Disrupting human auditory change detection: Chopin is superior to white noise

SARI LEVÄNEN AND MIKKO SAMS
Brain Research Unit, Low Temperature Laboratory, Helsinki University of Technology, Espoo, Finland

Abstract
A deviant sound in a sequence of standard sounds elicits a neuromagnetic mismatch field (MMF) reflecting change detection based on the auditory sensory memory trace. To illuminate the nature of this trace, we investigated the effects of white noise and music maskers on the MMF. The stimuli were delivered to the participant’s right ear, and the maskers were delivered to the same or contralateral ear. Only maskers containing transients (music) presented to either ear abolished the MMF. In parallel, the ability to discriminate the deviants decreased dramatically, probably because of integration of transient features of the music to the neural representations of standards and deviants. As a result, the similarity of these representations prevents change detection. White noise affected the MMF amplitude only when presented to the same ear to which the stimuli were presented. All maskers decreased the M100 but not the M50 amplitude, suggesting that the neural generators behind these responses are functionally separate.

Descriptors: Evoked potentials, Magnetoencephalography, Masking, Memory trace, Mismatch response, Temporal integration

Infrequent deviant sounds, occurring randomly among repetitive standard stimuli, elicit a response called the mismatch negativity (MMN; Näätänen, Gaillard, & Mäntysalo, 1978) in electroencephalographic recordings and the mismatch field (MMF; Hari et al., 1984) in magnetoencephalographic (MEG) recordings. This response, suggested to reflect change detection in the auditory system, peaks 100–300 ms from stimulus onset, usually partly overlapping the 100-ms onset response. Mismatch responses (MMRs) to changes in many stimulus features have been recorded (for reviews, see Hari, 1990; Näätänen, 1990; Sams, 1991).

MMRs are not elicited by infrequent deviants presented without intervening standards (Lounasmaa, Hari, Joutsiniemi, & Hämäläinen, 1989; Näätänen, Pääviläinen, Alho, Reinkainen, & Sams, 1989; Sams et al., 1985). This result supports the hypothesis that the MMF reflects the automatic comparison of the neural memory trace coding the physical features of the standard stimulus with the input caused by the deviant stimulus (Näätänen et al., 1978). Support for the memory trace hypothesis is also provided by studies where standard and deviant stimuli are presented at different interstimulus intervals (ISIs). For example, Sams, Hari, Rif, and Knuutila (1993) presented tones at ISIs ranging from 0.75 to 12 s.

A significant MMF was elicited at ISIs up to 9 s. When individuals were asked to judge the similarity of two successive tones, they performed above chance level at ISIs $\leq$ 9 s. In an auditory backward masking experiment by Winkler, Reinkainen, and Näätänen (1993), the MMR amplitude correlated well with performance in a recognition task, also suggesting that the MMR reflects traces in auditory sensory memory. Behavioral estimates of the duration of auditory sensory memory range from a few seconds to 20 s (Crowder & Morton, 1969; Watkins & Todd, 1980).

MMRs are specific to the changed auditory feature (Levänen, Ahonen, Hari, McEvoy, & Sams, 1996; Levänen, Hari, McEvoy, & Sams, 1993; Schröger, 1995), suggesting that single features of the standard stimuli leave functionally separate and simultaneous traces in the human auditory cortex. Our recent results (Levänen et al., 1996) also show that deviants activate feature-specific MMF components bilaterally in the supratemporal cortex and a right-hemisphere parietal component that probably reflects more global auditory change detection. According to MEG recordings, the supratemporal MMF is generated slightly anterior to the source of the magnetic 100-ms response, M100 (Hari et al., 1984; Sams et al., 1985).

The 100-ms response, evoked by the onset of any abrupt change in the auditory environment (Hari et al., 1987), may also reflect the existence of a sensory memory trace (Lü, Williamson, & Kaufman, 1992a; Näätänen & Picton, 1987). MEG studies with varying ISIs (Hari, Kaila, Katila, Tuomisto, & Varpula, 1982; Hari et al., 1987; Lü, Williamson, & Kaufman, 1992b; Mäkelä et al., 1993; Sams et al., 1993) have shown that the amplitude of the auditory M100 increases with ISIs up to 6–16 s, without accompanying changes in the perceived quality of the stimulus (Hari et al., 1982). The difference between response amplitudes at short and long ISIs reflects
the existing level of neural activation caused by the previous stimulus. An identical stimulus occurring during the lifetime of the existing trace need only partially reactivate the system (Picton, Cambell, Baribeau-Brown, & Proulx, 1978).

Ongoing auditory stimulation is known to have marked effects on both electric and magnetic 100-ms responses evoked by a probe stimuli (Hari & Mäkelä, 1988; Woods, Hillyard, & Hansen, 1984). The more frequency and amplitude transients the masker has, the more effective it is in damping the 100-ms response. The neural activity underlying the response, although stimulus specific, probably becomes refractory when frequent transients are available. Hari and Mäkelä (1988) also found that the amplitude decreases of M100 in different masking conditions were not accompanied by systematic changes in the psychoacoustic thresholds, supporting the suggestion that this response is not directly linked to stimulus perception.

The M100 to onset but not to offset of an auditory stimulus is typically preceded by a 50-ms response (M50), suggesting that the two responses are not causally linked (Hari et al., 1987). M50 is functionally distinct from both the earlier 30-ms and the M100 deflection (Hari et al., 1987; Mäkelä, Hämmäläinen, Hari, & McEvoy, 1994; McEvoy, Mäkelä, Hämmäläinen, & Hari, 1994) and has a separate neural generator in the supratemporal auditory cortex (McEvoy et al., 1994).

To clarify the nature of the assumed memory trace underlying the MMF, we examined the effects of different simultaneous maskers on the MMF amplitude and latency and on the discriminability of the deviants from the standards. The effects of maskers on the MMF were compared with those on the M100 and M50 responses to investigate possible differences in their neural generation mechanisms.

Methods

The participants were eight right-handed members of laboratory staff (28–49 years of age; five women, three men) without histories of hearing loss or neurological disease.

The stimuli were 50-ms tones of 1 kHz (standards, probability = .85) and 1.1 kHz (deviants, probability = .15) presented with an ISI of 0.6 s. The same stimulus sequence was delivered to the right ear in three different conditions: (a) in a control condition without a masker and (b) when continuous white noise or (c) music (Chopin: Piano Concerto No. 1, Op. 11 from the beginning) was presented to the same or to the contralateral (left) ear. In all conditions, participants read a self-chosen book to prevent them from voluntarily attending to the auditory stimuli. The sounds were conducted to the ears through plastic tubes and earpieces. Sound intensity was 75 dB SPL for the stimuli (measured from a continuous tone), 82 dB SPL for the noise masker, and on average 80 dB SPL for the music masker.

Neuromagnetic signals were recorded with the Neuromag 1227TM system (Ahonen et al., 1993) containing a helmet-shaped array of 122 planar first-order SQUID (Superconducting Quantum Interference Device) magnetometers. Each sensor unit contains a pair of gradiometers that measure the two orthogonal tangential derivatives of the magnetic field component normal to the helmet surface at the sensor location. These gradiometers detect the largest signal just above a local brain source, where the field gradient has its maximum. The two independent derivatives measured at each site also give the direction of the source current.

The participant was seated in a magnetically shielded room with the head tightly pressed against the bottom surface of the dewar. Special care was taken to ensure that the head was in the middle of the helmet, and the participant was instructed not to move. Location of the head with respect to the sensors was determined by measuring the magnetic fields produced by small currents delivered to three coils attached to the scalp. Location of the coils with respect to the preauricular points and the nasion was measured with a three-dimensional digitizer (Isotrac 351002, Polhemus Navigation Sciences, Colchester, VT, USA), thereby determining the head coordinate system. The x-axis of the coordinate system points towards the right preauricular point, the y-axis points towards nasion, and the z-axis points towards vertex. The origin of the coordinate system, although showing some interindividual variability, is typically slightly anterior to the conjunction point of pons and medulla in the midline.

The recording bandpass was 0.03–100 Hz (3 dB points; high-pass roll-off = 35 dB/decade, low-pass roll-off > 80 dB/decade) with a sampling rate of 0.4 kHz. The measured responses were averaged on line. The analysis period was 600 ms, including a 145-ms prestimulus period that furnished a baseline for the measurement of peak amplitudes. The vertical electrooculogram was recorded simultaneously, and epochs contaminated by eye movements or blinks were rejected. A minimum of 100 responses were averaged for the deviants in each condition. After digital low-pass filtering at 40 Hz, peak amplitudes and latencies of MMFs to deviant stimuli and of M100 and M50 responses to standards were measured from those channels over the contralateral and ipsilateral auditory cortices showing the largest deflections. MMFs were examined from difference waveforms obtained by subtracting responses to standards from those to deviants.

To identify neural sources of the evoked responses, an equivalent current dipole (ECD) best describing the measured field pattern at a given latency was found by a least-squares search. ECDs for MMF and M100 responses were identified from the data measured in the control condition. Unfortunately, the signal-to-noise ratio was often too low to allow reliable estimation of ECDs for the small M50 responses. After digital low-pass filtering at 40 Hz, single ECDs were found separately for the left- and right-hemisphere data using a subset of channels over each hemisphere. The goodness of fit (g; Kaukoranta, Hämmäläinen, Sarvas, & Hari, 1986) of the model was also calculated. ECDs explaining more than 70% of the field variance in each hemisphere during the response peak were accepted for further analysis. In the left hemisphere, one ECD usually described the measured field pattern for the deviants adequately. When one ECD was insufficient for explaining the measured field pattern for the deviants in the right hemisphere, MMF was divided into two time periods, during each of which a separate ECD was found. Then the analysis period was extended to the entire MMF duration and the two dipoles were used to describe the data in the right hemisphere. The criterion for including two right-hemisphere dipoles rather than one was based on clearly separable components peaking on different channels at different latencies, on the improved explanation during the later part of the measured response, and on the improved g value at that latency (cf. Levänen et al., 1996). A time-varying multipole model was computed in which the locations and orientations of the previously found ECDs were kept fixed while the strengths were allowed to change.

The ability to detect the deviant stimuli in the different masking conditions was investigated in 6 participants (30–50 years of age; 4 women, 2 men), 3 of whom participated in the MEG experiments. Behavioral data were obtained in a separate session. Participants were instructed to lift their right index finger as soon as possible after hearing the deviant and not to respond to the stan-
Figure 1. MMFs of all participants during all conditions, shown from the channel with the maximum response in each hemisphere. The bottom waveforms are the grand averages across all participants.

Results

Figure 1 shows maximum MMF signals (difference waveforms: responses to standards subtracted from those to deviants) in each hemisphere of the 8 participants during all conditions. The grand-average waveforms across participants are shown at the bottom of the figure.

In the control condition, clear MMFs were elicited in almost all participants in both hemispheres. In some participants, the MMF was larger over the contralateral (left) hemisphere, and in others it was larger over the ipsilateral (right) hemisphere. Music presented to either ear abolished MMFs bilaterally in all participants; only the data of S1 show any sign of an MMF. Continuous white noise presented to the left ear did not influence MMFs. However, white noise to the right ear somewhat diminished the MMF amplitudes over both hemispheres. Mean ($\pm$ SEM) peak amplitudes for MMFs in each condition are shown in Figure 2.

Because the MMF was totally abolished in the left (contralateral) hemisphere of all participants during right-ear music masking, this condition was excluded from the ANOVA. MMF peak amplitudes and latencies in the control, left-ear music masking, and both white-noise masking conditions were compared using two-way ANOVAs (Condition $\times$ Hemisphere). For the MMF peak amplitude, the main effect of condition was statistically significant, $F(3,21) = 46.4, p < .001, \epsilon = 0.83$, but that of hemisphere was not. Post hoc comparisons showed that the MMF amplitude ($M \pm SEM$; contra- and ipsilateral responses pooled: $3.9 \pm 3.9$ fT/cm) during left-ear music masking was reduced to 8.8% of its unmasked value in the control condition (mean difference: 40.4 fT/cm). Left-ear white noise reduced the MMF amplitude to 89% and right-ear white noise reduced it to 74% of its unmasked value (mean differences: 35.5 and 28.9 fT/cm, respectively). White noise delivered to the right ear significantly diminished the MMF amplitudes.
Disrupting auditory change detection

M100 was preceded by a small M50 response of opposite polarity (Figure 3). The amplitude of M50 was not affected by the different maskers, as is evident from Figure 4, where the grand-average responses (8 participants) and the responses of two participants with the most prominent M50 to standards are shown on a more sensitive scale. (In the following discussion, M50 always refers to responses to the standard stimuli.)

None of the maskers had a significant effect on the M50 peak amplitude; however, the peak latency was affected, $F(4,28) = 5.5$, $p < .01$, $e = 0.52$. Post hoc comparisons revealed that the M50 peak latency was prolonged by about 10 ms bilaterally when music or white noise was delivered to the right ear. The contralateral M50 was significantly earlier (8.0–10.0 ms), $F(1,7) = 16.4$, $p < .005$, and larger (7.5–13.5 $\text{ft} \cdot \text{cm}$), $F(1,7) = 11.7$, $p < .01$, than the ipsilateral response in all conditions. For the M50 amplitude, condition produced significant interactions with hemisphere, $F(4,28) = 3.4$, $p < .05$, $e = 0.67$: the amplitude difference between the contra- and ipsilateral responses was larger during music masking to either ear than in other conditions.

ECDs for M50 and M100 responses were identified from the data measured in the control condition. M100 locations in each hemisphere agreed with activation of the supratemporal auditory cortex (Hari, 1990). In the contralateral (left) hemisphere, the ECD location for M50 was significantly anterior (0.9 cm), $t(6) = 4.0$, $p < .01$, to the source of M100. The ECD locations did not differ in the other directions. In the ipsilateral (right) hemisphere, in 4 of the 8 participants, two ECDs were needed to explain the measured field patterns in response to the deviant stimuli. The ECD location for the first MMF component was significantly anterior (0.9 cm), $t(4) = 3.5$, $p < .05$, to the source of M100; the ECD locations did not differ in other directions. The second MMF

<table>
<thead>
<tr>
<th>Control</th>
<th>Music LE</th>
<th>Music RE</th>
<th>Noise LE</th>
<th>Noise RE</th>
</tr>
</thead>
<tbody>
<tr>
<td>LH</td>
<td>RH</td>
<td>LH</td>
<td>RH</td>
<td>LH</td>
</tr>
<tr>
<td>S1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

![Image](image_url)

**Figure 3.** M100s of all participants during all conditions, shown from the channel with the maximum response in each hemisphere. The bottom waveforms are the grand-averages across all participants.
Component was superior to the corresponding source of M100 (mean difference: 0.5 cm) and to the first MMF component (mean difference: 0.7 cm), but the differences were not significant. Figure 5 shows contra- and ipsilateral M100 and MMF sources in the control condition for 1 participant superimposed on his magnetic resonance images. The sources for both M100 and the first MMF are within the sylvian fissure. The second right-hemisphere MMF source is in the inferior parietal cortex.

The behavioral data indicated that for all participants, the detection of deviants was most difficult during music masking. Most participants reported that even when they could hear the stimuli, it was very difficult to discriminate between the deviants and the standards. The mean (±SEM) discrimination percentages in all conditions are shown in Figure 6.

For the discrimination accuracy, the main effect of condition was significant, \( F(4,20) = 51.5, p < .001, e = 0.52 \), but that of the masked ear was not. Post hoc comparisons showed that accuracy was significantly impaired during both left- and right-ear music masking (\( M ± SEM = 39.5% ± 6.8% \) and 23.3% ± 7.5%) when compared with the control condition (92.8% ± 2.8%) and with both left- and right-ear white-noise masking (89.2% ± 3.5% and 86.8% ± 4.1%, respectively).

Because the participants knew the relative amount of the deviants (15% of the 350 stimuli) from the previous trials, they could have reacted by chance to 7.8% of the deviants. In all but the right-ear music-masking condition, the discrimination accuracy was well above the chance level, control: \( t(5) = 31.8, p < .001 \); left-ear music: \( t(5) = 4.7, p < .001 \); right-ear music: \( t(5) = 2.1, ns \); left-ear white noise: \( t(5) = 23.3, p < .001 \); right-ear white noise: \( t(5) = 19.5, p < .001 \).

Discussion

The MMF was practically abolished during music masking containing transients. When such a masker was presented to the same ear as the stimuli, the deviants did not elicit any detectable MMF. This result had a parallel in the behavioral experiment; the participants could not detect the deviants above the chance level. When the music masker was presented to the contralateral ear, the deviants elicited a small MMF in some participants. In this condition, the detection of deviants was very difficult but the participants were able to perform the task at above the chance level. These results support the suggestion that MMR reflects detection of change in an ongoing monotonous stimulus sequence; in the presence of continuously changing music, there is no stable background against which the change would occur.

Cowan (1984) suggested that sensory memory has two phases: a short, literal phase of 200–300 ms and a longer phase that can last up to 20 s. MMRs most probably reflect the longer phase (Sams et al., 1993), which, however, is dependent upon the earlier literal phase (Ritter, Deacon, Gomes, Javitt, & Vaughan, 1995). The short phase is experienced as a sensation that persists for 200–300 ms regardless of stimulus duration and integrates sensory information over that period. The persistence of sensation both feeds the temporal integration process that determines a sound’s perceptual qualities, such as loudness, and prolongs an information extraction process that improves discriminability (Cowan, 1987). The short store can retain information about a brief, rapid sequence of stimuli, but only holistically as an integrated event, without identifying each component. Loveless, Hari, Hämäläinen, and Tiitinen (1989) suggested that the temporal integration process is nonspecific in that it depends on the time since the last stimulus, without regard to the ear of stimulation or to the similarity of the two stimuli. It has also been suggested that the short phase is susceptible to masking: transients falling within this time window alter the temporal integration process (Gerken, 1993) and thus might affect the feature analysis and the perceived quality of the test stimuli.

We propose that during music masking, standard stimuli are integrated with the constantly changing features of the masker into integrated events. Similar integration, of course, affects the neural representation of the deviants. Therefore, the neural representations of the standards and deviants are so similar that they do not allow change detection. This similarity is reflected in the absent or strongly reduced MMF and in the impaired discrimination performance. The noise level of auditory sensory memory could be described as continuation of feature analysis during which feature integration occurs. Therefore, in parallel to changes in the short-duration representations, the feature analysis is degraded because of masking, which is reflected in the diminished M100 amplitude.

Another possible way to explain our results is to assume that the transients in our music masker act as deviants and therefore would result in very frequent mismatch processes. As a result, the MMF generator might become refractory. However, it has recently been shown (for a review, see Ritter et al., 1995) that if two different deviants are presented successively, the second does not elicit a smaller MMR than does the first, which suggests that it is difficult to make the MMF generator refractory. Of course, in the present experiment the assumed mismatch processes would have followed each other much more rapidly.

When Schröger (1994) presented the same stimulus sequence at different intensity levels (10–70 dB), the MMN amplitude to frequency deviants was not significantly influenced by stimulus intensities above the 25-dB threshold necessary for the MMN elicitation. Therefore, the change detection mechanism seems to be largely independent of stimulus energy. However, when the signal-to-noise ratio was decreased in the present experiment by adding noise to the same ear as was receiving the stimuli, the MMF amplitude was diminished. The reason for this result probably is (peripheral) blurring of frequency coding in the presence of a white-noise masker.
One MMF source was enough to explain the MMF field pattern over the left (contralateral) hemisphere in all 8 participants. However, two temporally overlapping but spatially separate sources were needed to explain the right-hemisphere MMFs in 4 of the participants. These results confirm our previous findings (Levänen et al., 1996), on the basis of which we suggested stronger involvement of the right than the left hemisphere in memory-based change detection and that the bilateral MMF components originating in the supratemporal cortex are feature specific, whereas the right-hemisphere parietal component reflects more global auditory change detection. These parietal-lobe sources might reflect activation of polysensory parietal areas (Irvine & Phillips, 1982), which receive convergent acoustic, visual, and somatic input and are thought to be less affected by the physical attributes of the stimuli than by their significance. The observation that MMN is larger in amplitude in the right than in the left hemisphere, irrespective of the ear of stimulation (Giard, Perrin, & Pernier, 1991; Giard, Perrin, Pernier, & Bouchet, 1990; Paavilainen, Alho, Reinikainen, Sams, & Näätänen, 1991; Scherg, Vajsar, & Picton, 1989) also suggests hemispheric differences in MMN generation.
Using similar stimulus intensities as in the present experiment, Hari and Mäkelä (1988) found that continuous noise presented to the same ear as received the stimuli nearly abolished the M100 response to and the sensation of brief tones. This masking effect was suggested to derive mainly from the periphery. This stronger masking effect than that found in the present experiment probably can be explained by their use of shorter and lower frequency stimuli. In Hari and Mäkelä’s (1988) experiment, the contralateral masker had no effect on M100. In the present study, however, continuous white noise presented to either ear decreased the M100 amplitude. Therefore, the masking effect cannot be explained solely by the interaction of masker and signal in the cochlea. We suggest that the white-noise masker activated both ipsi- and contralateral auditory cortices, thereby decreasing their excitability to other simultaneous auditory stimulation.

In the present study, none of the maskers had a significant effect on the M50 amplitude; only its latency was slightly prolonged when maskers were delivered to the same ear as the stimuli. The absence of masking effects on the M50 amplitude provides further evidence that the 50-ms response is related to different aspects of auditory processing than is the 100-ms response (Hari et al., 1987; Mäkelä et al., 1994; McEvoy et al., 1994). The early response latency and the lack of contralateral masking effects in the present experiment support the suggestion that M50 reflects sensory input to the auditory cortex (Hari et al., 1987), which is relatively little affected by the maskers.

In the present study, the effects of music and white-noise maskers were different for the MMF, M100, and M50 responses, suggesting that the neural generators behind these responses are, in addition to being spatially distinct, functionally separate. On the basis of the present and earlier results, we suggest that sensory input to the auditory cortex, as reflected in M50, is not particularly affected by different maskers, but the attention-triggering mechanism reflected in M100 and especially memory-based change detection reflected in MMF are strongly affected by transient features of the masker.

REFERENCES


Disrupting auditory change detection


(Received January 5, 1996; Accepted July 19, 1996)
This document is a scanned copy of a printed document. No warranty is given about
the accuracy of the copy. Users should refer to the original published version of the
material.