A mismatch negativity study of local–global auditory processing

Alexandra List[a,b,c,*], **Timothy Justus**[a,b], **Lynn C. Robertson**[a,b], and **Shlomo Bentin**[d]

a*University of California, Berkeley, USA*

b*Veterans Affairs, Northern California Health Care System, Martinez, USA*

c*University of Wales, Bangor, UK*

d*Hebrew University, Jerusalem, Israel*

Abstract

We used mismatch negativity (MMN) to examine structural encoding of local and global auditory patterns in perceptual memory. Unlike previous MMN studies of local–global auditory perceptual organization that used *interval-contour* stimuli, here we presented hierarchical stimuli in which local pattern organization formed global patterns. Importantly, our stimuli allowed independent manipulation of the two structural levels. In separate blocks, participants were exposed to frequent local *standard* patterns and rare local *deviant* patterns, or to frequent global *standard* patterns and rare global *deviant* patterns. Within each deviant pattern, the variation from the standard pattern could occur at onset (*early*), towards the end of the pattern (*late*) or over both time windows (*both*). To isolate pattern indexing at one level, the other level continuously changed (e.g., in a global standard block, local elements varied trial-by-trial). MMN was found only for global deviant patterns, and only when deviation occurred late in the pattern. In a separate behavioral experiment, global deviants were detected more often than local ones, although initial similarity followed by a late deviation from the standard pattern was not required for explicit deviant detection (as with the MMN). This report demonstrates neural structural encoding for global information, when independently manipulated from local information. Furthermore, it extends previous MMN findings that have revealed indexing of complex abstract auditory information to the realm of hierarchical perceptual organization.

Keywords

Mismatch Negativity; ERP; Audition; Perceptual Organization; Local; Global

1. Introduction

1.1. Mismatch negativity

Numerous studies have employed electro- and magneto-encephalography (EEG and MEG, respectively) in mismatch negativity (MMN) experiments to examine the processing of various auditory stimulus attributes (for reviews, see Näätänen and Winkler, 1999; Näätänen, 1995; Näätänen and Alho, 1995; Tervaniemi and Huotilainen, 2003). Mismatch negativity has proven highly informative in assessing the automatic discrimination of both physical and abstract auditory properties, ranging from simple acoustic features such as the frequency, intensity or duration of pure tones (e.g., Novitski et al., 2004; Jaramillo et al., 2000; Paavilainen et al., 1991) to more complex musical or phonetic patterns (e.g., Fujioka et al., 2004; Näätänen et al.,

* *Corresponding author*. School of Psychology, Adeilad Brigantia, Penrallt Road, University of Wales, Bangor, Gwynedd, LL57 2AS, UK. E-mail address: a.list@bangor.ac.uk (A. List)..

1997; Paavilainen et al., 1999; Saarinen et al., 1992; Tervaniemi et al., 1994; van Zuijen et al., 2004, 2005).

The classic MMN design involves frequent presentation of a *standard* pattern and rare presentations of a *deviant* pattern (varied along an attribute of interest). In representative cases, the averaged event-related potential (ERP) elicited by the deviant condition deflects negatively relative to that elicited by the standard condition. This difference begins typically about 100 ms after deviant onset and lasts for about 100 ms (Näätänen and Winkler, 1999). This relative negativity is best illustrated by a difference waveform, i.e., by plotting the deviant minus standard voltage over time. Prevailing models posit that the MMN reveals a discrepancy between the incoming auditory deviant stimulus and a memory template established by repetition of the standard stimulus. With EEG, MMN is often measured over frontal sites. Sources of the signal have been identified bilaterally in the supratemporal plane, as well as in the right frontal lobe (e.g., Deouell et al., 1998; Giard et al., 1990; Kasai et al., 1999; Kropotov et al., 1995; Levänen et al., 1996; Rinne et al., 2000). Importantly, MMN is not evoked under conditions of continuously changing stimuli, unless the so-called deviant is itself a repetition (*repetition negativity*,[1] e.g., Horváth and Winkler, 2004). Although the MMN can be modulated by attention, its elicitation does not depend on it (Alain and Woods, 1997; Näätänen et al., 1993); indeed, it is frequently recorded while listeners are actively engaged in an unrelated task (e.g., Alho et al., 1994). More remarkably, it can even be recorded in comatose patients (Fischer et al., 2000; Kane et al., 1993), revealing the MMN to be a powerful tool for investigating what dimensions of auditory input are processed automatically.

1.2. Studies of musical structure

In the present study, we focused on the encoding of auditory structural information. More specifically, we investigated how the brain parses a barrage of complex auditory input into structural levels, i.e., into local and global aspects of the stimulus. These structural levels may carry different forms of meaningful or otherwise relevant information. The majority of previous investigations have been carried out in the domain of music cognition, in which structural levels have typically been investigated using *interval-contour*[2] stimuli (Fig. 1). *Interval-contour* stimuli were originally employed in behavioral investigations (e.g., Dowling and Fujitani, 1971), and were later used to investigate theories of hemispheric asymmetry for music processing (e.g., Peretz, 1987,1990;Zatorre, 1985). These investigations addressed questions related to the role of contour information in memory for melodies, and whether the cerebral hemispheres are asymmetrically biased to process different aspects of musical structural information (with equivocal results for the latter issue).

More recently, *interval-contour* stimuli have been used in MMN designs to further elucidate the neural underpinnings of melodic structural processing (e.g., Schiavetto et al., 1999; Trainor et al., 2002; Fujioka et al., 2004). In one such study (Schiavetto et al., 1999), only deviant contours (i.e., deviant "global" structures) reliably elicited MMN. In a second study, Trainor et al. (2002) addressed the potential difference(s) between explicit and implicit melodic structural processing, using both active and passive tasks. They showed that, collapsed, interval and contour deviants elicited an MMN, whether in the active or passive condition. However, they did not report whether interval or contour manipulations *alone* revealed MMN in either the active or the passive tasks. Lastly, Fujioka et al. (2004) compared passive structural processing of melodies in non-musicians vs. musicians, recording the magnetic MMN (mMMN). They aimed to determine whether musical expertise altered the neural indexing of

---

[1]In repetition negativity experiments, the standard pattern is one that always changes, and thus when a repetition occurs, the pattern has been violated. As in other MMN studies, a deviation in the pattern causes a negative deflection of the waveform.
[2]In *interval-contour* stimuli, the contour of a musical phrase is the pattern of pitch rises and falls, whereas an *interval* is the precise increment of a pitch change between two tones.

musical structure. These authors only found mMMN for interval conditions in non-musicians, whereas they found mMMN for both interval- and contour-deviance in musicians (mMMN for interval-deviance was larger than for contour-deviance). Interestingly, this study indicated that musical experience was required for *passive* contour-deviance mMMN to emerge, whereas the results from Schiavetto et al. (1999) showed that in an *active* task, contour-deviance MMN emerged more readily (than did interval-deviance MMN) in the musically naive. Taken together, these three studies have reported MMNs elicited by both interval and contour deviance, although not consistently across studies (and the results are not easily ascribed to musical expertise and/or active vs. passive tasks). Nevertheless, the findings do provide evidence for automatic structural encoding of interval and contour dimensions in melodic stimuli, as measured by the MMN.

Despite the contributions previous studies using *interval-contour* stimuli make to the domain of music cognition, from our present perspective there are a number of shortcomings inherent to using *interval-contour* stimuli in studies of general auditory structural processing, especially when couched in terms of local–global structure. One limitation is that the *interval-contour* mapping onto local–global structure is neither a transparent nor a unique one. In fact, so-called "global" manipulations of the melodic sequences have not only involved contour manipulations but also violations of key and scale (e.g., Peretz, 1987). The flexibility in terminology may suggest a lack of consensus on what constitutes a global manipulation.[3] More importantly, whichever the global manipulation, *interval-contour* stimuli do not allow for independent control of levels; a global change always requires a local one as well (see Fig. 1). Consequently, researchers employing *interval-contour* melodic variations have recently downplayed the mapping between melodic structure and local–global structure (e.g., Ayotte et al., 2000; Fujioka et al., 2004). This might indicate recognition of the failure for the *interval-contour* stimuli to adequately capture local–global perceptual structure. For these reasons, we took a new approach.

2. The MMN experiment

In the present experiment, we sought to investigate whether local and global structural information is encoded by the auditory system by analyzing MMN measures. Our approach did not involve the use of *interval-contour* stimuli and consequently avoided the abovementioned inherent limitations of those stimuli. To achieve our goal, we adopted the hierarchical stimuli developed by Justus and List (2005, Experiment 2) in which local elements are arranged to produce a global pattern. In these stimuli, hierarchical structure is created by sequences of tones unfolding at two different temporal scales: local patterns which emerge over 300 ms and global patterns which emerge over 900 ms. As in the widely used visual local–global stimuli (Navon, 1977), the auditory local and global levels are manipulated independently of one another (Fig. 2). With independent local and global manipulations, data interpretation does not rely on comparisons of local vs. local-with-global conditions (as in the *interval-contour* investigations) but instead relies on direct comparisons of local-only vs. global-only manipulations. This independence is of particular importance when trying to disentangle neural contributions to processing of each level.

---

[3]The terms "global" and "local" have also been used in other contexts. For instance, Horváth et al. (2001) explored temporal rules for eliciting a MMN. When presenting a series of pure tones, adjacent relationships between adjacent tones were deemed local, and relationships between temporally non-adjacent sounds were deemed global. They found that deviants to *both* types of rules elicited a MMN. Our study's aim was quite different than theirs: we examined the neural indexing of local and global hierarchical structure internal to the stimulus. In their study, however, the local elements did not combine to produce a global element. In other words, they examined temporal relationships between auditory objects/events whereas we are examining the structural relationships within auditory objects/events.

The hierarchical stimuli used in this experiment are visually depicted in Fig. 3. Three complex tones were presented sequentially without inter-stimulus intervals (ISIs) to form four different *local* patterns: rising–rising (rr), rising–falling (rf), falling–falling (ff) and falling–rising (fr). In each trial, one local pattern was repeated sequentially three times without ISIs, to form one of the four patterns at the *global* level (also rr, rf, ff and fr). Because there were no ISIs, presenting nine 100-ms duration tones resulted in a 900-ms sequence. We will refer to each 900-ms sequence as a trial. Global and local patterns were never redundant (i.e., a local rr pattern was never presented three times to produce a global rr pattern, resulting in 12 total stimulus configurations.)

While undergoing EEG recording, sixteen participants watched a silenced movie of their choice and were instructed to ignore the auditory stimuli presented through headphones. In each of eight blocks, one of the four patterns (rr, rf, ff or fr), at either the local or the global level, served as the standard. For example, in a local rr standard block, local rf, ff or fr patterns served as deviants. Importantly, by nature of the hierarchical stimuli, local and global patterns were never presented in isolation: all 900-ms sequences were composed of patterns at the local and global levels. In local-standard blocks, the global patterns were randomly varied and in global-standard blocks, the local patterns were randomly varied (Figs. 4A and B). This design was adopted to promote the isolation of local and global processing because local and global patterns were varied orthogonally, information from the irrelevant level could not contribute to indexing the standard pattern in auditory memory at the opposite level. This design allowed us to test whether the brain is able to extract and index a pattern at one level of structure while being presented with distracting (and always changing) information at another level of structure.

The deviant trials fell into three categories: *early*, *late* or *both* (Fig. 4). For instance, given an rr standard pattern, fr patterns deviated only at the first element (*early*), rf patterns deviated only at the third element (*late*) and ff patterns deviated at both the first and third elements (*both*). Within a block, the role of each stimulus was determined solely by context (e.g., in an rr standard block, fr patterns were *early* deviants, whereas in a ff standard block, fr patterns were *late* deviants). Fig. 4C illustrates the standard patterns and the *early*, *both* and *late* deviant patterns for each of the eight blocks (4 patterns × 2 levels of structure). For analyses, the waveforms for individual patterns were collapsed into *early*, *late* or *both* conditions.

3. Results

3.1. ERP results

To determine whether local and global structure was indexed by the auditory system as indicated by MMN, planned comparison paired sample *t*-tests were carried out at site Fz (a frontal midline site). Local and global standards were compared to each of the three deviant conditions (*early, late* and *both*) over matched time periods (collapsed over pattern). Since the number of deviant trials was necessarily smaller than the number of standard trials, we randomly selected 12 standard trials from each block to match the number of deviants per condition. The resulting means of 48 responses to standard trials (12 trials for each of 4 blocks, for local and global blocks separately) formed the standard conditions to which local and global deviants were compared.

The dependent variable in this experiment was the mean amplitude[4] at Fz across 100-ms time windows starting 100 ms after the onset of each of the nine tones in a trial (see Figs. 5A and 6A). Relative to the onset of the first tone in each trial, *early* local deviations began at times

---

[4]Mean amplitude measures were chosen as they are more stable than peak measures (e.g., as argued by Luck, 2005, Chapter 6) and because the measure is widely used in MMN studies (e.g., Näätänen et al., 2004; van Zuijen et al., 2004).

0, 300 and 600 ms (mean amplitude [μV] was taken over the time windows: 100–200, 400–500 and 700–800 ms) and *late* local deviations began at 200, 500 and 800 ms (mean amplitude [μV] was taken over the time windows: 300–400, 600–700 and 900–1000 ms). *Early* global deviations occurred at 0, 100 and 200 ms (mean amplitude [μV] was taken over the time window: 100–400 ms) and *late* global deviations occurred at 600, 700 and 800 ms (mean amplitude [μV] was taken over the time window: 700–1000 ms). Whether analyzing local or global conditions, in the *early* and *late* analysis windows, the mean voltage was calculated over a total of 300 ms. In the *both* condition, however, the analysis windows included the *early* and *late* windows and thereby reflected the mean voltage over 600 ms. Mean voltage amplitude for deviant conditions were always compared to the mean voltage amplitude for the standard condition *over the same time window*.

As shown in Fig. 5, none of the ERPs associated with local deviance were different from the ERPs associated with local standard pattern presentation ($-1.00 <$ *t*-values< 1.00). Fig. 6 illustrates the global level waveforms, in which only *late* global deviant conditions elicited a significant MMN [$t(15) = -2.24$, $p < 0.05$]. Neither *early* nor *both* global deviant waveforms differed reliably from the standard waveform ($-1.00 <$ *t*-values< 1.00).

The *late* global condition was further explored to examine whether the MMN varied by hemisphere (based on theories proposing a global perceptual bias of the right hemisphere, e.g., Ivry and Robertson, 1998). The data were analyzed in a 2×2 repeated measures ANOVA with condition (standard/deviant) and electrode site (F5-left frontal site/F6-right frontal site) as factors. None of the effects were reliable (all *F*-values < 1.00; Fig. 7).

As previously stated, the introduction of new frequencies can elicit an MMN (shown even when using complex sequences of stimuli, as in van Zuijen et al., 2005). A secondary analysis was performed to verify that the MMN was attributable to deviations from the abstracted global structure, and not simply carried by conditions in which the deviant introduced or reduced the frequencies presented relative to the standard condition. The *late* global level deviant waveforms were separated into those that introduced new frequencies not present in the standard (e.g., an rr deviant in an rf standard block) and those that presented fewer frequencies than did the standard (e.g., an rf deviant in an rr standard block). For *late* global level conditions, deviants with more or fewer frequencies did not differ ($-1.00 <$ *t*-values< 1.00). Our data provide no evidence for introduction or reduction of frequencies carrying the *late* global MMN effect (Fig. 8).

3.2. Behavioral experiment

Since the detection of MMN has sometimes been associated with perceptual thresholds (Sams et al., 1985) and because we failed to find an MMN for local conditions, we ran a behavioral experiment to examine whether explicit perceptual thresholds varied between local and global conditions. If explicit report mirrors the MMN data, the hypotheses are that local deviance sensitivity is worse than global deviance sensitivity, and that *late* global deviant identification is better than for *early* or *both* global deviants.

In this experiment, twelve new participants were instructed to identify deviants among blocks of stimuli with local or global standards (with a 1:9 ratio, as in the EEG experiment). D-prime (d') measures were calculated from the z-score transformed proportions of detected deviants and false alarms to standard stimuli. In a structural level (local, global) $\times$ time window (*early, both, late*) within-subjects ANOVA for d', both main effects and their interaction were significant [level: $F(1,11) = 7.69$, $p < 0.05$; time window: $F(2,22) = 7.20$, $p < 0.01$; interaction: $F(2,22) = 4.01$, $p < 0.05$]. Consistent with our hypothesis, we found that participants were better able to detect deviants in the global condition [$d' = 1.05$] than in the local condition [$d' = 0.68$]. Regardless of level, *both* deviants [$d' = 1.01$] were detected more reliably than *early* [$d' = 0.86$;

$t(11)=2.43$, $p<0.01$] or *late* deviants [$d' =0.70$; $t(11)=3.87$, $p<0.01$], with no reliable difference between the latter two [$t(11)=1.58$, $p>0.10$]. Examining the interaction, for local and global deviants separately, *late* local deviants [$d' =0.39$] were detected worse than *early* [$d' = 0.76$, $t(11)=2.82$, $p=0.017$] or *both* local deviants [$d' =0.88$, $t(11)= 4.57$, $p=0.001$]. Unlike local deviants [$F(2,22)=9.88$, $p=0.001$], no effect of time interval was found for global deviants [*early* $d' = 0.96$, *both* $d' =1.14$ and *late* $d' =1.02$; $F(2,22)=1.40$, $p=0.27$] when participants were actively involved in detecting deviants. In sum, explicit global deviant detection was better than explicit local deviant detection (as with the MMN), although no advantage for *late* vs. *early* or *both* global deviant detection was found (unlike with the MMN).

4. General discussion

The present investigation explored whether the structure of hierarchical local and global patterns was indexed in auditory memory, as manifested by the MMN. Local and global stimulus structures were independent from one another and were presented to participants while they watched a silenced movie. No electrophysiological evidence was found for local structural indexing at any time window, whereas global patterns varying in the final local element elicited a reliable MMN. Below, we discuss several possibilities for why we found evidence for global (but not local) deviance indexing, as well as potential reasons for the MMN found at *late* (but not at *early* or *both*) temporal intervals.

4.1. Why global and not local?

Our results are in line with studies using *interval-contour* stimuli, in which contour variations are better detected than interval-only variations, as shown in performance by larger d' measures (Zatorre, 1985) and electrophysiologically by larger or earlier MMNs (Schiavetto et al., 1999; Trainor et al., 2002, respectively). However, as argued in the introduction, studies using *interval-contour* stimuli are not conclusive due to confounds in manipulating the global and local dimensions independently (Fig. 1). In contrast, in the current experiments, local and global levels were independently manipulated and only block context determined whether a given stimulus was considered a local or global deviant. Therefore, compared to the results from *interval-contour* studies, the differences found between global and local conditions in our study provide stronger evidence for differences of structural indexing in auditory memory.

The detection of global, but not local, MMN might be related to the greater salience of the global manipulation compared to the local one (given the performance results). We did rule out the possibility that new or fewer frequencies carried the MMN effect found with *late* global deviants. Another possible frequency difference between global and local patterns, that might have enhanced the salience of the former, is that local pattern tones spanned a (necessarily) smaller perceptual range. Spanning smaller frequency and temporal ranges may have made these patterns more difficult for the system to resolve than their global counterparts. However, there are at least two arguments against this as the explanation for the differences between local and global conditions. Firstly, the MMN has been recorded for stimuli whose frequency or temporal deviations were within the local ranges presented, or smaller (e.g., Jaramillo et al., 2000; Novitski et al., 2004; Näätänen et al., 1993; Sams et al., 1985; Tervaniemi et al., 1999). Secondly, from behavioral studies, we know that it is at least possible for individuals to discriminate between local patterns varying in just one tone (Justus and List, 2005), despite the behavioral data reported here revealing an asymmetry between structural levels when actively detecting deviants.

Salience differences between local and global conditions may have alternatively been due to silences creating structural boundaries. Namely, global three-element patterns were delimited by silence (the inter-trial intervals), whereas there were no such breaks between local three-tone elements. The silence intervals may have consequently had the effect of emphasizing both

the beginning and end of the global patterns, whereas the local patterns benefited from no such delimitation.

However, it is also worth considering that the global patterns in this experiment were inherently more perceptually salient than the local patterns were, and that this difference is not merely the result of absolute range differences or an extra structural boundary cue at the global level. There are two interrelated aspects worth considering.

Firstly, principles of auditory perceptual organization do not operate equivalently over all temporal scales. At the short extreme, auditory perceptions unfolding over 100 ms or less will likely be perceived as a single event, regardless of any additional internal structure. Beyond about 1500 ms, when echoic memory has reached its limit, principles of sequence perception begin to yield to higher order principles of temporal organization (McAdams and Drake, 2002). Within the sequence perception range of 100–1500 ms, further qualitative subdivisions have been proposed. For instance, Fraisse (1963; see Clarke, 1999) argued that an interval of approximately 600 ms marks a qualitative boundary in temporal perception, a window which falls between our local and global temporal scales.

Secondly, it could also be the case that our global patterns were more perceptually salient not only because of the specific temporal range over which they unfolded, but because they were global patterns per se. Two kinds of temporal hierarchies were proposed by Lerdahl and Jackendoff (1983): a grouping hierarchy operating over units and a metrical hierarchy operating over beats. Units and beats at subordinate levels of the hierarchy help define and reinforce the global level as it unfolds over time, and it may be the case that any locally reinforced global form (if occurring within the range of sequence perception) will be more salient to the listener.

It is also interesting to note that, because of the hierarchical structure of the stimuli, the system is given three opportunities to resolve the local patterns in every trial compared to a single opportunity to resolve the global pattern. This attribute of the stimuli generates an opposite prediction to what was found: that local patterns and deviations thereof would be better registered than global ones. Importantly, this repetition hypothesis was ruled out by our findings.

4.2. Why no lateralization?

Despite the fact that we used stimuli with independent local and global levels, we were unable to detect a hemispheric asymmetry for MMN, even for *late* global processing. This result is consistent with previous findings using *interval-contour* stimuli in similar designs (Fujioka et al., 2004; Schiavetto et al., 1999; Trainor et al., 2002) and is not consistent with the proposed right hemisphere bias for global processing that applies to *both* vision and audition (Ivry and Robertson, 1998). Hemispheric asymmetries have been reported for MMN when testing other hypotheses and using other stimuli (e.g., Deouell et al., 1998), although they are more often found using MEG (e.g., Levänen et al., 1996; Shtyrov et al., 1998). Hence, although we did not observe laterality effects in our passive MMN design, it remains possible that laterality effects would emerge under other conditions (e.g., when using MEG or when participants engage in an active discrimination task while undergoing EEG). It is also possible that the auditory system does operate asymmetrically for processing structural information, only over temporal ranges that we did not probe. In his asymmetric sampling in time theory, Poeppel (2003) has suggested that hemispheric asymmetries exist for the processing of *absolute* temporal ranges, contrary to Ivry and Robertson's *relative* theory (1998). Poeppel argues that the left hemisphere is biased to process information within temporal windows of 25–50 ms, and the right is biased to process information within temporal windows of 150–250 ms (neither of which were probed here).

In a recent study by Sanders and Poeppel (2006), they employed stimuli similar to ours in an ERP study with the aim of addressing whether local and global temporal processing occurs asymmetrically in the brain. Instead of patterns being defined by sequences of individual tones, they used local frequency-modulated (FM) sweeps that rose or fell and arranged these to form rising or falling global patterns. The stimuli were either redundant (e.g., a 40-ms rising local FM sweep repeated three times with 190-ms ISIs to form a 500-ms globally rising pattern) or in conflict (e.g., a rising local FM sweep repeated three times to form a globally falling pattern). Participants in the study performed a directed-attention task in which they reported the direction of the local or global patterns in separate blocks. When comparing the conditions of interest, they reported a sustained negativity between 250 and 700 ms for the global condition (compared to the local condition). They also found ERP evidence for global interference in local blocks that was not mirrored by local interference in global blocks (despite equal interference effects in performance). Lastly, no hemispheric lateralization was observed in their study. It should be noted, however, that they did not probe the global range (150–250 ms) presented in Poeppel's (2003) original work. So, while lateralization of local–global auditory processing remains somewhat elusive, their study does converge with ours and others' in finding stronger electrophysiological evidence for global vs. local processing under the conditions tested.

4.3. Why only late global?

The presence of *late* MMN but not *early* or *both* MMNs suggests something unique about the formation of an auditory template for global processing. It is notable that the *both* global condition deviated from the standard pattern in the late time window as well. However, global pattern indexing was only reliably apparent electrophysiologically when the initial elements of a pattern met expectation, only to be later violated. At first, this seems paradoxical: the *late* violation only elicited an MMN when it initially matched, whereas no MMN was elicited when it violated the standard initially and finally. By design, because the local level constantly changed, it was impossible for precise tones or local patterns to anchor the global stimulus at its onset. In other words, the initial element of the global pattern was not enough to determine whether the global pattern matched or not. In the case of *late* deviations; however, the pattern established by the first two local constituent elements could form a template match, only to be violated by the third element. This finding suggests that the *late* global MMN was in fact indexing global structure. This interpretation remains tentative because a similar advantage of *late* global deviance was not observed in behavior (*late* was undifferentiated from *early* and *both* behavioral conditions). However, as we described in the introduction, implicit and explicit processing do not always produce parallel results (nor need they).

In conclusion, avoiding methodological weaknesses evident with the use of *interval-contour* stimuli, the present study shows that the emergent global structure of a hierarchically organized sequence of tones can be indexed in auditory memory. Further, the findings suggest that the detection of global deviance is limited to stimuli that matched the standard pattern at onset. Stimuli initially conforming to the standard pattern arguably enable the system to relate the auditory hierarchical input to the global template established by the repeated standard pattern, i.e., the stimuli require anchoring to the initial portion of the global template.

5. Experimental procedures

5.1. MMN experiment

5.1.1. Participants—Twenty-one participants gave informed consent before undergoing EEG recording. They were financially compensated for participation. Five participants' data were excluded from analyses due to excessive movements and/or noise in critical EEG channels. All remaining participants (mean age = 23 years, age range = 19–29 years) were

right-handed, as determined by the Edinburgh handedness inventory (mean = 87%, range = 41–100%; Oldfield, 1971).

5.1.2. Auditory stimuli—The auditory stimuli were constructed from complex tones, ranging in fundamental frequency from 185 to 467 Hz, in whole step increments (F#$_3$–A#$_4$). The next four harmonics were presented at 1/n loudness. Each 100-ms complex tone ramped on and off over 10 ms. Stimuli were presented to *both* ears at ~70 dB SPL.

Three complex tones were grouped sequentially without inter-stimulus interval (ISI) to form four different 300-ms local patterns (shown in Fig. 3): rising–rising (rr), rising–falling (rf), falling–rising (fr) and falling–falling (ff). Each 300-ms pattern could occur over one of three different frequency ranges: 185–233 Hz (F#$_3$–A#$_3$), 261–330 Hz (C$_4$–E$_4$) and 370–467 Hz (F#$_4$–A#$_4$).

One 300-ms local pattern was repeated sequentially three times without ISI's over the frequency ranges to form the same four patterns at the *global* level (rr, rf, fr and ff). Consequently, global patterns emerged over 900 ms. For example, a local rr pattern and a global fr pattern was formed from [F#$_4$–G#$_4$–A#$_4$]–[C$_4$–D$_4$–E$_4$]–[F#$_4$–G#$_4$–A#$_4$]. Global and local patterns were never redundant, resulting in 12 stimulus configurations (see Fig. 3).

5.1.3. Design—Each participant was presented with eight, 489-trial blocks. Each block contained one *standard* pattern (rr, rf, fr or ff), either local or global. Blocks with local or global standards alternated were separated by brief breaks and were counter-balanced across participants (e.g., global-rf, local-fr, global-ff, local-rr, global-rr, local-ff, global-fr, local-rf).

Each block began with nine standard trials. In the remainder of the block, trials were randomly interleaved. Standard stimuli were presented nine times more often than deviant patterns (432:48). Silent inter-trial intervals lasted 500–600 ms.

In local standard blocks, the global patterns randomly varied and in global standard blocks, the local patterns randomly varied. This design was adopted to promote the isolation of local and global processing; information from one level could not contribute to pattern identification at the opposite level. In each block type, only nine of the 12 stimuli were used to avoid pattern redundancy between levels, e.g., in a local rr standard block, none of the global rr stimuli were presented (Fig. 3). The role of each stimulus within a block was determined solely by context.

5.1.4. Procedure—After being prepared for EEG scalp recording, participants viewed a self-selected silenced movie while auditory stimuli were presented binaurally through headphones. Participants were instructed to ignore the sounds, and to keep blinking and eye movements to a minimum. The experiment was conducted in a dim sound-treated chamber.

5.1.5. EEG recording and processing—A 64-channel Biosemi system was used to record EEG signals. Six additional external electrodes were used: one as a nose reference, two to record from each of the left and right mastoids, and three others for detecting eye movements (one placed lateral to the outer canthus of each eye for the identification of horizontal eye movements, and one placed below the left eye for the identification of blinks and vertical eye movements).

Signals from each 12-min block were recorded at a sampling rate of 256 Hz. Because of the necessarily small number of trials for deviant conditions, we used independent component analysis (ICA) to isolate and correct for eye blinks and movements. For each block, for each participant, the data were referenced to the nose and high-pass filtered at 0.5 Hz. ICA was then run on the data and produced a set of components. When a component was consistent with

reflecting ocular artifacts (in *both* topographical distribution and timing registry with the EEG), the component was set to 0, and the data were restored to their original format. Other artifacts were marked for removal through manual inspection or if a 100-µV change occurred within a 100-ms time window. The data were segmented into 1500-ms epochs (200 ms pre-stimulus to 1300 ms post-stimulus onset). To equate cell sizes, 12 (the average number of deviant segments) randomly selected standard segments were averaged for each block and each individual.[5] Deviant segments were also averaged for each block and individual. The averaged data were filtered (1–30 Hz, slope 24 dB/octave) and baseline corrected to 100 ms pre-stimulus.

At a given electrode site, the average amplitude was calculated over 100-ms increments beginning 100 ms after each tone's onset (the time range over which the MMN was expected). This was done for each of ten conditions: 2 (local, global)×5 (standard, standard randomly selected subset, *early* deviant, *late* deviant and *both* deviant). For example, the local standard measure was calculated over times 100, 400 and 700 ms (each of 100-ms duration). That measure was compared to the local *early* deviant, at times 100, 400 and 700 ms (over 100 ms), for a total of 300 ms. Conditions were collapsed over specific patterns.

5.2. Behavioral experiment

5.2.1. Participants—Twelve participants from UC Berkeley's summer research participation pool in the Department of Psychology gave informed consent and received course credit for their volunteered time. All participants (mean age=21 years, age range=18–24 years) were right-handed, as determined by the Edinburgh handedness inventory (mean=89%, range=67–100%).

5.2.2. Stimuli—The stimuli used in the EEG experiment were also employed here.

5.2.3. Design—Each block was composed of one repeated standard pattern, either at the local or global level. As in the EEG experiment, patterns continuously changed at the other level. *Standard* patterns were always presented for the first nine trials to establish it as standard. Following those trials, 54 standards were randomly presented with six deviants (9:1 standard: deviant) in each block. As before, all participants were exposed to eight standard types (2 levels × 4 patterns). Block order was randomized.

5.2.4. Procedure—Each participant was informed that within a block of trials, one pattern would be repeated frequently. They were instructed that when the pattern in a trial varied, they were to press a mouse button to indicate the presence of a deviant. Participants were then shown a visual aid depicting the local–global hierarchical stimuli (similar to Fig. 3). The experimenter demonstrated the difference between a local and global pattern, and verified that each participant understood by prompting them to indicate one example on the visual aid of a local pattern and one example of a global pattern. Once participants demonstrated their understanding, they were presented with two practice blocks. One practice block had a local standard pattern, and one had a global standard pattern. In each practice block, the first nine trials were standard patterns and one deviant followed within four trials (for a total of 13 trials). Participants then proceeded through eight blocks of experimental trials. They were informed

---

[5]MMN designs inherently demand that standard stimuli be presented far more frequently than deviants (a 9:1 ratio in the present experiment). To account for possible discrepancies in signal-to-noise ratios between standard and deviant conditions, standard EEG trials were sub-sampled for averaging (12 trials/block, which matched the average deviants/block). To verify that the randomly selected subset standard averages were adequate approximations of the standard all-trial averages, a comparison was made between all-trial standard averages and the subset standard averages at site Fz. Barring one comparison at 600–700 ms from local blocks, neither local nor global standard subset averages showed differences from standard all-trial averages (−1.00< *t*-values<1.00, at all time windows). The only reliable difference found indicated that the subset average was more negative than the all-trial average, which if anything would result in a more conservative estimate of the standard condition (since deviant waveforms are expected to deflect negatively with respect to the standard waveform). We therefore deemed the randomly selected standard subsets satisfactory and used them as standards in the analyses.

that each block contained more than one deviant, to encourage them to attend to the stimuli for the duration of the block.

5.2.5. Data analysis—Proportions of hits and false alarm were calculated for each participant in local and global conditions separately. These proportions were z-score transformed and d' scores were calculated. A 2×3 repeated measures ANOVA with level (local/global) and time window (*early*/*both*/*late*) was performed on the d' scores.

References

Alain C, Woods DL. Attention modulates auditory pattern memory as indexed by event-related brain potentials. Psychophysiology 1997;34(5):534–546. [PubMed: 9299908]

Alho K, Woods DL, Algazi A. Processing of auditory stimuli during auditory and visual attention as revealed by event-related potentials. Psychophysiology 1994;31(5):469–479. [PubMed: 7972601]

Ayotte J, Peretz I, Rousseau I, Bard C, Bojanowski M. Patterns of music agnosia associated with middle cerebral artery infarcts. Brain 2000;123:1926–1938. [PubMed: 10960056]

Clarke, EF. Rhythm and timing in music. In: Deutsch, D., editor. The Psychology of Music. second edition. Academic Press; San Diego: 1999. p. 473-500.

Deouell LY, Bentin S, Giard MH. Mismatch negativity in dichotic listening: evidence for interhemispheric differences and multiple generators. Psychophysiology 1998;35(4):355–365. [PubMed: 9643050]

Dowling WJ, Fujitani DS. Contour, interval, and pitch recognition in memory for melodies. J Acoust Soc Am 1971;49(2 Suppl 2):524–531. [PubMed: 5541747]

Fischer C, Morlet D, Giard M. Mismatch negativity and N100 in comatose patients. Audiol Neuro-otol 2000;5:192–197.

Fraisse, P. The Psychology of Time. Harper and Row; New York: 1963.

Fujioka T, Trainor LJ, Ross B, Kakigi R, Pantev C. Musical training enhances automatic encoding of melodic contour and interval structure. J Cogn Neurosci 2004;16:1010–1021. [PubMed: 15298788]

Giard MH, Perrin F, Pernier J, Bouchet P. Brain generators implicated in the processing of auditory stimulus deviance: a topographic event-related potential study. Psychophysiology 1990;27(6):627–640. [PubMed: 2100348]

Horváth J, Winkler I. How the human auditory system treats repetition amongst change. Neurosci Lett 2004;368(2):157–161. [PubMed: 15351440]

Horváth J, Czigler I, Sussman E, Winkler I. Simultaneously active pre-attentive representations of local and global rules for sound sequences in the human brain. Brain Res Cogn Brain Res 2001;12(1):131–144. [PubMed: 11489616]

Ivry, RB.; Robertson, LC. The Two Sides of Perception. The MIT Press; Cambridge, MA, USA: 1998.

Jaramillo M, Paavilainen P, Näätänen R. Mismatch negativity and behavioural discrimination in humans as a function of the magnitude of change in sound duration. Neurosci Lett 2000;290(2):101–104. [PubMed: 10936687]

Justus T, List A. Auditory attention to frequency and time: an analogy to visual local–global stimuli. Cognition 2005;98:31–51. [PubMed: 16297675]

Kane NM, Curry SH, Butler SR, Cummins BH. Electrophysiological indicator of awakening from coma. Lancet 1993;341:688. [PubMed: 8095587]

Kasai K, Nakagome K, Itoh K, Koshida I, Hata A, Iwanami A, Fukuda M, Hiramatsu KI, Kato N. Multiple generators in the auditory automatic discrimination process in humans. Neuroreport 1999;10(11):2267–2271. [PubMed: 10439446]

Kropotov JD, Näätänen R, Sevostianov AV, Alho K, Reinikainen K, Kropotova OV. Mismatch negativity to auditory stimulus change recorded directly from the human temporal cortex. Psychophysiology 1995;32(4):418–422. [PubMed: 7652119]

Lerdahl, F.; Jackendoff, R. A Generative Theory of Tonal Music. MIT Press; Cambridge, MA: 1983.

Levänen S, Ahonen A, Hari R, McEvoy L, Sams M. Deviant auditory stimuli activate human left and right auditory cortex differently. Cereb Cortex 1996;6:288–296. [PubMed: 8670657]

Liégeois-Chauvel C, Peretz I, Babaï M, Laguitton V, Chauvel P. Contribution of different cortical areas in the temporal lobes to music processing. Brain 1998;121:1853–1867. [PubMed: 9798742]

Luck, S. An Introduction to the Event-Related Potential Technique. MIT Press; Cambridge, MA: 2005.

McAdams, S.; Drake, C. Auditory perception and cognition. Stevens' Handbook of Experimental Psychology, Volume 1. In: Yantis, S.; Pashler, H., editors. Sensation and Perception. third edition. Wiley; New York: 2002. p. 397-452.

Näätänen R. The mismatch negativity: a powerful tool for cognitive neuroscience. Ear Hear 1995;16(1): 6–18. [PubMed: 7774770]

Näätänen R, Alho K. Mismatch negativity—A unique measure of sensory processing in audition. Int J Neurosci 1995;80(1–4):317–337. [PubMed: 7775056]

Näätänen R, Winkler I. The concept of auditory stimulus representation in cognitive neuroscience. Psychol Bull 1999;125(6):826–859. [PubMed: 10589304]

Näätänen R, Paavilainen P, Tiitinen H, Jiang D, Alho K. Attention and mismatch negativity. Psychophysiology 1993;30(5):436–450. [PubMed: 8416070]

Näätänen R, Lehtokoski A, Lennes M, Cheour M, Huotilainen M, Iivonen A, Vainio M, Alku P, Ilmoniemi RJ, Luuk A, Allik J, Sinkkonen J, Alho K. Language-specific phoneme representations revealed by electric and magnetic brain responses. Nature 1997;385(6615):432–434. [PubMed: 9009189]

Näätänen R, Pakarinen S, Rinne T, Takegata R. The mismatch negativity (MMN): towards the optimal paradigm. Clin Neurophysiol 2004;115(1):140–144. [PubMed: 14706481]

Navon D. Forest before trees: the precedence of global features in visual perception. Cogn Psychol 1977;9:353–383.

Novitski N, Tervaniemi M, Huotilainen M, Näätänen R. Frequency discrimination at different frequency levels as indexed by electrophysiological and behavioral measures. Cogn Brain Res 2004;20:26–36.

Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 1971;9(1):97–113. [PubMed: 5146491]

Paavilainen P, Alho K, Reinikainen K, Sams M, Näätänen R. Right hemisphere dominance of different mismatch negativities. Electroencephalogr Clin Neurophysiol 1991;78(6):466–479. [PubMed: 1712282]

Paavilainen P, Jaramillo M, Näätänen R, Winkler I. Neuronal populations in the human brain extracting invariant relationships from acoustic variance. Neurosci Lett 1999;265(3):179–182. [PubMed: 10327160]

Peretz I. Shifting ear differences in melody comparison through transposition. Cortex 1987;23:317–323. [PubMed: 3608525]

Peretz I. Processing of local and global musical information by unilateral brain-damaged patients. Brain 1990;113:1185–1205. [PubMed: 2397389]

Poeppel D. The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. Speech Commun 2003;41(1):245–255.

Rinne T, Alho K, Ilmoniemi R, Virtanen J, Näätänen R. Separate time behaviors of the temporal and frontal mismatch negativity sources. Neuroimage 2000;12(1):14–19. [PubMed: 10875898]

Saarinen J, Paavilainen P, Schröger E, Tervaniemi M, Näätänen R. Representation of abstract attributes of auditory stimuli in the human brain. Neuroreport 1992;3(12):1149–1151. [PubMed: 1493229]

Sams M, Paavilainen P, Alho K, Näätänen R. Auditory frequency discrimination and event-related potentials. Electroencephalogr Clin Neurophysiol 1985;62(6):437–448. [PubMed: 2415340]

Sanders LD, Poeppel D. Local and global auditory processing: behavioral and ERP evidence. Neuropsychologia 2006;45(6):1172–1186. [PubMed: 17113115]

Schiavetto A, Cortese F, Alain C. Global and local processing of musical sequences: an event-related brain potential study. Neuroreport 1999;10:2467–2472. [PubMed: 10574353]

Shtyrov Y, Kujala T, Ahveninen J, Tervaniemi M, Alku P, Ilmoniemi RJ, Näätänen R. Background acoustic noise and the hemispheric lateralization of speech processing in the human brain: magnetic mismatch negativity study. Neurosci Lett 1998;251:141–144. [PubMed: 9718994]

Tervaniemi M, Huotilainen M. The promises of change-related brain potentials in cognitive neuroscience of music. Ann N Y Acad Sci 2003;999:29–39. [PubMed: 14681115]

Tervaniemi M, Maury S, Näätänen R. Neural representations of abstract stimulus features in the human brain as reflected by the mismatch negativity. Neuroreport 1994;5:844–846. [PubMed: 8018861]

Tervaniemi M, Lehtokoski A, Sinkkonen J, Virtanen J, Ilmoniemi RJ, Naatanen R. Test-retest reliability of mismatch negativity for duration, frequency and intensity changes. Clin Neurophysiol 1999;110 (8):1388–1393. [PubMed: 10454274]

Trainor LJ, McDonald KL, Alain C. Automatic and controlled processing of melodic contour and interval information measured by electrical brain activity. J Cogn Neurosci 2002;14:430–442. [PubMed: 11970802]

van Zuijen TL, Sussman E, Winkler I, Näätänen R, Tervaniemi M. Grouping of sequential sounds—An event-related potential study comparing musicians and nonmusicians. J Cogn Neurosci 2004;16:331–338. [PubMed: 15068601]

van Zuijen TL, Sussman E, Winkler I, Näätänen R, Tervaniemi M. Auditory organization of sound sequences by a temporal or numerical regularity—A mismatch negativity study comparing musicians and non-musicians. Cogn Brain Res 2005;23:270–276.

Zatorre RJ. Discrimination and recognition of tonal melodies after unilateral cerebral excisions. Neuropsychologia 1985;23:31–41. [PubMed: 3974853]
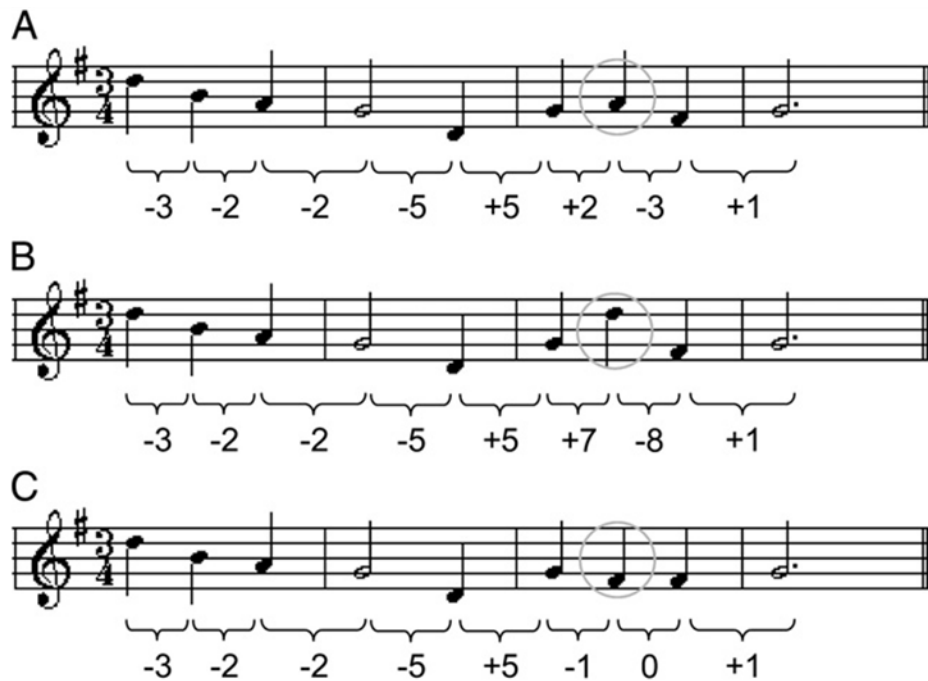
Fig. 1. Examples of *interval-contour* stimuli, illustrating the impossibility of dissociating a contour from an interval change
The seventh note of (A) the standard phrase is changed to create either (B) an interval-changed (contour-preserved) phrase or (C) a contour-changed (and therefore also interval-changed) phrase. Notation below the phrases indicates direction and increment of the change in semitones (one semitone corresponds to a frequency change of $2^{1/12}$). Adapted from Liégeois-Chauvel et al. (1998).
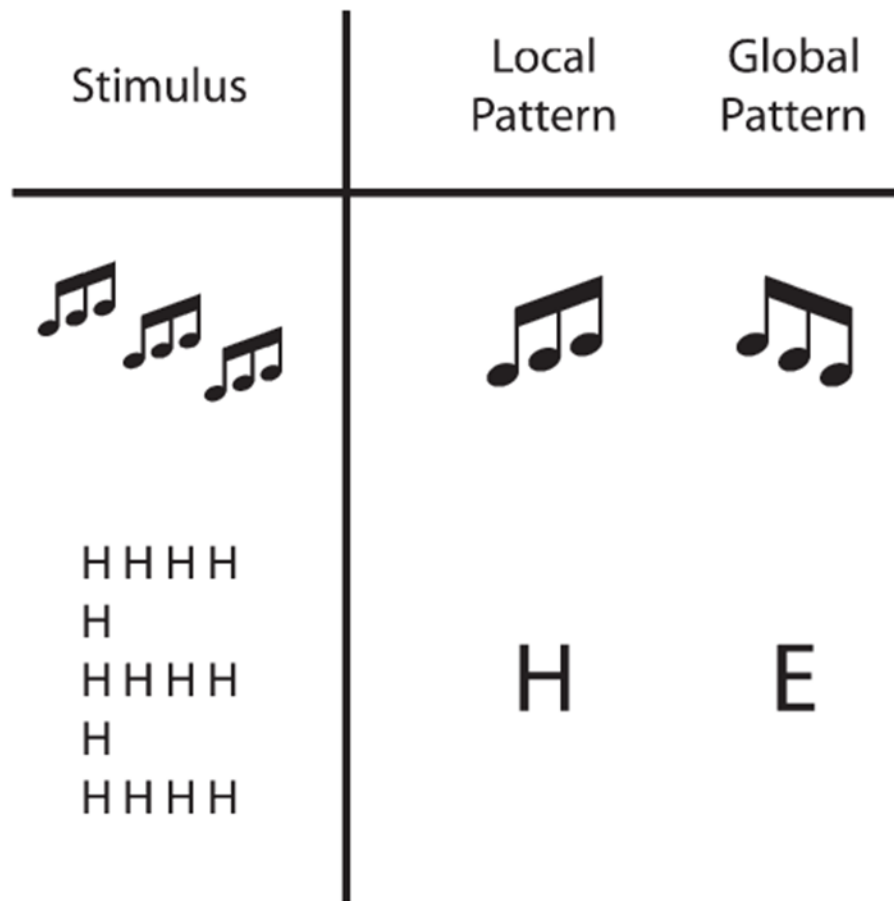
Fig. 2. Illustration of the auditory–visual hierarchical stimulus analogy
In each stimulus, one local pattern (rising–rising for audition, H for vision) is repeated to form the global pattern (falling–falling for audition, E for vision). In audition, the global pattern unfolds over a wider temporal scale, whereas in vision, the global pattern unfolds over a wider spatial scale. Whether in the auditory or visual modality, the local and global levels can be orthogonally manipulated, i.e., the pattern at one level is independent of the pattern at the opposite level.

Fig. 3. The stimulus matrix (visually) depicting the auditory stimuli employed in the current investigations
Patterns were never redundant across local and global levels. For example, local rr patterns were never arranged to form global rr patterns.
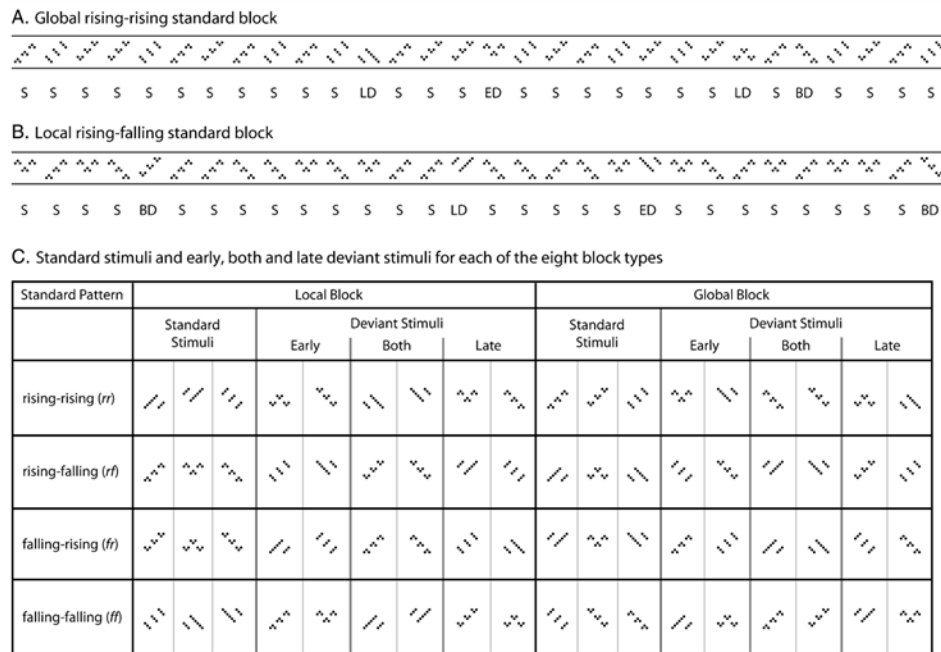
Fig. 4. Stimulus blocks

Thirty-trial sequence examples from (A) a global rr standard block and (B) a local rf standard block. Standard (S) stimuli occurred nine times more often than did deviants (ED = *early* deviant, BD = *both* deviant, LD = *late* deviant). The standard pattern was never presented at the opposite level, e.g., in a global rr block, no local rr patterns were ever presented as part of the standard or deviant stimuli. (A) Global rising–rising standard stimuli are interspersed with falling–rising (ED), falling–falling (BD) and rising–falling (LD) patterns. (B) Local rising–falling standard stimuli are interspersed with falling–falling (ED), falling–rising (BD) and rising–rising (LD) patterns. (C) The full stimulus matrix for each of the eight blocks (two structural levels by four patterns). In each block, three possible standard stimuli and six possible deviant stimuli were presented. Deviant stimuli were divided into three categories: *early*, *both* and *late*. Note that the category to which each of the twelve stimuli is assigned is fully determined by the block context.
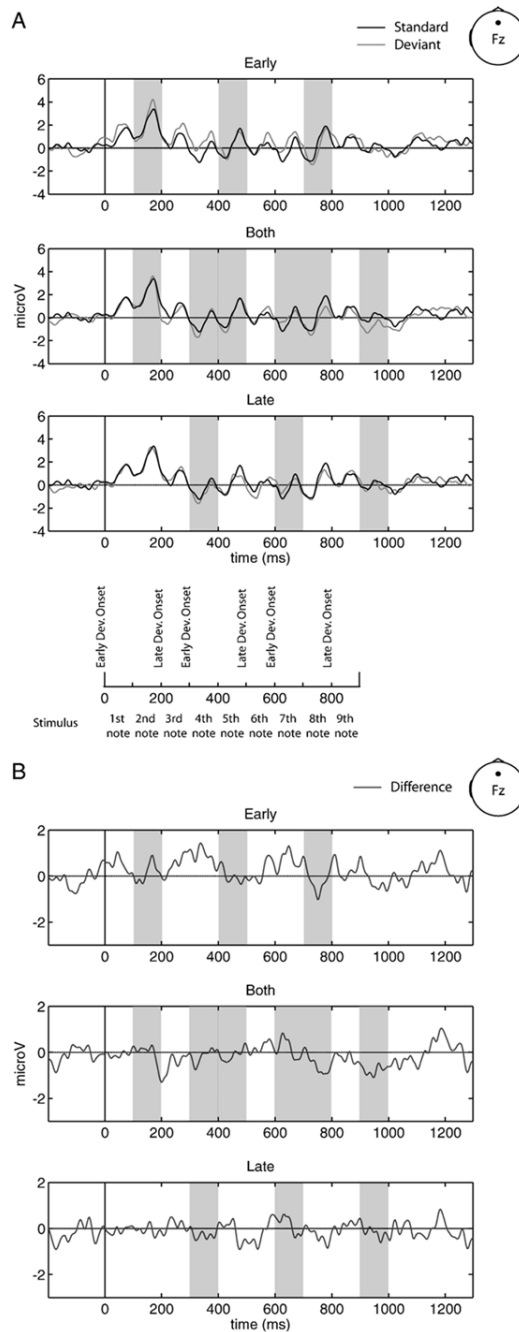
Fig. 5. Local waveforms

(A) ERPs of local standard (black) and *early* (top), *both* (middle) and *late* (bottom) deviant conditions (light gray), recorded at site Fz. Temporal unfolding of stimuli is shown below for reference. (B) Local difference waves (deviant minus standard) for *early* (top), *both* (middle) and *late* (bottom) conditions, recorded at site Fz. Shading indicates the relevant time windows over which statistical comparisons were made. All waveforms are collapsed over pattern.
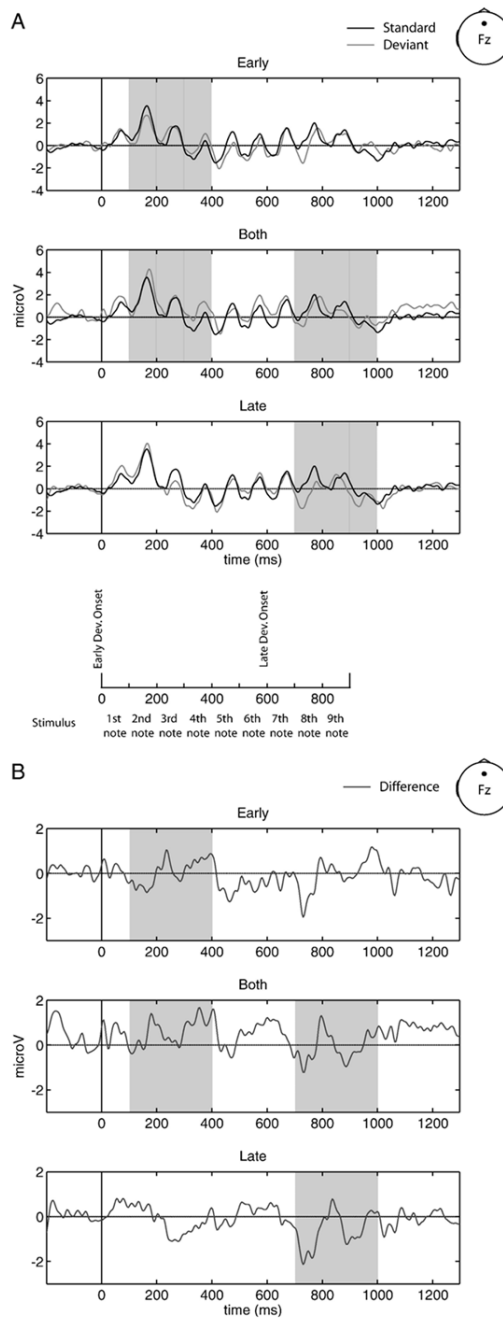
Fig. 6. Global waveforms
(A) ERPs of global standard (black) and *early* (top), *both* (middle) and *late* (bottom) deviant conditions (light gray), recorded at site Fz. Temporal unfolding of stimuli is shown below for reference. (B) Global difference waves (deviant minus standard) for *early* (top), *both* (middle) and *late* (bottom) conditions, recorded at site Fz. Shading indicates the relevant time windows over which statistical comparisons were made. All waveforms are collapsed over pattern.
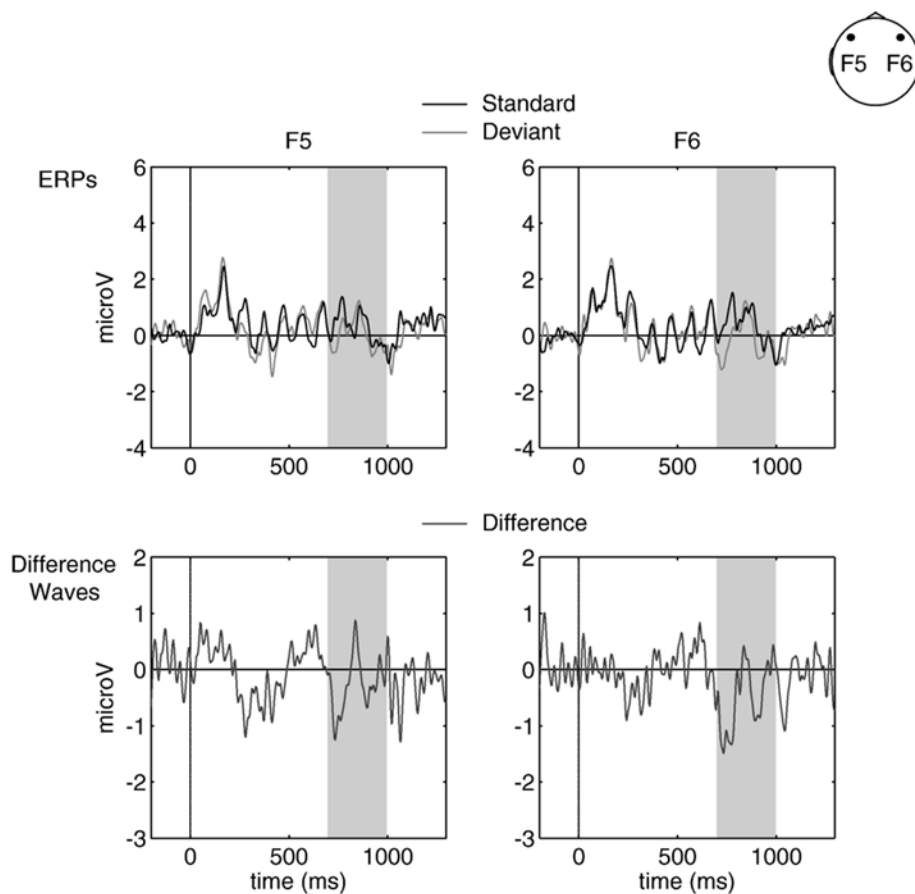
Fig. 7. Laterality
Global *late* deviants vs. standards (top row) and difference waves (bottom row) over left (F5, left column) and right (F6, right column) frontal electrode sites. Standards are shown in black, deviants in light gray and difference waves (deviant minus standard) in dark gray. Shading indicates the relevant time windows over which statistical comparisons were made. All waveforms are collapsed over pattern.
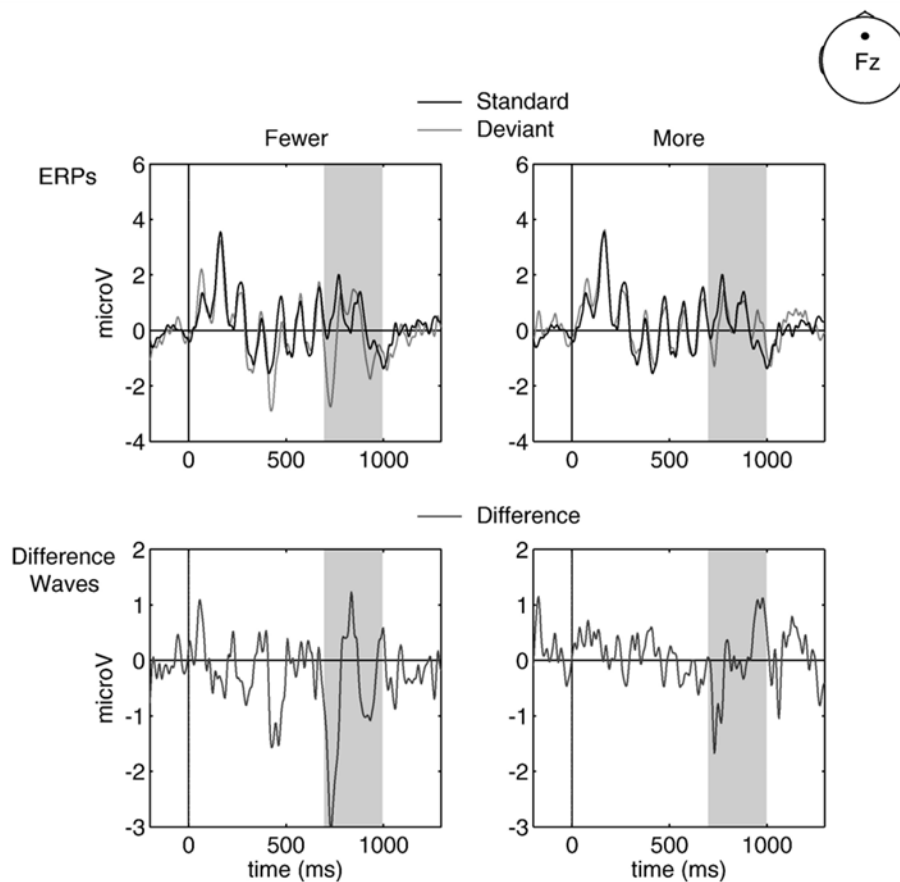
Fig. 8. Frequencies
Global *late* deviants vs. standards (top row) and difference waves (bottom row) for conditions introducing fewer frequencies (left column) and conditions introducing more frequencies at site Fz. Standards are shown in black, deviants in light gray and difference waves (deviant minus standard) in dark gray. Shading indicates the relevant time windows over which statistical comparisons were made. All waveforms are collapsed over pattern.