Psychophysiology, 44 (2007), 671–679. Blackwell Publishing Inc. Printed in the USA.

DOI: 10.1111/j.1469-8986.2007.00554.x

Representation of harmonic frequencies in auditory memory: A mismatch negativity study

ELANA ZION-GOLUMBIC,[a] LEON Y. DEOUELL,[b,c] DOUGLAS H. WHALEN,[d,e] AND SHLOMO BENTIN[b,c,d]

[a]Department of Cognitive Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel
[b]Department of Psychology, The Hebrew University of Jerusalem, Jerusalem, Israel
[c]Interdisciplinary Center for Neural Computation, The Hebrew University of Jerusalem, Jerusalem, Israel
[d]Haskins Laboratories, New Haven, Connecticut, USA
[e]National Science Foundation, Arlington, Virginia, USA

Abstract

Most natural sounds are composed of a mixture of frequencies, which activate separate neurons in the tonotopic auditory cortex. Nevertheless, we perceive this mixture as an integrated sound with unique acoustic properties. We used the Mismatch Negativity (MMN), a marker of auditory change detection, to determine whether individual harmonics are represented in sensory memory. The MMN elicited by duration and pitch deviations were compared for harmonic and pure tones. Controlled for acoustic differences between standards and deviants and their relative probabilities, the MMN was larger for harmonic than pure tones for duration but not for pitch deviance. Because the magnitude of the MMN reflects the number of concurrent changes in the acoustic input relative to a preexistent acoustic representation, these results suggest that duration is represented and compared separately for individual frequencies, whereas pitch comparison occurs after integration.

Descriptors: MMN, Auditory, Harmonic, Pitch, Duration

Most periodic sounds generated by humans, animals, musical instruments, and other ecological sources are spectrally rich, made up of a fundamental frequency and integer-multiplications of the fundamental (“harmonics”). Although each frequency activates different neurons in the tonotopically organized auditory cortex, a harmonic tone is usually characterized by one pitch, which in most cases corresponds to its fundamental frequency (Moore & Moore, 2003). This unitary perception implies integration of the representations of individual frequencies into a single representation of the complex sound. The question posed in the present study was whether individual harmonics comprising a complex sound are independently represented in sensory memory and whether auditory dimensions such as pitch and duration are extracted from each of these harmonics prior to their integration. Alternatively, only the integrated complex sound might be represented in the auditory sensory memory trace and auditory features are extracted following integration. To address this question we used a Mismatch Negativity (MMN) paradigm (Näätänen, 1990; Näätänen, Gaillard, & Mäntysalo, 1978).

In a typical MMN paradigm, a deviant auditory stimulus is infrequently interspersed in a sequence of another, frequently repeated, “standard” auditory stimulus (oddball condition). The deviance can be obtained by manipulating basic auditory features such as pitch, intensity, duration, interstimulus interval, and location, as well as more complex dimensions of deviation, such as phonetic information and temporal order. The MMN traditionally results from the subtraction of the ERP elicited by the standard stimulus from that elicited by the deviant stimulus. This waveform presents a negative deflection, the MMN, peaking roughly between 100 and 250 ms after deviance onset. The MMN magnitude is determined by the size and the dimension of the deviance: the larger the change, the earlier and bigger the MMN (e.g., Deouell, Parnes, Pickard, & Knight, 2006; Tiitinen, May, Reinikainen, & Näätänen, 1994). Importantly, studies have pointed to a major source of the MMN in the associative auditory cortex, suggesting that the acoustic change is already detected in the unimodal cortex and is affected relatively little by top-down processes (Sussman, Winkler, & Schröger, 2003) or, arguably, by attention (Näätänen, 1991; Näätänen, Päävilainen, Tiitinen, Jiang, & Alho, 1993; Restuccia, Della Marca, Marra, Rubino, & Valeriani, 2005; but see Woldorff, Hackley & Hillyard, 1991; Woldorff, Hillyard, Gallen, Hampson, & Bloom, 1998; for reviews, see Näätänen, 2003; Näätänen & Escera, 2000).

According to a prevalent account, the MMN manifests the discrepancy between the auditory percept formed by an incoming deviant stimulus and a memory template established by the repetitive, standard stimulus. Several lines of evidence confirm that deviance along different auditory dimensions (e.g., duration, pitch, location) generate independent MMNs (e.g., Deacon, Nousak, Pilotti, Ritter, & Yang 1998; Tervaniemi et al., 1999),

---

Address reprint requests to: Leon Y. Deouell, Department of Psychology, Hebrew University of Jerusalem, Jerusalem 91905, Israel. E-mail: Leon.Deouell@huji.ac.il

which add linearly (Schröger, 1995). This suggests that discrete features are extracted from the incoming auditory stimulus and independently compared with the respective features embedded in parallel in the memory template (Ritter, Deacon, Gomes, Javitt, & Vaughan, 1995).

Consequently, if the harmonics of a complex tone are represented separately in the sensory memory template, then multiple traces should be registered for each auditory dimension (one for each harmonic). If this is the case, then the MMN should be larger for harmonic than for pure tones, because for harmonic tones the MMN would reflect a composite effect of multiple deviance detections. If the alternative option is true, and the comparison is based on a single merged sound representation, then the MMN elicited by harmonic and pure tones should not differ. Using pitch as the dimension of deviance, Tervaniemi and her colleagues (Tervaniemi, Ilvonen, et al., 2000; Tervaniemi, Schröger, Saher, & Näätänen, 2000) found that the amplitude of the MMN elicited by harmonic tones was higher compared to pure tones, was increased as additional harmonics were added, and its peak latency was earlier (albeit the difference in latency was not affected by the number of harmonics added). These authors also found that behavioral discrimination performance was better for harmonic than for pure tones. Based on these data and the above described logic, one might conclude that the harmonics are represented in the sensory memory template independently and that they are extracted from the acoustic input and compared to this template prior to their perceptual integration. However, this conclusion might be premature given a fundamental caveat about the method of estimating the MMN.

The basic rationale underlying the "Classic MMN" method, as was described above, is that the subtraction of the standard ERP from the deviant ERP nullifies the effect of the evoked sensory activity elicited by all stimuli regardless of the context (e.g., the N1 potential). Moreover, the assumption is that the difference waveform manifests only the neural activity resulting from the mismatch between the deviant input and the memory template established by the repetitive standard. The validity of this rationale, however, depends on the veracity of two assumptions: One is that the evoked sensory activities elicited by the standard and by the deviant in the oddball sequence are very similar, and the second is that this evoked sensory response and the response to the mismatch are additive. Whereas evidence for or against additivity is not available, there are good reasons to believe that the first assumption is probably not always true (Jacobsen & Schröger 2001). First, the standard and deviant stimuli are, by definition, acoustically different. Second, stimulus repetition causes adaptation of neural response and consequently affects the evoked response in the time range of the MMN, resulting in a reduction in the amplitude of the exogenous N1 component (Budd, Barry, Gordon, Rennie, & Michie, 1998; Haenschel, Vernon, Dwivedi, Gruzelier, & Baldeweg, 2005; Picton, Hillyard, & Galambos, 1976). This process is likely to affect the response to the highly repetitive standards more than to the infrequent deviants. Therefore, the difference waveform resulting from subtracting the standard from the deviant should manifest sensory differences between these stimuli in addition to the detection of the mismatch between them. Indeed, some authors suggest that the MMN is no more than a result of differences in the N1 component between the deviant and standard stimuli rather than a separate component in itself (Jääskeläinen et al., 2004; May et al., 1999; Ulanovsky, Las, & Nelken, 2003; but see Näätänen, Jacobsen, & Winkler, 2005, for a rebuttal of this argument). At a minimum, it is quite possible that the magnitude of the deviant–standard difference is determined by both N1 adaptation and an independent MMN.

The "contamination" of the MMN with sensory-based activity differences between the standard and the deviant is particularly conspicuous when the onset of the deviance and the onset of the stimuli coincide, as is the case with pitch-deviance MMN. Hence, the higher deviance-related negativity found by Tervaniemi, Ilvonen, et al. (2000) for harmonic tones relative to pure tones might have reflected a difference in N1. In an attempt to circumvent this problem, we explored whether the MMN elicited by harmonic tones is larger than that elicited by pure tones, while disentangling the MMN and the putative adaptation effects, using the method developed by Jacobsen and Schröger (2001, 2003) for assessing MMN. In Jacobsen and Schröger's method, the response to the ERP elicited by the deviant stimulus is not compared to the ERP elicited by the standard in the same oddball sequence, but to the ERP elicited by a control stimulus that is physically identical to the deviant but presented in a separate "control" block, where it is presented with the same probability as in the oddball sequence, but interspersed among several equiprobable stimuli. The rationale behind this method is that the sensory-evoked response to the control and the deviant stimuli is identical because they are physically the same. In addition, the control and the deviant stimuli are equally adapted because in both blocks (oddball and control) they are repeated with equal relative probabilities. Thus, the only difference between the two stimuli is the context within which they are presented; whereas in the oddball block the infrequently occurring stimulus is deviant, in the control block all stimuli have the same probability. Consequently, a purely memory-based MMN can be obtained by subtracting the ERP elicited by the deviant in the oddball block from the identical control stimulus presented in the "control" block. We will refer to the MMN calculated using this method as the "Controlled MMN," and compare the MMN elicited by harmonic and pure tones calculated by the classic and controlled methods. Additional dissociation between the N1 effects and MMN effects resulting from deviance from a memory template can be obtained by separating the onset of the deviance from the onset of the deviant stimulus. Such separation can be achieved using stimulus duration as the dimension of deviance. In this case, the onset of the deviance is the offset of the shorter stimulus.

To summarize, the goal of the study was to investigate the representation of complex tones in sensory memory. We reasoned that if there are separate representations of the complex tone's harmonics in the sensory memory template, then a greater memory-based MMN should be seen for harmonic than for pure tones. This should be true for both duration and pitch deviants and regardless of the method used to derive the MMN. In contrast, if the previously observed effect of complexity of the sound on the MMN is due to an N1-associated effect, then this effect should be absent when the MMN is calculated using the controlled method and should be smaller for duration than for pitch even in the classic design.

Method

Participants

The participants were 29 undergraduate students (12 women) who were paid or received course credit for taking part in this

Table 1. *List of Tones Used as Stimuli*

Stimulus	Frequencies (Hz)[a]	Duration (ms)
Standard (S)	500 (+1000+1500)	150
Duration deviant (DD)	500 (+1000+1500)	100
Pitch deviant (PD)	550 (+1100+1650)	150
Control 1	328 (+656+984)	175
Control 2	364.5 (+729+1093.5)	50
Control 3	405 (+810+1215)	225
Control 4	450 (+900+1350)	75
Control 5	605 (+1210+1815)	200
Control 6	665.5 (+1331+1996.5)	25
Control 7	732 (+1464+2196)	250
Control 8	805 (+1610+2415)	125

[a] The pitch values added in parentheses are the values of the harmonics added in the harmonic set.

study. The median age was 25 (range 20–37). All participants reported normal hearing and normal or corrected-to-normal sight. One of the subjects did not have a discernable MMN in over five conditions and therefore was excluded from the analysis.

Stimuli

The standard pure tone stimulus (S) had a frequency of 500 Hz and duration of 150 ms. These were also the median values for frequency and duration in the entire set of stimuli. Pure tone deviant stimuli differed from the standard in only one dimension, either pitch or duration. The pitch deviant (PD) was 10% higher in frequency (550 Hz), and the duration deviant (DD) was 50 ms shorter (100 ms) than the standard stimulus. The pure-tone stimuli in the control condition included 10 pure tones with frequencies between 328 Hz and 805 Hz and duration from 25 ms to 225 ms including one tone that was identical to the pitch deviant described above (550 Hz/150 ms) and one to the duration deviant (500 Hz/100 ms; Table 1). With the exception of these two stimuli, the pitch and duration values of the control stimuli were paired in a pseudorandom fashion to maintain independence of the two dimensions (Table 1). The corresponding harmonic stimuli had identical fundamental frequencies as the pure tones, with the addition of the first and the second harmonics, with linearly decreasing amplitudes. All stimuli included 5 ms rise and fall times.

EEG Recording

The EEG analog signals were recorded continuously by 64 Ag-AgCl pin-type active electrodes mounted on an elastic cap (Electro-Cap International, Inc.) according to the extended 10–20 system (American Electroencephalographic Society, 1994), and from two additional electrodes placed at the right and left mastoids. Eye movements, as well as blinks, were monitored using bipolar horizontal and vertical EOG derivations via two pairs of electrodes, one pair attached to the external canthi and the other to the infraorbital and supraorbital regions of the right eye. Both EEG an EOG were sampled at 256 Hz using a Biosemi Active 2 system (www.biosemi.com). A 67-Hz low-pass filter was used during recording to avoid aliasing of high frequencies. The digitized EEG was saved and processed off-line.

Experimental Procedure

After the application of the electrode cap, the participants were seated in a reclining chair in a sound-attenuated and electrically shielded chamber. They watched a silent animated film on a computer screen positioned approximately 70 cm in front of them. While watching the movie, auditory stimuli were presented through earphones (Sennheiser model HD 570) at 70 dB SPL with stimulus onset asynchrony ranging randomly between 450 and 700 ms. Subjects were instructed to focus on the movie and ignore the sounds.

For each type of tone (pure/harmonic) there were two experimental blocks.

Control block. The control stimuli for the two deviants (PD and DD) and the eight additional control stimuli were presented 280 times each, in random order (the relative frequency of each stimulus in this block was thus equal to 10%).

Oddball block. The standard stimulus was presented 2240 times (80%) whereas the deviants PD and DD were presented 280 times each (hence the relative frequency of each of the two deviants was 10%, as in the control block). The order of presentation was pseudorandom so that a deviant could appear only after at least three consecutive standards.

The order of tone type presentation (pure/harmonic) was counterbalanced across subjects, but the control block was always presented before the oddball block.[1]

Data Analysis

The continuous EEG data were digitally referenced off-line to the tip of the nose and filtered with a Butterworth zero-phase high pass filter (0.5 Hz, 24 dB/octave) and a 50-Hz notch filter.

Independent Component Analysis (ICA) was carried out for each subject to isolate blink artifacts, and components that could be attributed to such artifacts based on their topography and time course were eliminated from the data (Jung et al. 2000). Additional artifact rejection was applied and epochs with EEG amplitude exceeding ± 150 μV at Fz, F3, F4, both mastoids, and the bipolar EOG channels were excluded (300 ms symmetrical around the artifact's occurrence time).

Data were segmented into 600-ms epochs starting 100 ms before stimulus onset and averaged separately for each type of stimulus: pure and harmonic standard stimuli, pure and harmonic duration, and pitch deviants presented in oddball block and their respective identical controls, which were presented in the control block. To approximate equal number of trials in each condition, in the oddball block only standards immediately preceding a deviant were included in the averaged bins.

After averaging, frequencies lower than 1 Hz and higher than 12 Hz were digitally filtered out (24 dB/octave). The ERP epochs were baseline corrected by subtracting the mean of the prestimulus activity of each ERP waveform.

For each of the four deviants, the classic MMNs were calculated by subtracting the ERPs elicited by the standard stimuli from the ERPs elicited by each of the deviant types presented within the same oddball block. The controlled MMNs were calculated by subtracting the ERPs elicited by the corresponding equiprobable control stimuli presented in the control blocks from the deviant stimuli presented in the oddball blocks.

---

[1] Because in the classic paradigm the deviant tone stands out, there is a risk that this specific stimulus might retain a special status if the control block is presented after the classic. Hence we presented the control blocks first.

ERP Quantification and Statistical Analysis

MMN peak amplitudes for each participant were determined at Fz by finding the maximum negativity in the difference waveform within a time window of 100–200 ms after *deviance* onset (100–200 ms after stimulus onset for pitch MMN, 200–300 ms after stimulus onset for duration MMN). Of the 28 subjects included in the analysis, 15 did not show a discernible MMN in one to three of the eight conditions (i.e., the difference waveform did not show negative voltage throughout the relevant time window). In these cases, a value of zero was assigned as the MMN amplitude. In these subjects, analysis of MMN peak latency was precluded. A three-way within-subject ANOVA was carried out for the amplitude of the MMN at electrode Fz with factors Tone Complexity (pure vs. harmonic), Deviance Type (pitch vs. duration), and Subtraction Method (classic vs. controlled).

Because in this study some of our subjects did not show a clear MMN in all conditions, in order to analyze latency affects we quantified peak latencies of the MMN only from the 13 subjects who had a clear MMN in all the conditions. On these latencies we performed a three-way within-subject ANOVA with the factors Tone Complexity (pure vs. complex), Deviance Type (pitch vs. duration), and Subtraction Method (classic vs. controlled).

Greenhouse-Geisser (G-G) correction was used where indicated; the G-G epsilon is reported along with the uncorrected degrees of freedom.

Results

The ERP waveforms for standards, deviants, and their control stimuli in all conditions are presented in Figure 1. These ERPs were used to calculate difference waves using both the classic method (deviant minus standard) and the controlled method (deviant minus control). Visual inspection of the grand average difference waves obtained in both classic and controlled procedures revealed, across participants, a clear MMN in all conditions peaking between 164 and 199 ms for pitch deviants and between 227 and 257 ms for duration deviants (Figure 2). Table 2 summarizes the mean amplitude and latency values of the MMNs for all conditions and for both methods. All MMNs were most negative at fronto-central electrodes with a polarity reversal at the mastoids. Topographical distributions of the MMNs in all conditions were similar and are presented in Figure 2.

The three-way within-subject ANOVA of the MMN amplitudes revealed a main effect for Tone Complexity, $F(1,26) = 5.055$, G-G = .557, $p < .05$, indicating that overall the MMN was larger for harmonic tones than for pure tones. None of the other main effects were significant: Deviance Type, $F(1,26) < 1.0$, Subtraction Method $F(1,26) < 1.0$. However, a significant interaction between Deviance Type and Subtraction Method was found, $F(1,26) = 10.43$, G-G = .361, $p < .005$. This interaction was further qualified by a significant three-way interaction between all factors, $F(1,26) = 4.838$, G-G = .358,

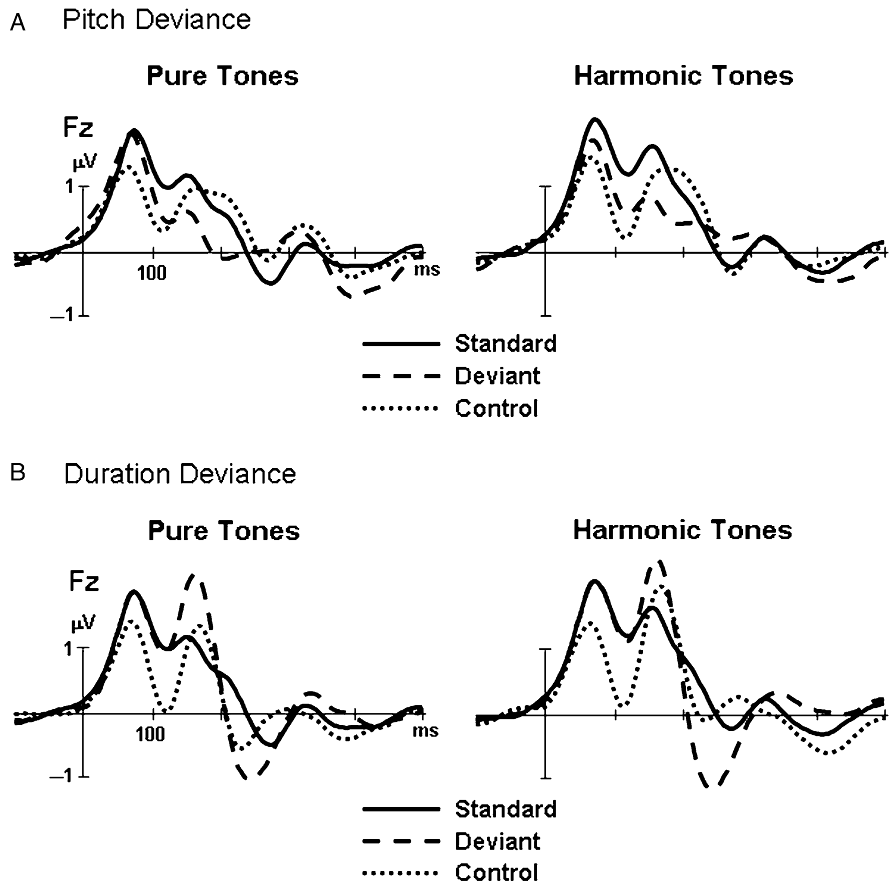

Figure 1. Grand average ERP waveforms elicited by pure and harmonic standard, deviant, and control stimuli in (A) the pitch deviance condition and (B) the duration deviance condition, evaluated at electrode Fz.

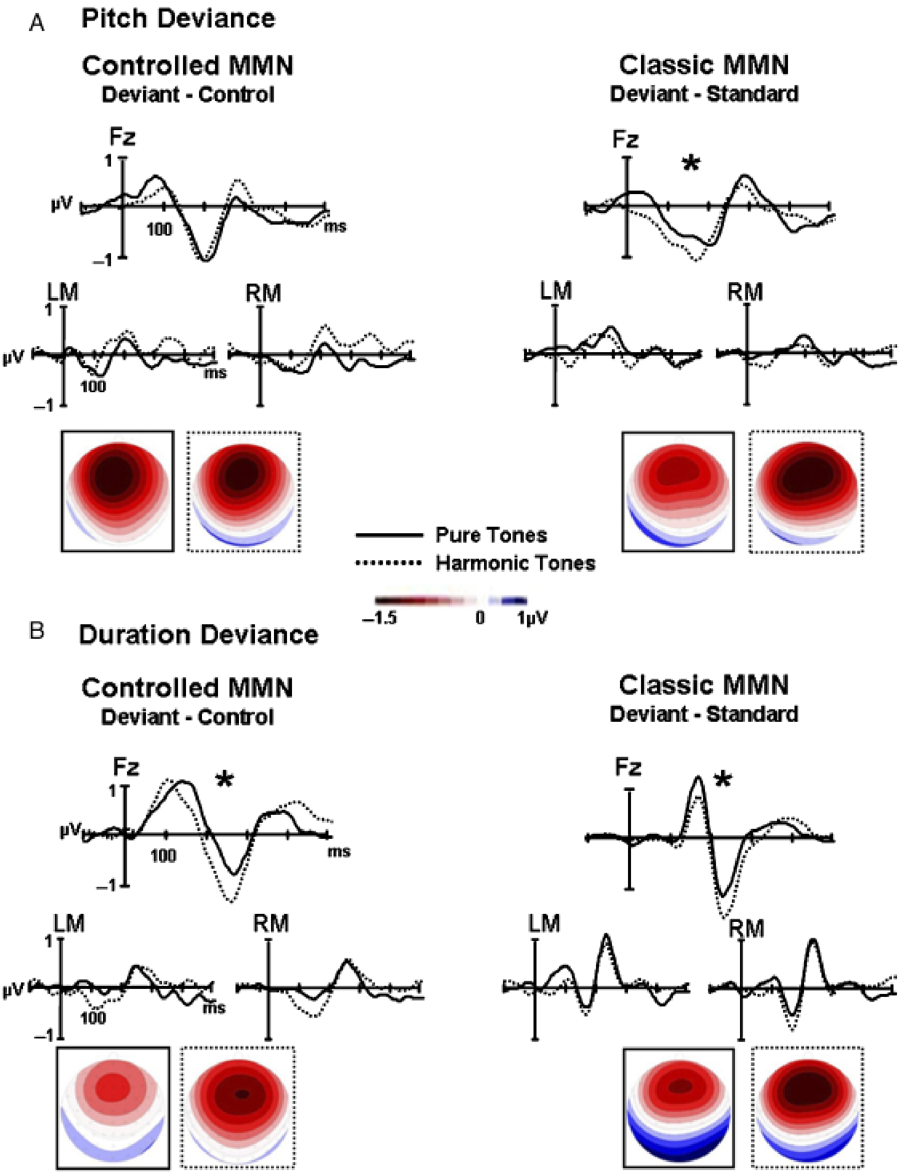

Figure 2. Difference waves derived using the classic (deviant minus standard) and control (deviant minus control) methods, for (A) pitch and (B) duration deviants, at electrode Fz and the mastoids. Clear MMNs are observed at Fz in all conditions with an inversion at the mastoids, and their topographical distributions are presented. The difference waves elicited by pure and harmonic tones in each calculation method and for each deviant type are compared, exhibiting the tone complexity effects on the MMN (significant effects are marked with a star). An ANOVA performed on the voltage at the mastoids at the MMN peak did not reveal any effect for tone complexity.

$p < .05$. As shown in Figure 2, whereas the amplitude differences between the MMN elicited by harmonic and pure tones exists for duration deviance regardless of method used, this difference was eliminated for pitch deviance when using the controlled method.

To assess the origin of the second order interaction a two-way ANOVA (Tone Complexity × Subtraction Method) was applied within each deviance type. Duration deviants showed a main effect for Tone Complexity, $F(1,26) = 5.303$, G-G = .535 $p < .05$, indicating that the MMN amplitudes were larger for harmonic than for pure tones, across subtraction methods. In addition classic MMNs were of a larger amplitude than controlled MMNs for both tone types, $F(1,26) = 6.365$, G-G = .581, $p < .05$. The interaction between these two factors was not significant, $F(1,26) = 1.877$, G-G = .264, $p > .1$.

In contrast, pitch deviance showed no main effect of tone complexity, $F(1,26) < 1$, or of subtraction method, $F(1,26) = 1.943$, G-G = .466, $p > .1$, but the interaction between the factors was significant, $F(1,26) = 5.070$, G-G = .299, $p < .05$, indicating that the complexity effect was different for pitch deviants when using the two methods. This interaction was followed up further by individual one-tailed t tests[2] between the MMN to pure and harmonic tones within each method. Although the MMN was

---

[2] The predicted direction was determined by comparison to Tervaniemi, Ilvonen, et al. (2000).

Table 2. *Grand Average MMN Peaks Derived Using Classic and Control Methods for All Conditions*

	Duration deviance			
	Classic MMN		Control MMN	
	Amplitude (μV)	Latency (ms)	Amplitude (μV)	Latency (ms)
Pure	– 1.106	226	– 0.746	257
Harmonic	– 1.5	230	– 1.25	250
	Pitch deviance			
	Classic MMN		Control MMN	
	Amplitude (μV)	Latency (ms)	Amplitude (μV)	Latency (ms)
Pure	– 0.72	195	– 1.013	199
Harmonic	– 1.014	164	– 0.986	187

significantly larger for harmonic tones than for pure tones when using the classic method, $t(27) = 1.720$, $p < .05$, this difference was no longer significant when using the controlled method, $t(27) = .821$, $p > .1$.

The three-way within-subject ANOVA of the MMN latencies revealed a main effect for deviance type, $F(1,12) = 131.47$, $p < .001$, which is expected due to the difference in deviance onset (for pitch it is with the tone onset whereas for duration it is at tone offset). In addition we found that the controlled MMN peaked later than classic MMN, $F(1,12) = 4.412$, $p < .001$. This is in line with the findings of Jacobsen and Schröger (2001, 2003), who introduced the controlled MMN. The MMN to pure tones tended to be later than for complex tones, but this effect was not quite significant, $F(1,12) = 4.41$, $p = .058$. There were no significant interactions.

Inspection of the difference waveforms (Figure 2) reveals unanticipated positive-going differences prior to the MMN latencies. First, a broad positivity starting as early as 30 ms and continuing up to $\sim$ 200 ms is seen when comparing deviants to their control condition, most evidently for the duration deviations. This type of difference is absent when deviants are compared to standards in the classic comparison. Inspection of the raw waveforms (Figure 1) reveals the source for this difference: The early waveform (up to $\sim$ 250 ms) elicited by repetitive standards is characterized by a typical broad positivity (cf. Haenschel et al., 2005). The same effect is seen in response to the duration deviants, which were indistinguishable from the standards up to 100 ms. In contrast, this positivity is absent (or much smaller) in the control condition, and therefore the subtraction of control from the deviant reveals a positive difference. The pitch deviant, which is very similar to the standard except for a slight pitch change, shows the initial part of positive deflection, which ends shortly, possibly when the deviance can be detected. Thus, the difference wave shows only a small and short-lived positivity. Most importantly for the present study, post hoc t tests showed no significant difference between harmonic and pure tones for this positivity. Second, a sharp positive peak is seen in the difference between 100 ms duration deviant and the 150 ms standard in the classic comparison. This difference reflects larger P2 in response to the deviant than the standard (Figure 2B). Although there is a trend for a smaller positivity for harmonic tones, this trend did not survive correction for post hoc multiple comparisons.

Discussion

Using the "classic" subtraction procedure, we found that pitch deviants elicited larger and earlier MMN when the stimuli were complex (harmonic) tones than when they were pure tones, consistent with the report of Tervaniemi, Ilvonen, et al. (2000). However, no such effect was found for pitch deviants if the MMN was calculated using the controlled subtraction paradigm. The cardinal difference between the two subtraction methods is that effects derived by the classic method may include a contribution due to N1 adaptation, whereas effects derived by the controlled method arguably reflect only memory-based MMN (Jacobsen & Schröger, 2001). Thus, it seems that the effect of added harmonics on the pitch MMN may be due to N1 adaptation effects. In contrast, duration deviance elicited larger MMNs for harmonic deviants than for pure tones regardless of subtraction method. This three-way interaction between deviance type, subtraction method, and stimulus complexity indicates that detection of deviance in pitch and duration is based on different forms of representation in auditory sensory memory.

The Analysis of Harmonic Tones in the Auditory System— Two Models

According to the "memory-based" comparison account for the MMN (Näätänen, 1990; Näätänen et al., 2005), a memory template is created for the frequently presented sound, and every subsequent sound is compared with this template to identify changes. The question addressed in this study was whether the different frequencies are preserved in the memory template or, alternatively, if the frequencies comprising the sound are integrated at an early stage so that it is represented in memory only as a singular sound object. We surmised that if the frequency components are encoded separately in memory, then each representation would be compared independently to the frequency components of the input (Figure 3A). Hence, the number of comparisons performed on each auditory input would be proportional to the number of components. Each of these comparisons would elicit an MMN, and, assuming additivity, this would result in larger overall MMN amplitude for harmonic than for pure tones. This is, in fact, what we found for duration deviants. Alternatively, if only integrated representations are compared rather than individual frequencies (Figure 3B), the resulting MMNs should not be influenced by the number of harmonics. The pattern of MMN elicited by pitch deviance seems to conform better to this model.

The present results suggest that pitch and duration comparisons are carried out by mechanisms that treat harmonic components in different manners. The pitch comparison is based on the combined stimulus, whereas the duration comparison is performed for each frequency channel. The notion that pitch traces reflect the emergent property (pitch) rather than the separately computed pitch of the spectral components is consistent with the MMN findings employing the missing fundamental phenomenon (Winkler et al., 1995; Winkler, Tervaniemi, & Naatanen, 1997). In those studies, the only common determinant among the standards was their virtual pitch, determined by their missing fundamental, whereas their spectral content varied. The elicitation of an MMN by a different pitch sound in this condition indicated that the relevant memory trace indeed stores the emergent pitch.

Consistent with the existence of integral pitch representation, simultaneous activation of individual frequencies in mammal

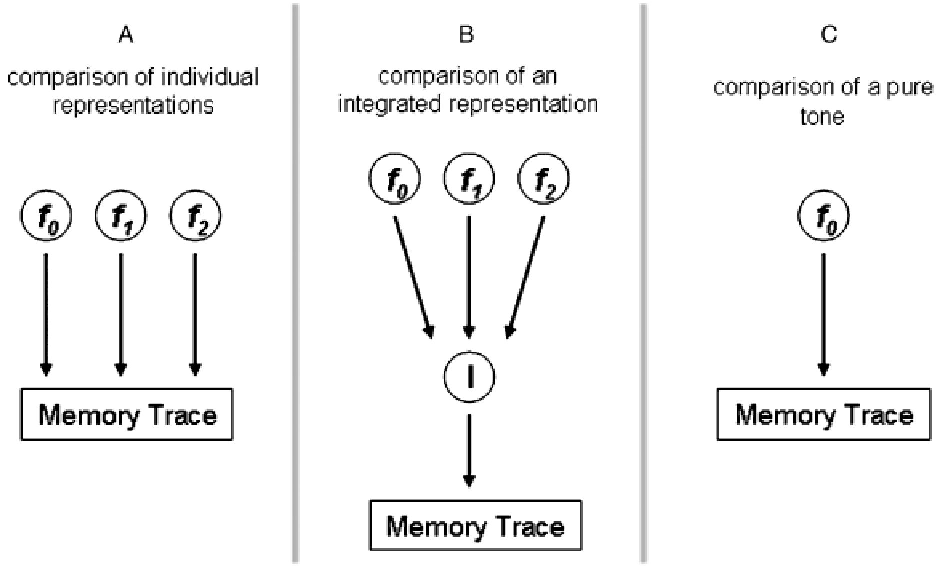

Figure 3. Models for mechanism dealing with complex sounds in the auditory system. A: Model A suggests individual representations for each frequency components, which are separately compared with the memory trace. B: The individual frequencies are integrated (I) prior to comparison with the memory trace. C: For pure tones only one comparison is needed regardless of mechanism.

auditory cortex facilitates a nonlinear ''combination-sensitive'' neural response (Kanwal, Fitzpatrick, & Suga, 1999; Rauschecker, Tian, & Hauser, 1995). The lateral areas of the monkey auditory cortex have been suggested to be organized in a hierarchical sequence in which neurons prefer increasingly complex stimuli. ''Combination sensitive'' neurons have been suggested as the neural basis for grouping harmonics within a complex sound and extracting the pitch regardless of whether the fundamental frequency is present (Medvedev, Chiao, & Kanwal, 2002) and may play an important role in communication. This hypothesis receives further support from recent imaging studies on humans revealing cortical pitch processing antero-lateral to primary auditory cortex (Penagos, Melcher, & Oxenham, 2004) and from the discovery of ''pitch-specific'' neurons in primate auditory cortex (Bendor & Wang, 2005).

Environmental sounds are harmonic by nature, and it makes ecological sense for our system to treat the pitch of the harmonic sound as the pitch of an integrated object. In contrast to pitch, our data suggest that the duration of individual frequencies is maintained and utilized for detection of change. Arguably, determining that different frequencies have different durations may be an important tool for auditory scene analysis.

Finally, this study joins a number of scalp-topography and source modeling findings suggesting that there is not one MMN generator but rather multiple generators depending on the dimension of deviance (Giard et al., 1995; Rosburg, 2003). A recent fMRI study (Molholm, Martinez, Ritter, Javitt, & Foxe, 2005) has pointed to distinct networks of auditory cortices that are activated in response to duration and pitch deviance. Because each dimension seems to rely on specific calculations to assess changes, it would not be surprising to find that the networks involved in deviance detection in different dimension have at least partially distinct neural implementation.

Before concluding, we consider the unexpected differences between deviant and controls that occurred prior to the MMN. Because these differences were not anticipated, we can only hypothesize on their origin post hoc. Similar positive differences were found comparing duration deviants with control stimuli by Jacobsen and Schröger (2003) and pitch deviants with control stimuli (Jacobsen & Schröger, 2001). Analyzing the structure of the earlier positivity, we suggest that it might reflect a modulation of a slow positive potential (50–250 ms), which has been associated with the formation of the echoic memory template and is correlated with the number of repetitions in a sequence (''repetition positivity'' [RP]; Haenschel et al., 2005). The RP was evident in our study for the repeating standards and for the duration deviants, which were identical to the standards up to 100 ms, and initially also for the pitch deviants, which were similar to the standards on many dimensions. Because the amplitude of the RP is a function of the number of consecutive repeating stimuli, it should be expected to be absent (or much smaller) in the control condition, in which 10 different stimuli were presented with equal probabilities, and indeed this was the case. Therefore, subtraction of the controls from the deviants elicits a positive difference (cf. Jacobsen & Schröger, 2001, 2003). The sharper positive peak elicited by the 100-ms-duration deviants relative to the 150-ms standards at about 170 ms probably manifests a different process. This difference, resulting from larger P2 for the deviant than standard (Figure 1) was also observed in previous studies with similar stimuli (Jacobsen & Schröger, 2003) and might be the response to the deviant offset (at 100 ms). However, both accounts are post hoc in nature, and, therefore, additional research is necessary to elucidate the origin of these electrophysiological manifestations.

These pre-MMN positivities showed apparent differences between the harmonic and pure tones. Although it remains for future research to directly target these effects, we note that in the current experiment these differences were not significant. Moreover, they cannot explain the critical finding of this study: When the controlled method is implemented, there is a larger deviance response for harmonic than pure tones for duration deviance, but not for pitch deviance. This is because if the MMN is measured from the previous positive peak (peak-to-peak), an even smaller response for harmonic than pure tones would be found in the case of pitch deviance, and the robust finding of a larger response to harmonic than pure deviants in the case of duration changes remains unaltered (see the control conditions in Figure 2).

Conclusion

This study suggests that the mechanisms generating the MMN for the different features differ not only in the location of their sources but might also have different functional structures. The results show that adding harmonics affects the memory-based MMN for duration but not for pitch. We propose that the amplitude modulation of duration MMN is due to a mechanism comparing the duration of each individual frequency, whereas pitch comparison is carried out after an integration stage. How other dimensions of sounds are treated within the mismatch detection apparatus remains to be investigated.

REFERENCES

American Electroencephalographic Society. (1994). Guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, *11*, 111–113.

Bendor, D., & Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature*, *436*, 1161–1165.

Budd, T. W., Barry, R. J., Gordon, E., Rennie, C., & Michie, P. T. (1998). Decrement of the N1 auditory event-related potential with stimulus repetition: Habituation vs. refractoriness. *International Journal of Psychophysiology*, *31*, 51–68.

Deacon, D., Nousak, J. M., Pilotti, M., Ritter, W., & Yang, C. M. (1998). Automatic change detection: Does the auditory system use representations of individual stimulus features or gestalts? *Psychophysiology*, *35*, 413–419.

Deouell, L. Y., Parnes, A., Pickard, N., & Knight, R. T. (2006). Spatial location is accurately tracked by human auditory sensory memory: Evidence from the mismatch negativity. *European Journal of Neuroscience*, *24*, 1488–1494.

Giard, M. H., Lavikainen, J., Reinikainen, K., Perrin, F., Bertrand, O., Pernier, J., et al. (1995). Separate representation of stimulus frequency, intensity, and duration in auditory sensory memory: An event-related potential and dipole-model analysis. *Journal of Cognitive Neuroscience*, *7*, 133–143.

Haenschel, C., Vernon, D. J., Dwivedi, P., Gruzelier, J. H., & Baldeweg, T. (2005). Event-related brain potential correlates of human auditory sensory memory-trace formation. *Journal of Neuroscience*, *25*, 10494–10501.

Jacobsen, T., & Schröger, E. (2001). Is there pre-attentive memory-based comparison of pitch? *Psychophysiology*, *38*, 723–727.

Jacobsen, T., & Schröger, E. (2003). Measuring duration mismatch negativity. *Clinical Neurophysiology*, *114*, 1133–1143.

Jääskeläinen, I., Ahvenien, J., Bonmassar, G., Dale, A., Ilmoniemi, R., & Levänen, S. (2004). Human posterior auditory cortex gates novel sounds to consciousness. *Proceedings of the National Academy of Sciences, USA*, *17*, 6809–6814.

Jung, T. P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2000). Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clinical Neurophysiology*, *111*, 1745–1758.

Kanwal, J. S., Fitzpatrick, D. C., & Suga, N. (1999). Facilitatory and inhibitory frequency tuning of combination-sensitive neurons in the primary auditory cortex of mustached bats. *Journal of Neurophysiology*, *82*, 2327–2345.

May, P., Tiitinen, H., Ilmoniemi, R. J., Nyman, G., Taylor, J. G., & Näätänen, R. (1999). Frequency change detection in human auditory cortex. *Journal of Computational neuroscience*, *6*, 99–120.

Medvedev, A., Chiao, F., & Kanwal, J. (2002). Modeling complex tone perception: Grouping harmonics with combination-sensitive neurons. *Biological Cybernetics*, *86*, 497–505.

Molholm, S., Martinez, A., Ritter, W., Javitt, D. C., & Foxe, J. J. (2005). The neural circuitry of pre-attentive auditory change-detection: An fMRI study of pitch and duration mismatch negativity generators. *Cerebral Cortex*, *15*, 545–551.

Moore, B. C. J., & Moore, G. A. (2003). Discrimination of the fundamental frequency of complex tones with fixed and shifting spectral envelopes by normally hearing and hearing-impaired subjects. *Hearing Research*, *182*, 153–163.

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Science*, *13*, 201–288.

Näätänen, R. (1991). Mismatch negativity outside strong attentional focus: A commentary on Woldorff et al. *Psychophysiology*, *28*, 478–484.

Näätänen, R. (2003). Mismatch negativity: Clinical research and possible applications. *International Journal of Psychophysiology*, *48*, 179–188.

Näätänen, R., & Escera, C. (2000). Mismatch negativity: Clinical and other applications. *Audiology and Neuro-otology*, *5*, 105–110.

Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early selective attention effect on evoked potential reinterpreted. *Acta Psychologica*, *42*, 313–329.

Näätänen, R., Jacobsen, T., & Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (MMN): A review of evidence. *Psychophysiology*, *42*, 25–32.

Näätänen, R., Päävilainen, P., Tiitinen, H., Jiang, D., & Alho, K. (1993). Attention and mismatch negativity. *Psychophysiology*, *30*, 436–450.

Penagos, H., Melcher, J. R., & Oxenham, A. J. (2004). A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. *Journal of Neuroscience*, *24*, 6810–6815.

Picton, T. W., Hillyard, S. A., & Galambos, R. (1976). Habituation and attention in the auditory system. In E. de Boer, W. D. Keidel, & W. D. Neff (Eds.), *Auditory System: Clinical and special topics, Handbook of sensory physiology vol. V/3* (pp. 343–390). New York: Springer-Verlag.

Rauschecker, J. P., Tian, B., & Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science*, *268*, 111–114.

Restuccia, D., Della Marca, G., Marra, C., Rubino, M., & Valeriani, M. (2005). Attentional load of the primary task influences the frontal but not the temporal generators of mismatch negativity. *Cognitive Brain Research*, *25*, 891–899.

Ritter, W., Deacon, D., Gomes, H., Javitt, D. C., & Vaughan, H. G. Jr. (1995). The mismatch negativity of event-related potentials as a probe of transient auditory memory: A review. *Ear and Hearing*, *16*, 52–67.

Rosburg, T. (2003). Left hemispheric dipole locations of the neuromagnetic mismatch negativity to frequency, intensity and duration deviants. *Cognitive Brain Research*, *16*, 83–90.

Schröger, E. (1995). Processing of auditory deviants with changes in one versus two stimulus dimensions. *Psychophysiology*, *32*, 55–65.

Sussman, E., Winkler, I., & Schröger, E. (2003). Top-down control over involuntary attention switching in the auditory modality. *Psychonomic Bulletin and Review*, *10*, 630–637.

Tervaniemi, M., Lehtokoski, A., Sinkkonen, J., Virtanen, J., Ilmoniemi, R. J., & Näätänen, R. (1999). Test-retest reliability of mismatch negativity for duration, frequency and intensity changes. *Clinical Neurophysiology*, *110*, 1388–1393.

Tervaniemi, M., Ilvonen, T., Sinkkonen, J., Kujala, A., Alho, K., Huotilainen, M., et al. (2000). Harmonic partials facilitate pitch discrimination in humans: Electrophysiological and behavioral evidence. *Neuroscience Letters*, *279*, 29–32.

Tervaniemi, M., Schröger, E., Saher, M., & Näätänen, R. (2000). Effects of spectral complexity and sound duration on automatic complex-sound pitch processing in humans—A mismatch negativity study. *Neuroscience Letters*, *290*, 66–70.

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

Ulanovsky, N., Las, L., & Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nature Neuroscience*, *6*, 391–398.

Winkler, I., Tervaniemi, M., Huotilainen, M., Ilmoniemi, R., Ahonen, A., Salonen, O., et al. (1995). From objective to subjective: Pitch representation in the human auditory cortex. *NeuroReport*, *6*, 2317–2320.

Winkler, I., Tervaniemi, M., & Näätänen, R. (1997). Two separate codes for missing-fundamental pitch in the human auditory cortex. *The Journal of the Acoustical Society of America*, *102*, 1072–1082.

Woldorff, M. G., Hackley, S. A., & Hillyard, S. A. (1991). The effects of channel-selective attention on the mismatch negativity wave elicited by deviant tones. *Psychophysiology*, *28*, 30–42.

Woldorff, M. G., Hillyard, S. A., Gallen, C. C., Hampson, S. R., & Bloom, F. E. (1998). Magnetoencephalographic recordings demonstrate attentional modulation of mismatch-related neural activity in human auditory cortex. *Psychophysiology*, *35*, 283–292.

(Received November 24, 2006; Accepted May 3, 2007)