sound according to number does not have relevance for non-musicians and their auditory system was not able to encode the four-tone regularity. It can, however, not be excluded that the auditory system of non-musicians would be able to encode numerosity in a different, perhaps more simple situation.

The stimuli used in this study contained temporal and numerical aspects that varied independently of each other, unlike most music in which temporal and numerical aspects are co-varying. To estimate measure duration in music, both temporal (note duration) and numerical cues (number of beats) can be used. The finding that the auditory system of musicians can keep track of number independent of temporal processing cues indicates the importance of number processing for the perceptual organization of music. We have demonstrated a specific isolated ability of the auditory system of musicians to detect numerical regularities.

Visual inspection of the MMN scalp distributions of the time and number condition in musicians suggested that the MMNs had different neural generators. We did, however, not obtain a statistical confirmation of this. Future research specifically aimed at characterizing neural generators will be needed to shed light on whether the neural generators of MMNs to numerical regularity violations differ from the generators of MMNs to other types of regularity violations.

In conclusion, the auditory system can encode a temporal regularity even when formed by a varying number of consecutive tones irrespective of musical skill. This indicates that organizing sound by tracking time is part of common involuntary auditory processing abilities. The auditory system of musicians is, in addition, specialized to encode numerosity and organize sound by keeping track on number.

Acknowledgements

This research was supported by The Academy of Finland (77322, 79820, 79821), the Finnish Graduate School of Psychology, the National Institutes of Health grant (R01 DC04263) and the Hungarian National Research Fund (OTKA 034112).

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

[3] M. Atienza, J.L. Catero, C. Grau, C. Gomez, E. Dominguez-Marin, C. Escera, Effects of temporal encoding on auditory object formation: a mismatch negativity study, Cogn. Brain Res. 16 (2003) 359–371.

[4] M. Besson, F. Faita, J. Requin, Brain waves associated with musical incongruities differ for musicians and non-musicians, Neurosci. Lett. 168 (1994) 101–105.

[5] E.M. Burns, A.J.M. Houtsma, The influence of musical training on the perception of sequentially presented mistuned harmonics, J. Acad. Soc. Am. 106 (1999) 3564–3570.

[6] C. Drake, M. Botte, Tempo sensitivity in auditory sequences: evidence for a multiple-look model, Percept. Psychophys. 54 (1993) 277–286.

[7] C. Drake, A. Penel, E. Bigand, Tapping in time with mechanically and expressively performed music, Music Percept. 18 (2000) 1–23.

[8] S. Koelsch, E. Schröger, M. Tervaniemi, Superior attentive and pre-attentive auditory processing in musicians, NeuroReport 10 (1999) 1309–1313.

[9] G. McCarthy, C.C. Wood, Scalp distribution of event-related potentials: an ambiguity associated with analysis of variance models, Electroencephalogr. Clin. Neurophysiol. 62 (1985) 203–208.

[10] E.K. Miller, A. Nieder, D.J. Freedman, J.D. Wallis, Neural correlates of categories and concepts, Curr. Opin. Neurobiol. 13 (2003) 198–203.

[11] C. Mizuno, J.L. Schwartz, Y. Cazals, Periodicity of long-term context can influence gap detection, Hear. Res. 78 (1994) 41–48.

[12] T.F. Münte, C. Kohlmetz, W. Nager, E. Altenmuller, Superior auditory spatial tuning in conductors, Nature 409 (2001) 580.

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

[14] R. Näätänen, M. Tervaniemi, E. Sussman, P. Paavilainen, I. Winkler, 'Primitive intelligence' in the auditory cortex, Trends Neurosci. 24 (2001) 283–288.

[15] H. Nordby, W.T. Roth, A. Pfefferbaum, Event-related potentials to breaks in sequences of alternating pitches or interstimulus intervals, Psychophysiology 25 (1988) 262–268.

[16] P. Paavilainen, J. Simola, M. Jaramillo, R. Näätänen, I. Winkler, Preattentive extraction of abstract feature conjunctions from auditory stimulation as reflected by the mismatch negativity (MMN), Psychophysiology 38 (2001) 359–365.

[17] C. Pantev, R. Oostenveld, A. Engelien, B. Ross, L.E. Roberts, M. Hoke, Increased auditory cortical representation in musicians, Nature 23 (1998) 811–814.

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

[19] W. Ritter, D. Deacon, H. Gomes, D.C. Javitt, H.D. Vaughan Jr., The mismatch negativity of event-related potentials as a probe of transient auditory memory: a review, Ear Hear. 16 (1995) 52–67.

[20] J. Rüsseler, E. Altenmüller, W. Nager, C. Kohlmetz, T.F. Münte, Event-related brain potentials to sound omissions differ in musicians and non-musicians, Neurosci. Lett. 308 (2001) 33–36.

[21] R.V. Shannon, F.G. Zeng, V. Kamath, J. Wygonski, M. Ekelid, Speech recognition with primarily temporal cues, Science 270 (1995) 303–304.

[22] E. Sussman, W. Ritter, H.G. Vaughan Jr., Predictability of stimulus deviance and the mismatch negativity, NeuroReport 9 (1998) 4167–4170.

[23] E. Sussman, W. Ritter, H.G. Vaughan Jr., An investigation of the auditory streaming effect using event-related brain potentials, Psychophysiology 36 (1999) 22–34.

[24] E. Sussman, I. Winkler, M. Huoutilainen, W. Ritter, R. Näätänen, Top-down effects on the initially stimulus-driven auditory organization, Cogn. Brain Res. 13 (2002) 393–405.

[25] M. Tervaniemi, M. Rytkönen, E. Schröger, R.J. Ilmoniemi, R. Näätänen, Superior formation of cortical memory traces for melodic patterns in musicians, Learn. Mem. 8 (2001) 295–300.

[26] L.J. Trainor, R.N. Desjardins, C. Rockel, A comparison of contour and interval processing in musicians and nonmusicians using event-related potentials, Aust. J. Psychol. 51 (1999) 147–153.
[27] L.J. Trainor, K.L. McDonald, C. Alain, Automatic and controlled processing of melodic contour and interval information measured by electrical brain activity, J. Cogn. Neurosci. 14 (2003) 430–442.
[28] D.J. van Tasell, S.D. Soli, V.M. Kirby, G.P. Widin, Speech waveform envelope cues for consonant recognition, J. Acoust. Soc. Am. 82 (1987) 1152–1161.
[29] T.L. van Zuijen, E. Sussman, I. Winkler, R. Näätänen, M. Tervaniemi, Grouping of sequential sounds—an event-related potential study comparing musicians and non-musicians, J. Cogn. Neurosci. 16 (2004) 331–338.