Simultaneous pitches are encoded separately in auditory cortex: an MMNm study

Takako Fujioka[a], Laurel J. Trainor[a,b,c] and Bernhard Ross[a]

[a]Rotman Research Institute, Baycrest, University of Toronto, [b]Department of Psychology, Neuroscience and Behaviour and [c]McMaster Institute for Music and the Mind, McMaster University, Hamilton, Canada

Correspondence to Takako Fujioka, PhD, Rotman Research Institute, 3560 Bathurst Street, Toronto, Ontario, M6A 2E1, Canada
Tel: + 1 416 785 2500 x 3413; fax: + 1 416 785 2862; e-mail: tfujioka@rotman-baycrest.on.ca

This study examined whether two simultaneous pitches have separate memory representations or an integrated representation in preattentive auditory memory. Mismatch negativity fields were examined when a pitch change occurred in either the higher-pitched or the lower-pitched tone at 25% probability each, thus making the total deviation rate of the two-tone dyad 50%. Clear MMNm was obtained for deviants in both tones confirming separate memory traces for concurrent tones. At the same time, deviants to the lower-pitched, but not higher-pitched, tone within the two-tone dyad elicited a reduced MMNm compared to when each tone was presented alone, indicating that the representations of two pitches are not completely independent. *NeuroReport* 19:361–366 © 2008 Wolters Kluwer Health | Lippincott Williams & Wilkins.

Keywords: auditory cortex, auditory scene analysis, chord, magnetoencephalography, mismatch negativity, pitch

Introduction

Auditory scene analysis involves two automatic complementary processes of segregating sounds into concurrent objects (or streams) and integrating sounds into a single object (or stream) [1] based on sound properties such as frequency, pitch, timbre, and temporal synchrony. These processes can be investigated with the mismatch negativity (MMN) component in the evoked potential or its magnetoencephalographic (MEG) counterpart, the mismatch negativity field (MMNm). MMNm is elicited mainly in auditory cortices in response to occasional changes (deviants) in the auditory environment and reflects memory traces that encode invariant aspects of the recent acoustic past [2,3]. MMNm becomes larger with an increased size of deviation and a decreased rate of deviant occurrence. Memory traces can store different acoustic features concurrently. MMN was found in response to deviation in each of five different acoustic features of a single repeating tone, with each feature altered at a rate of 10%, despite the global deviance rate of 50% [4]. Previous studies have also shown that tone sequences containing tones with disparate pitch levels are encoded into multiple memory traces, as indexed by MMN responses [5,6]. For example, in alternating high- and low-pitched tones (e.g. H-L-H-L…), pitch deviations in one pitch level (H—H—…) produce MMN regardless of the number of deviants in the other pitch level (L—L—…) [6], suggesting separate memory traces. MMN, however, is reduced in the alternating two-tone case compared to the case where only one pitch is presented, suggesting that encoding of separate pitch levels is not completely independent [6].

Perceptual stream segregation, the phenomenon of two alternating tones being perceived as two separate streams, depends on an interaction between stimulus-driven parameters (bottom-up process) and a listener's intention (top-down process) [1]. That is, one can choose between listening to an integrated stream or two segregated streams when the presentation rate of the tones is slow and when the pitch interval between the tones is not too close or too far apart. The variety of MMN results reflects this ambiguity. MMN was elicited without focused attention to the pitch changes in alternating tones at a fast rate that promoted strong segregation, while MMN was absent at a slower rate [5], unless participants were instructed to attend one of the streams [7]. MMN, however, was still generated even when participants did not experience strong perceptual segregation in passive listening [6]. Focusing on an auditory detection task in one stream suppressed the MMN responses to changes in the other two streams, even though they were clearly present in an unattended condition [8]. Thus, it appears that the memory trace system does not reflect either bottom-up or top-down process exclusively. Rather, it might function to optimize auditory analysis needed for subsequent higher cognitive processes which extract the 'meaning' of sounds such as speech and music.

In the real-world, understanding music requires segregation of *simultaneous* sounds as well as *alternating* tones. To date, it has not been investigated whether two *simultaneous* tones of different pitch are encoded separately in auditory memory. MMNm has been shown in response to a single pitch change within a musical chord of several pure tones [9,10], but it remains open as to whether the MMNm was elicited by a change in separate representations for each concurrent tone, or in their unified representation, or a combination of the two. To address this question in the present study, we make use of the fact that MMNm

decreases with increased probability of a deviant and set up a situation where the rate of pitch change in individual tones is 25%, but the global rate across simultaneous unified tones is 50%. If each tone has a separate memory representation, MMNm would be expected, as the deviance rate would be 25% for each tone; if there is only a unified representation, a small, or no MMNm would be expected, as the deviance rate would be 50%. Our previous study partially addressed this issue [11] by using two five-note melodies presented synchronously (e.g., five different combinations of two pitches presented in a row), with 25% deviants in each melody for a global rate of 50%. A significant MMNm was obtained for deviants in each melody, confirming separate encoding of each melody, although MMNm was larger for changes in the higher-pitched than in the lower-pitched melody in both musicians and nonmusicians. This was consistent with behavioural data showing the perceptual dominance of the highest melody in multivoiced music [12,13]. It was, however, not clear whether two melodies are required for separate traces, or whether each tone of a single simultaneous tone pair also would be encoded separately. Thus, in the present study, we tested whether two separate pitch representations exist when two notes are simultaneously presented, and whether the higher-pitched tone has a more robust representation than the lower-pitched tone as it does in a melodic context.

We used a fully crossed design to compare two-tone with single-tone conditions, matching pitch level (high, low) and deviant rate (two-deviant, one-deviant) as described in Fig. 1. In the two-tone two-deviant condition, two repeating pitches were presented simultaneously as a two-tone dyad (Fig. 1a). For one deviant, the high-pitched tone was raised 2 semitones; for the other deviant the low-pitched tone was lowered 2 semitones. Despite the 50% global deviation rate, each tone had a deviance rate of 25%. Thus, separate tone encoding would result in MMNm to both deviants. Each repeating tone of this original stimulus was presented alone (Fig. 1c), thereby having a local and global deviance rate of 25% to test the extent to which simultaneous tones are encoded separately. Thus, MMNm in these 'alone' conditions were compared to those when both tones were presented simultaneously. In order to confirm that MMNm is weak with an overall deviance rate of 50% in a single standard tone, the upward and downward pitch changes

were applied in a single stream (Fig. 1b). Finally, these two deviants were applied separately to ensure that MMNm was obtained with global deviance rates of 25% (Fig. 1d). From these comparisons, we examined the extent to which individual representations exist for simultaneously presented tones of different pitch.

Methods

Eleven right-handed adults (7 women, 24–37 years of age, mean 29.7) with normal hearing (250–8000 Hz) and without history of neurological and psychological disorders participated after giving informed consent. None had post-secondary musical education. The Research Ethics Board at Baycrest Centre approved the study.

The four conditions are outlined in Fig. 1 as described in the Introduction. Tones were 300 ms computer-synthesized piano tones (Creative SB), presented with stimulus onset asynchrony (SOA) of 750 ms. Each sequence was 5.6 min, containing 450 stimuli in pseudorandom order, avoiding identical deviants in a row. Standard tones in the two-tone conditions had fundamental frequencies of 466.2 Hz (B-flat-4, international standard notation) and 196.0 Hz (G3), which are 15 semitones apart and form a minor third interval with an additional octave. The two-semitone deviations result in 17 semitone intervals, which comprise a perfect fourth interval with an octave. Note that since we used Equal-temperament tuning (12 semitones=1 octave), even the 17-semitone interval was not perfectly consonant. A minor third interval is widely used in current Western tonal music, and considered to be reasonably consonant even though it is less consonant than an octave or perfect fourth. The standard pitch in the one-tone conditions was 293.7 Hz (D4), midway between the standards in the two-tone conditions. Intensity was set 60 dB above the thresholds for each ear for the D4 note. The order of conditions was counterbalanced across participants.

Neuromagnetic fields were recorded with a 151-channel whole-cortex magnetometer (OMEGA, VSM MedTech, Coquitlam, Canada) in a quiet magnetically shielded room, after 100 Hz lowpass filtering at a sampling rate of 312.5 Hz. The participants were seated in an upright position and instructed to stay awake but to pay no specific attention to the stimuli while watching a subtitled movie.
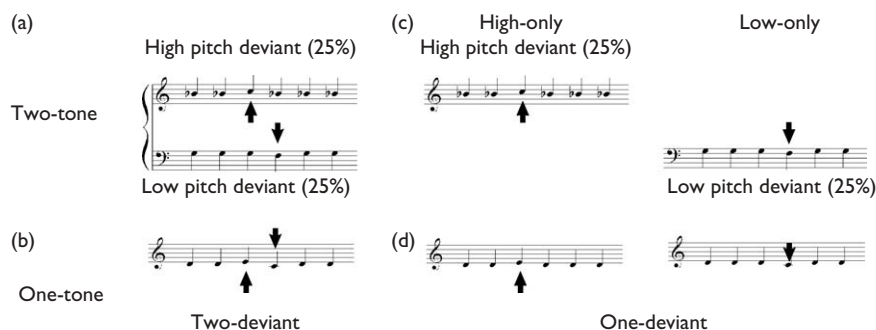
Fig. I Stimulus sequences illustrated in musical notation. (a) Two-tone two-deviant condition. The standard stimulus was a pair of two notes, B-flat 4 and G3 (fundamental frequency of 466.2 and 196.0 Hz). In one deviant, the pitch of the higher note was raised by two semitones (C5, indicated by the upward arrow), while in the other the pitch of the lower tone was lowered by two semitones (F3, indicated by the downward arrow) indicated by arrows. (b) One-tone two-deviant condition. A single note, D4 (293.7 Hz) was used. As in the two-tone case, the deviants went up (E4) or down (C4) by two semitones. (c) Two-tone one-deviant conditions. The high-only and low-only sequences were derived by separating the tones in the two-deviant case. (d) One-tone one-deviant condition. Two sequences were derived by including either only the high deviant or only the low deviant from the one-tone case.

Their compliance was verified by video monitoring. The MEG data were segmented into 750-ms epochs including a 150-ms prestimulus interval. Trials contaminated with eye-blink or movement artifacts were rejected from averaging based on a 1.5 pT threshold criteria, resulting in 93.6% accepted trials. Averaged data across conditions in individual participants were used to estimate equivalent current dipoles (ECD) in the left and right auditory cortices using the whole evoked response. A dipole was accepted based on the criteria of goodness-of-fit more than 85% and being located in auditory area overlaid to individual structural magnetic resonance images (MRI) of the brain, acquired with a 1.5 T scanner (Signa, General Electric Medical Systems, Waukesha, WI). On the basis of these dipoles, the signal space projection method (SSP) [14] extracted standard, deviant, and difference waveforms (deviant minus standard) for the auditory cortical sources for each condition. Offset correction based on the prestimulus interval and 30-Hz low-pass filtering was applied. The 99% confidence intervals for the grand-averaged evoked responses were estimated from nonparametric bootstrap resampling [15] and served as indices for the noise level. The same technique was used to examine significant differences between conditions.

Individual MMNm peak latency was identified in the 90 to 200 ms interval. MMNm amplitudes were defined as the mean across a 40 ms interval around the peak latency of the grand-average waveform for each condition. The amplitude and latency in the two-deviant cases were assessed by a repeated measures analysis of variance (ANOVA) with three factors: Number-of-tones (one, two), Deviance-type (high, low), and Hemisphere (left, right). For comparison between the two-deviant and the one-deviant conditions, three-way ANOVAs with the factors Number-of-deviants (one, two), Deviance-type (high, low), and Hemisphere (left, right) were performed separately for the two-tone and the one-tone conditions. Post-hoc comparisons used Fisher's PLSD test at 5% level of significance.

Results

MMNm is larger and later when two deviants are spread across two simultaneous tones than when they are both in a single tone

MMNm was larger and later for two-tone (Fig. 1a) compared to one-tone (Fig. 1b) conditions as illustrated in Fig. 2, even though in both cases there were two deviants, each presented on 25% of trials, for an overall deviance rate of 50%. This provides evidence that separate memory traces exist for the two simultaneous tones. The ANOVA for MMNm amplitude revealed a main effect of Number-of-tones [$F(1,10)=5.42$, $P<0.04$] because of the larger response in the two-tone case (6.57 nAm) than in the one-tone case (3.74 nAm). No other main effects or interactions were significant. The peak latency [$F(1,10)=7.9$, $P=0.018$] was longer in the two-tone case (132 ms) than in the one-tone case (115 ms). Hemisphere was significant [$F(1,10)=10.2$, $P=0.009$] due to a shorter latency in the left (114 ms) than in the right (132 ms). The interaction of Number-of-tones × Deviance-type × Hemisphere [$F(1,10)=9.8$, $P=0.011$] was explained by the absence of a hemispheric difference only for the low deviant in the one-tone case.

MMNm is similar to a deviant in a single tone and the higher of two simultaneous notes, but MMNm is reduced in the lower of two simultaneous tones

Figure 3a shows MMNm in the two-tone condition with 25% deviants in each pitch (Fig. 1a) overlaid with MMNm in the single-tone condition (Fig. 1c). The responses to deviants in the higher-pitched tones were almost identical regardless of the presence or absence of the lower-pitched tones, while the responses to the deviants in the lower-pitched tones were larger and earlier, especially in the left hemisphere when the higher tone was absent compared to when it was present. This illustrates that the memory trace for the lower tone is affected by the presence of the higher tone, but not vice versa. For latency, Hemisphere was significant

Difference waveforms (two-deviant conditions)
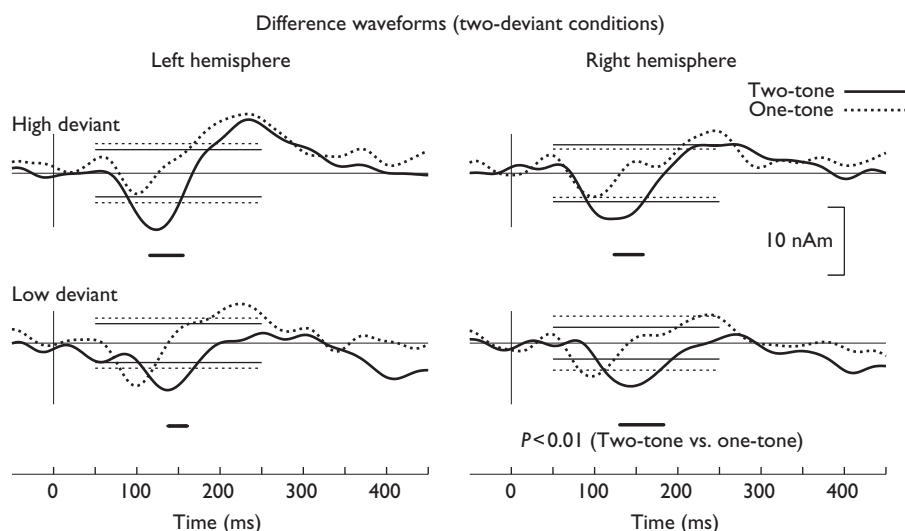
Fig. 2 Grand averaged difference waveforms in left and right hemispheres for the two-deviant conditions, plotted separately for the high- and the low-pitched deviants, with two-tone (Fig. la) and one-tone (Fig. lb) conditions overlaid. The horizontal lines above and below zero show the upper and lower limits of 99% confidence interval for in the two-tone condition (thin line) and one-tone condition (thin dotted line) as indices of noise level at the whole time interval, thus showing that the waveform exceeding these lines are significantly different from zero. The horizontal bar below each difference waveform indicates time points where the difference between two-tone and one-tone conditions was significant based on 99% confidence limits.
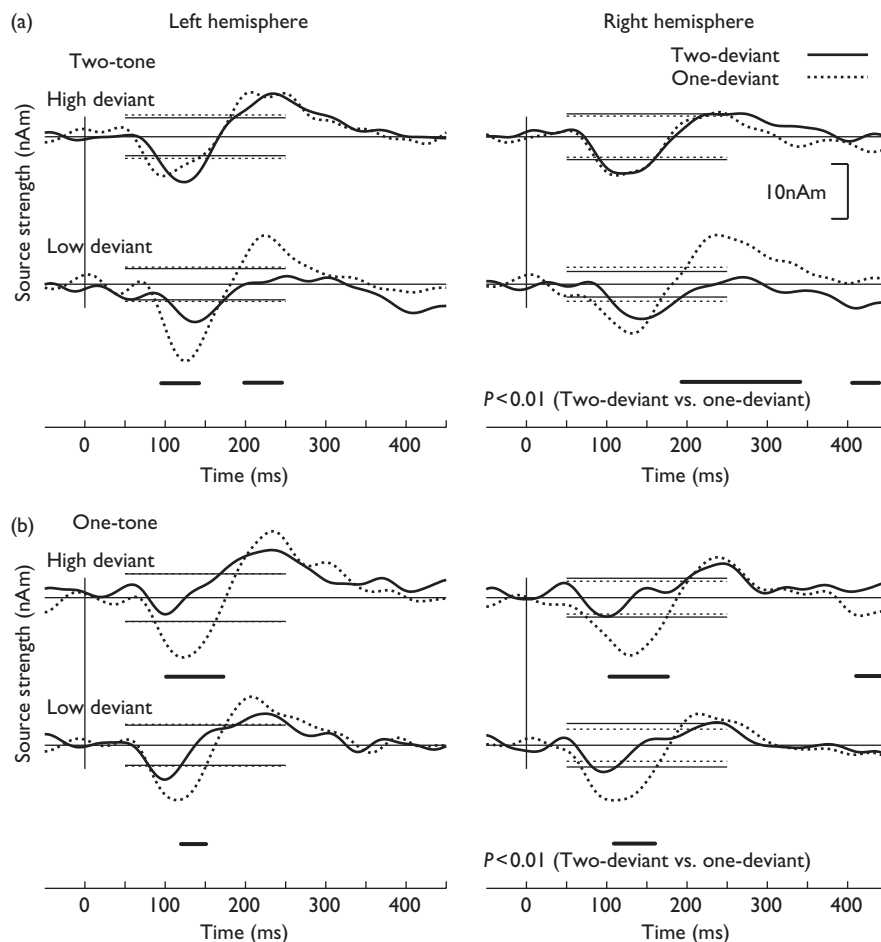
Fig. 3 (a) Grand averaged difference waveforms for the two-tone two-deviant condition, plotted separately for the high and low deviants (thick lines), and the corresponding separate tone one-deviant conditions (dotted lines). The horizontal lines above and below zero show the upper and lower limits of 99% confidence interval for in the two-deviant condition (thin line) and one-deviant condition (thin dotted line) as indices of noise level. (b) Grand averaged difference waveforms for the one-tone two-deviant condition, plotted separately for the high and low deviants, and the corresponding one-deviant conditions. The black horizontal bar below each trace indicates time intervals of significant difference between the two responses based on 99% confidence limits.

$[F(1,10)=18.3, P=0.002]$, with peaks 13 ms earlier in the left than in the right. The interaction of Number-of-deviants $\times$ Hemisphere $[F(1,10)=5.64, P<0.039]$ was caused by shorter latencies in the left hemisphere only in the two-deviant case $(P<0.01)$, as described in the previous section with the effect expressed to a greater degree for low than for high deviants. The latter contributed also to the interaction of Number-of-deviants $\times$ Deviance-type $\times$ Hemisphere $[F(1,10)=5.26, P<0.045]$, which arose because the hemispheric difference was not present in the one-deviant case.

Within a single tone, MMNm is larger for a single deviant (25% probability) than for two deviants
Figure 3b shows MMNm in one-tone conditions, demonstrating smaller responses in two-deviant (Fig. 1b) than in the corresponding one-deviant conditions (Fig. 1d) around 120–150 ms, as predicted by the global 50 and 25% deviance rates in the two conditions, respectively. This was confirmed by the ANOVA, revealing a main effect of Number-of-deviants $[F(1,10)=6.58, P=0.028]$ with a larger MMNm for the one-deviant (9.71 nAm) than for the two-deviant

conditions (3.74 nAm). No other main effects or interactions were found. For latency, there was a tendency for an effect of Number-of-Deviants $(P=0.062)$ with earlier peak (115 ms) for the two-deviant condition than for the one-deviant condition (126 ms).

Discussion
When two pitch changes (25% probability of each change) are spread across two repeating simultaneously presented tones, MMNm is larger and later compared to when the two pitch changes are contained in a single repeating tone, despite the same global deviation rate of 50% (Fig. 2). This indicates separate memory traces for each of the two *simultaneous* pitches at the level of auditory cortex. Previously, the memory trace system has been shown to encode *sequential* high and low tones separately [5,6], and to extract the interval between two simultaneous pitches regardless of the absolute pitch level [16]. Thus, our data extend these findings by showing that separate pitch representations exist for each tone of a simultaneous dyad, and that these representations likely coexist with an

integrative process. Our stimuli used an interval between the tones that was wider than that used by Paavilainen *et al.* [16]. An interesting question for future research, therefore, is how interval size affects separate and integrated representations.

A significant reduction in MMNm magnitude for the 50% compared to the 25% global deviation rate was found in the case of a single tone (Fig. 3b), but not in the case of two simultaneous tones (Fig. 3a). This strengthens the support for separate memory traces for each tone in the two-tone stimulus. For a single tone, it has been repeatedly shown that a decreased number of standard stimuli before a deviant results in a decreased amplitude of MMN in both frontal and temporal MMN components [17,18]. Our data replicated these reports for the component of MMNm originating from auditory cortex (temporal component).

The presence of a concurrent tone attenuated MMNm to deviants to the lower-pitched tone but not to deviants to the higher-pitched tone (Fig. 3a), indicating that the encoding of the lower-pitched tone is less robust when presented with a higher pitch, as we found previously in a two-melody context [11]. As we recorded brain response but not behavioural measures, we do not know whether the two tones were perceived differently. It has, however, been shown behaviourally that the degree of perceptual distinctiveness of simultaneous tones depends on a number of factors including consonance, relative pitch height, and musical experience [19,20]. We used piano-timbre tones, each of which elicits a clear pitch perception without separate perception of the harmonics. Furthermore, we used a widely separated interval (15 semitones) between the tones, which was not perfectly consonant. These factors likely contributed to separation of the two pitch representations in memory, and the individuality of the tones in perception.

The difference in encoding strength between the high and low tones reported here is unlikely to be the result of peripheral encoding. Asymmetric shape of the tuning curves of the auditory nerve around a centre frequency predicts a lower-pitch dominance, because low frequencies produce greater masking on high frequencies than vice versa [21]. This is actually reflected in our results showing that in the single-tone case MMNm was greater for downward than for upward pitch changes. Thus, the finding that MMNm to deviants in the lower tone, but not to deviants to the higher tone, was reduced in magnitude by the presence of the concurrent tone suggests that the lower-pitched of two concurrent tones is not encoded entirely independently from the higher-pitch tone. This also suggests a possible interaction between memory traces for simultaneous tones, consistent with previous literature showing similar reduction of MMN in multiple streams compared to a single stream alone [6]. Moreover, deviants in single-tone sequences involving different sound features produce smaller MMN than predicted from the summation of responses to each deviant presented alone [22].

Throughout the results, MMNm tended to be earlier in the left than in the right hemisphere. This is in contrast to the data of Tervaniemi *et al.* [10], who reported larger MMNm responses on the right to a change in one frequency of a chord consisting of four pure tones, without any difference in MMNm latency. It is possible that their study elicited processing related to a change in timbre, whereas the present study elicited processing related to individual tone tracking. For example, MMN to duration change in tones, which requires such a tracking process, was attenuated in patients with left-hemisphere damage [23] but not with right [24]. There is, however, no prior evidence of stream segregation causing left-lateralized response in MMN [5] or obligatory auditory evoked magnetic fields [25]. The full interpretation of the lateralization results must thus await further study.

Conclusion

We demonstrate that at the level of preconscious memory traces in the auditory cortex, two concurrent pitches (which are not perfectly consonant) are encoded separately to a large extent, but that the lower tone is encoded less robustly when in the presence of the higher tone. These results indicate that the two separate memory traces are not entirely independent, and that the emergence of a unified entity in the form of an interval is likely occurring by this stage of processing.

Acknowledgements

References

1. Bregman AS. *Auditory scene analysis*. Cambridge, MA: MIT Press; 1990.
2. Näätänen R. *Attention and brain function*. Hillsdale, NJ: Erlbaum; 1992.
3. Picton TW, Alain C, Otten L, Ritter W, Achim A. Mismatch negativity: different water in the same river. *Audiol Neurootol* 2000; **5**:111–139.
4. Näätänen R, Pakarinen S, Rinne T, Takegata R. The mismatch negativity (MMN): towards the optimal paradigm. *Clin Neurophysiol* 2004; **115**: 140–144.
5. Sussman E, Ritter W, Vaughan HG Jr. An investigation of the auditory streaming effect using event-related brain potentials. *Psychophysiology* 1999; **36**:22–34.
6. Shinozaki N, Yabe H, Sato Y, Sutoh T, Hiruma T, Nashida T, *et al*. Mismatch negativity (MMN) reveals sound grouping in the human brain. *NeuroReport* 2000; **5**:1597–1601.
7. Sussman E, Ritter W, Vaughan HG Jr. Attention affects the organization of auditory input associated with the mismatch negativity system. *Brain Res* 1998; **789**:130–138.
8. Sussman ES, Bregman AS, Wang WJ, Khan FJ. Attentional modulation of electrophysiological activity in auditory cortex for unattended sounds within multistream auditory environments. *Cogn Affect Behav Neurosci* 2005; **5**:93–110.
9. Alho K, Tervaniemi M, Huotilainen M, Lavikainen J, Tiitinen H, Ilmoniemi RJ, *et al*. Processing of complex sounds in the human auditory cortex as revealed by magnetic brain responses. *Psychophysiology* 1996; **33**:369–375.
10. Tervaniemi M, Kujala A, Alho K, Virtanen J, Ilmoniemi RJ, Näätänen R. Functional specialization of the human auditory cortex in processing phonetic and musical sounds: a magnetoencephalographic (MEG) study. *Neuroimage* 1999; **9**:330–336.
11. Fujioka T, Trainor LJ, Ross B, Kakigi R, Pantev C. Automatic encoding of polyphonic melodies in musicians and non-musicians. *J Cogn Neurosci* 2005; **17**:1578–1592.
12. Crawley EJ, Acker-Mills BE, Pastore RE, Weil S. Change detection in multi-voice music: the role of musical structure, musical training, and task demands. *J Exp Psychol Hum Percept Perform* 2002; **28**:367–378.
13. Zenatti A. Le developpement genetique de la perception musicale chapter II, La Perception Polyphonique. *Monog Francais Psychol* 1969; **17**:33–47.
14. Tesche CD, Uusitalo MA, Ilmoniemi RJ, Huotilainen M, Kajola M, Salonen O. Signal-space projections of MEG data characterize both distributed and well-localized neuronal sources. *Electroencephalogr Clin Neurophysiol* 1995; **95**:189–200.

15. Efron B, Tibshirani RJ. *An introduction to the bootstrap*. Boca Raton: Chapman & Hall; 1998.

16. Paavilainen P, Jaramillo M, Näätänen R, Winkler I. Neuronal populations in the human brain extracting invariant relationships from acoustic variance. *Neurosci Lett* 1999; **265**:179–182.

17. Sams M, Alho K, Näätänen R. Short-term habituation and dishabituation of the mismatch negativity of the ERP. *Psychophysiology* 1984; **21**:434–441.

18. Matuoka T, Yabe H, Shinozaki N, Sato Y, Hiruma T, Ren A, *et al*. The development of memory trace depending on the number of the standard stimuli. *Clin EEG Neurosci* 2006; **37**:223–229.

19. DeWitt LA, Samuel AG. The role of knowledge-based expectations in music perception: evidence from musical restoration. *J Exp Psychol Gen* 1990; **119**:123–144.

20. Platt JR, Racine RJ. Perceived pitch class of isolated musical triads. *J Exp Psychol Hum Percept Perform* 1990; **16**:415–428.

21. Egan JP, Hake HW. On the masking pattern of a simple auditory stimulus. *J Acoust Soc Am* 1950; **22**:622–630.

22. Wolff C, Schröger E. Human pre-attentive auditory change-detection with single, double, and triple deviations as revealed by mismatch negativity additivity. *Neurosci Lett* 2001; **311**:37–40.

23. Ilvonen T, Kujala T, Kozou H, Kiesilainen A, Salonen O, Alku P, *et al*. The processing of speech and non-speech sounds in aphasic patients as reflected by the mismatch negativity (MMN). *Neurosci Lett* 2004; **366**: 235–240.

24. Deouell LY, Bentin S, Soroker N. Electrophysiological evidence for an early (pre-attentive) information processing deficit in patients with right hemisphere damage and unilateral neglect. *Brain* 2000; **123** (Pt 2):353–365.

25. Gutschalk A, Micheyl C, Melcher JR, Rupp A, Scherg M, Oxenham AJ. Neuromagnetic correlates of streaming in human auditory cortex. *J Neurosci* 2005; **25**:5382–5388.