NIH-PA Author Manuscript

Auditory Scene Analysis: the interaction of stimulation rate and frequency separation on pre-attentive grouping

Pierfilippo De Sanctis[1], **Walter Ritter**[1,2], **Sophie Molholm**[1,2], **Simon P. Kelly**[1], and **John J. Foxe**[1,2]

[1]*The Cognitive Neurophysiology Laboratory, Nathan S. Kline Institute for Psychiatric Research, Program in Cognitive Neuroscience and Schizophrenia, 140 Old Orangeburg Road, Orangeburg, NY 10962, USA*

[2]*Program in Cognitive Neuroscience, Department of Psychology, City College of the City University of New York, New York, NY, USA*

Abstract

Segregation of auditory inputs into meaningful acoustic groups is a key element of auditory scene analysis. Previously, we showed that two interwoven sets of tones differing widely along multiple feature dimensions (duration, pitch and location) were pre-attentively separated into different groups, and that tones separated in this manner did not elicit the mismatch negativity component with respect to each other. Grouping was studied with human subjects using a stimulus rate too slow to induce streaming. Here, we varied the separation of tone sequences along a single feature dimension, i.e. frequency. Frequency differences were either 24 Hz (small) or 1054 Hz (large). Two relatively slow stimulus rates were used (2.7 or 1 tone/s) to explicitly investigate grouping outside the so-called 'streaming effect', which requires rates of about 4 tones/s or faster. Two tones were presented in a quasi-random manner with embedded trains of one to four identical tones in a row. Deviants were defined as frequency switches after trains of four identical tones. Mismatch negativity was only elicited for small frequency switches at the slower stimulation rate. The data indicate that pre-attentive grouping of tones occurred when the frequency difference that separated them was large, regardless of stimulation rate. For small frequency differences, inputs were only grouped separately when the stimulation rate was relatively fast.

Introduction

Our acoustic environment is often noisy and complicated, comprising several or more interwoven sequences of sounds originating from several sources at the same time. To achieve a coherent perception it is essential to segregate the mixture of sounds from each other and group them according to their respective sources, a process termed auditory scene analysis (Bregman, 1990). Neurophysiological evidence using event-related potentials (ERPs) has revealed that auditory scene analysis is in part determined by pre-attentive brain mechanisms in the auditory cortex (Sussman *et al.*, 1999; Ritter *et al.*, 2000, 2006; Gaeta *et al.*, 2001;

Correspondence: Dr John J. Foxe or Walter Ritter, The Cognitive Neurophysiology Laboratory, Nathan S. Kline Institute for Psychiatric Research, Program in Cognitive Neuroscience and Schizophrenia, 140 Old Orangeburg Road, Orangeburg, NY 10962, USA. E-mail: E-mail: foxe@nki.rfmh.org or E-mail: monterey@bcn.net.

Winkler *et al.*, 2001). The mismatch negativity (MMN) is a component of the auditory ERP elicited pre-attentively when one or more changes in previously repeating stimuli are detected (Näätänen, 1992; Ritter *et al.*, 2002). The MMN usually has a peak latency between about 100 and 250 ms and is largest at or near electrode Fz. Evidence that the MMN can be elicited pre-attentively by auditory stimuli is that it can be elicited when subjects ignore the stimuli, such as watching a silent movie as in the present study, when performing a difficult unrelated, visual task (Winkler *et al.*, 2003b), during sleep (Atienza *et al.*, 1997) and in newborn infants (Winkler *et al.*, 2003a).

Several MMN studies have used the streaming effect to investigate auditory scene analysis (Sussman *et al.*, 1999; Yabe *et al.*, 2001). Streaming is a perceptual phenomenon where a rapidly presented sequence of high- and low-frequency tones splits into separate streams as if emitted by different sound sources (Bregman, 1990). However, streaming as a model to investigate auditory scene analysis is limited as the streaming effect is only perceived when stimuli are presented at a rate of 4 Hz or faster (Bregman, 1990). Not all sound sources in the natural environment occur at such high stimulus rates and it is unlikely that grouping of interwoven sounds to their respective sources is limited to such circumstances.

What variables determine grouping of auditory events in an acoustic scene when the presentation rate is low? In a pair of previous studies (Ritter *et al.*, 2000, 2006), closely similar paradigms were used that incorporated the fact that acoustic events typically differ along several feature dimensions simultaneously. The stimulus onset asynchrony (SOA) in both studies was set at 370 ms (~3 Hz), which is outside the typical range used to induce streaming, and subjects reported not experiencing streaming. Two sets of tones were presented that differed in frequency, duration and ear of presentation. In Ritter *et al.* (2000) two sets of tones were alternated between ears. In Ritter *et al.* (2006) tones were presented to the left and right ears in a quasi-random order, where one to four tones in a row of one set (in one ear) were followed by one to four tones of the alternative set (to the other ear). The latter paradigm incorporated two types of deviants. The first was an infrequent duration deviant presented randomly. This deviant was shorter than the standard tones of one ear and longer than the standard tones of the other ear. In this way, this deviant was presumably capable of eliciting an MMN with respect to the tones in its own ear of presentation and also with respect to the tones in the other ear of presentation. The second deviant type occurred naturally in those trials when the ear of presentation switched after a prolonged series of tones had already occurred in one ear (i.e. after four identical tones in a row). In this case, as well as the obvious location change from left to right ear or vice versa, both the frequency and duration of the tone also changed. Under a standard MMN design, such an arrangement, using a deviant fully incorporating three feature changes, would be expected to elicit a robust MMN. The paradigm used in Ritter *et al.* (2006) is presented in Fig. 1.

In fact, Ritter *et al.* (2000, 2006) found that the first kind of deviant just described (duration deviant) elicited MMN only with respect to the set of tones in which it was embedded, those within the same ear of presentation, even though the deviant differed with respect to the standards of the tones in the other ear. On this basis, it was concluded that tones had been separately grouped. The result contrasts with previous studies (Winkler *et al.*, 1996; Sussman *et al.*, 2003), which found that a duration deviant that differed from two standard tones, each with different durations, elicited two MMNs, one with respect to each standard. In the latter studies no attempt was made to study grouping (see Discussion). More importantly for the present study, in addition to finding only one MMN as in Ritter *et al.* (2000), Ritter *et al.* (2006) found that, when grouping occurs, MMN is not elicited by the multifeature deviant on switch trials. This result implied that when two sets of tones have been pre-attentively grouped they do not elicit MMN with respect to each other.

Shinozaki *et al.* (2000) found ERP evidence of grouping of tones based solely on a frequency difference between two sets of tones using an SOA outside the range of streaming (see the Discussion). A variety of studies have shown that frequency differences between tones and stimulus rate both play a role in inducing streaming (Bregman, 1990). In the present experiment, we endeavored to determine whether frequency differences and rate of presentation still play an important role even when the SOAs are well outside the rates used to induce streaming. The design was similar to that of Ritter *et al.* (2006) except that stimuli only differed in frequency, a longer SOA was added (1000 ms) and there were no duration deviants. Based on Ritter *et al.* (2006), the presence and absence of MMN in the switch trials was used as an index of the non-occurrence and occurrence of grouping, respectively.

Materials and methods

Thirteen healthy adults (mean age 26.7 years, SD 4.3, nine females) participated in the experiment. Participants reported that they had normal hearing and had no known neurological deficits. All but two of the 13 participants were right handed as assessed using the inventory of Oldfield (1971). All participants provided written informed consent and the Institutional Review Board at the Nathan S. Kline Institute for Psychiatric Research approved all procedures in accordance with the Declaration of Helsinki.

Pure tones with a frequency of 440, 464 or 1494 Hz and an intensity of 75 dB sound pressure level were presented binaurally via headphones (Sennheiser HD-600). Stimulus duration was 100 ms with rise and fall times of 20 ms. Four separate blocked conditions were run, in which the difference in frequency between the two tones was either 24 or 1054 Hz and SOA was either 370 or 1000 ms. The two tones were delivered equiprobable in a quasi-random manner with one, two, three or four identical tones presented in a row followed by a switch wherein one, two, three or four identical tones of the alternative set were presented. There was a total number of 800 stimuli per run and four runs per condition. Conditions were counterbalanced within participants. Participants ignored the sounds and watched a silent movie.

Brain activity was recorded and analysed using the Neuroscan Synamp I system from 128 tin scalp electrodes (impedances < 5 kΩ), referenced to the nose and digitized at 500 Hz. The data were bandpass filtered from 1.5 to 45 Hz (24 dB/octave). Epochs of 500 ms were extracted with a 100 ms pre-stimulus baseline. Trials in which eye movements were made were rejected off-line on the basis of the horizontal and vertical electro-oculogram. We used an automatic artifact rejection criterion of ± 70 µV for artifact exclusion (e.g. blinks and large muscle activity), applied to all electrodes in the array. In the rare case where recordings at individual electrodes were corrupted or contacts were broken, recordings were interpolated using Brain Electrical Source Analysis software (BESA 5.1.6, MEGIS Software GmbH, Munich, Germany).

A tone was defined as a standard tone when it was preceded by three identical tones in a row and as a 'switch' tone when preceded by four identical tones of the alternative set. Note that it is not necessary to use low-probability deviants as in the oddball paradigm to effectively elicit the MMN. In three previous studies (Sams *et al.*, 1983; Giese-Davis *et al.*, 1993; Molholm *et al.*, 2005), MMNs were obtained simply by varying the sequencing of two equiprobable tones. These studies relied on analysis of local sequences and found that the MMN was elicited when one of the two possible tones followed several instances of the other tone. Hence, both tones elicited the MMN depending on the immediately preceding sequence of tones. The average number of accepted trials for ERPs elicited by a tone preceded by three identical tones in a row (standard) was 222. The average number of accepted trials for ERPs elicited by a tone preceded by four identical tones of the alternative pitch (switch) was 248.

To test for MMNs associated with switch trials, ERPs elicited by switch tones were subtracted from the ERPs elicited by standard tones. Note that tones of the same frequency served as both standards and deviants, so this computation could be performed for identical tones, obviating any possible effects due to simple differences in frequency characteristics. MMN was statistically assessed by applying simple two-tailed *t*-tests in a 50 ms time window centred on the peak latency in the grand mean waveforms and derived from the fronto-central scalp site Fz where MMN is typically found to be of maximal amplitude. These peaks were 109 and 110 ms for the set of tones with large frequency separation and SOAs of 370 and 1000 ms, and 198 ms for the set with small frequency separation with an SOA of 1000 ms.

Switching between the two sets of tones should also elicit an enhanced N1, which is an early sensory component of the ERP. This is because a switch in frequency after repeated stimulus presentation induces a release of refractoriness of N1 generators (Butler, 1968). If an MMN is elicited in the conditions with a large frequency separation, it is possible that the latency of MMN could overlap with N1. To distinguish N1 enhancement from a genuine MMN, we conducted a segmentation analysis (Pascual-Marqui *et al.*, 1995; Foxe *et al.*, 2005) and a topographic ANOVA (TANOVA) (Brandeis & Lehmann, 1989), which enabled us to assess whether N1 scalp topography was altered by an overlapping MMN.

Segmentation analysis and tanova

Segmentation analysis (Pascual-Marqui *et al.*, 1995) aims to identify the most dominant scalp topographies appearing in the group-average ERPs. Here, segmentation analysis was conducted across the entire 400 ms post-stimulus epoch. To compare switch and standard trials, the topographies appearing in the group-averaged ERPs were pooled over trials and time points. Through an iterative process, those topographies are identified for which the global explained variance is at its maximum. From such pattern analysis, it is possible to reduce real-time ERP map series into time segments of stable map topography. Each segment is thought to represent a given 'functional microstate' during information processing (Lehmann & Skrandies, 1980; Michel *et al.*, 1999). Described results are limited to those temporal segments where an enhanced negativity for switch trials was observed (for a detailed description see Pascual-Marqui *et al.*, 1995), i.e. we only considered segmentation results for the N1 and MMN timeframes.

The TANOVA assesses topographical differences statistically. To quantify differences in topography between standard and switch ERPs, we computed the global dissimilarity (Lehmann & Skrandies, 1980). Global dissimilarity is an index of configuration differences between two scalp distributions, independent of their strength as the data were normalized using the global field power. For each subject and time point, the global dissimilarity indexes a single value, which varies between 0 and 2 (0, homogeneity; 2, inversion of topography). To create an empiric probability distribution against which the global dissimilarity can be tested for statistical significance, the Monte Carlo MANOVA was applied (for a more detailed description see Manly, 1991).

Results

Figure 2A shows the grand mean ERPs for the standard and switch tones in the four conditions (red and blue lines, respectively). On the far left, the gray line shows the difference wave obtained by subtracting the ERPs elicited by the standard tones from the ERPs elicited by the switch tones. Enhanced negativities for the switch-tone ERPs compared with the standard-tone ERPs were evident in three of the conditions. An MMN can be seen in the condition with small frequency separation (24 Hz) and an SOA of 1000 ms [left column, Fig. 2A, $t(12) = -5.4$, $P < 0.001$]. For the small frequency separation with an SOA of 370 ms (third column from left), there was no evidence of an enhanced negativity associated with the switch trials in the latency

region of the MMN obtained for the same frequency separation of the previous condition. Thus, it is clear that an MMN was not elicited by the switch tones in this condition. For the two conditions in which the frequency separation was large (1054 Hz), enhanced negativities appeared in the N1 time frame (second and fourth columns from left) that were possibly due to refractory effects expected for large frequency changes [1054 Hz/1000 ms SOA: $t(12) = -5.7$, $P < 0.005$; 1054 Hz/370 ms SOA: $t(12) = -3.1$, $P < 0.05$].

Figure 2B presents the isopotential maps for the conditions where enhanced negativities were obtained for the switch tones. The maps in the upper half were based on the peak latency of N1 observed for the standard tones. The maps in the lower half were based on the switch tones. The far left lower map was based on the MMN and the other two lower maps were based on N1.

As can be seen in the voltage maps for the 24 Hz/1000 ms SOA condition on the far left of Fig. 2B, the MMN elicited by the switch tones shows a more frontal oriented maximum than the earlier latency N1 elicited by the standard tones, a common finding for these components (see Paavilainen *et al.*, 1989,1991;Picton *et al.*, 2000). Four of the six maps look highly similar to the map for N1 on the far left. The map on the upper right does not have a strong focus because the N1 for the standard trials of this condition is of very low amplitude with a somewhat more parietal distribution. A three-way ANOVA was conducted on these six maps, with factors of Condition (24 Hz/1000 ms SOA, 1054 Hz/1000 ms SOA and 1054 Hz/370 ms SOA), Trial (Standard and Switch) and electrode (AFz, Fz, Cz and CPz). Activity at midline electrodes AFz and Fz, and at Cz and CPz was averaged (AFz/Fz vs. Cz/CPz) and defined in terms of region of interest with the levels anterior vs. posterior.

The ANOVA revealed main effects of Condition ($F_{2,24} = 26.2$, $P < 0.001$), Trial ($F_{1,12} = 8.6$, $P = 0.012$) and Region of Interest ($F_{1,12} = 12.9$, $P = 0.004$), and a Condition by Trials by Region of Interest interaction ($F_{2,24} = 3.9$, $P = 0.034$). A protected analysis separately for each condition showed that an interaction between Trial and Region of Interest was only significant in the 24 Hz/1000 ms SOA condition ($F_{1,12} = 8.1$, $P = 0.015$). These results suggest that the MMN was only elicited in the latter condition.

Figure 2C shows the results of the segmentation analysis in the conditions where enhanced negativities were obtained for the switch tones. In all three conditions, the same segments were assigned to the standard and switch tones during the latency window of N1 (gray-colored bars). In the condition where an MMN was observed (far left), different segments were assigned to the standard and switch tones in the latency window of the MMN (orange- and yellow-colored bars, respectively).

Figure 2D presents the results of the TANOVAS. No significant differences in scalp topography between the standard and switch tones were obtained in any of the conditions with enhanced negativity until 140 ms (see color coding in Fig. 2D for *P*-values).

The segmentation analysis and TANOVAS yielded parallel results. For the 24 Hz/1000 ms SOA condition, in the N1 latency window the same segments were assigned to, and TANOVA found no significant difference in topography for, the standard and switch tones. In the MMN latency window, different segments were assigned to, and TANOVA revealed significant differences in topography for, the standard and switch tones. In the N1 latency window in the 1054 Hz/370 ms SOA and 1054 Hz/1000 ms SOA conditions, the same segments were also assigned to, and TANOVA found no significant difference in topography for, the standard and switch tones. Taken together, the data and analyses associated with Fig. 2 support the view that the enhanced negativity obtained in the 24 Hz/1000 ms SOA condition was due to MMN, whereas the enhanced negativities for the switch tones obtained in the 1054 Hz/370 ms SOA and 1054 Hz/1000 ms SOA conditions were due to release from refractoriness of the N1.

It is worth noting that the TANOVA for the 1054 Hz/1000 ms SOA condition did reveal a topographic difference between standard and switch ERP at approximately 140 ms, which could potentially indicate an MMN response. In a follow-up test, we assessed whether this difference in topography was driven by amplitude differences between standard and switch ERP over frontal scalp. A post-hoc t-test revealed no reliable difference in amplitude [$t(12) = 0.667, P = 0.52$] in the time window from 130 to 180 ms. What then is the cause of the TANOVA finding at 140 ms? The TANOVA is sensitive to topographic changes across the entire scalp and, as such, it is equally possible that this effect was driven by changes over more posterior scalp. We inspected topographic maps of the baseline sensory responses (not shown), which indicated that this topographic shift was due to the earlier rise of the second positive (P2) component for the standard trials than for the deviant trials. Follow-up t-tests over more posterior central-parietal scalp sites confirmed that the TANOVA effect was probably a function of this posterior difference.

Discussion

The present study explored conditions under which auditory grouping occurs by varying the separation of tone sequences along only a single feature dimension, i.e. frequency. We used the presence and absence of MMN to indicate whether separation and grouping of an interwoven sequence of tones had occurred. This approach is based on our previous finding (Ritter *et al.*, 2006) where we concluded that pre-attentively grouped tones do not elicit MMN with respect to each other. Our data indicate that grouping occurs pre-attentively when the frequency that separates two sets of tones is large, regardless of the stimulus rate used. However, when the frequency separation is small, inputs will be grouped only if the stimulus rate is relatively high. Our results are in agreement with previous studies (Gaeta *et al.*, 2001; Winkler *et al.*, 2001; Ritter *et al.*, 2006) showing that, once pre-attentive grouping takes place, MMNs are not elicited between groups.

In the Introduction it was pointed out that Winkler *et al.* (1996) and Sussman *et al.* (2003) found that a deviant duration tone elicits two MMNs when it differs from the durations of two standards. In contrast, Ritter *et al.* (2000, 2006) found that a duration deviant delivered to one ear elicited only one MMN, one with respect to the duration of the standards presented to the same ear but no MMN with regard to a different duration of the standards delivered to the other ear. The study of Winkler *et al.* (1996) presented all tones to the right ear and the study of Sussman *et al.* (2003) delivered all tones binaurally. The basic intent was to establish two standard tones of different durations that were alike in all other respects and an infrequent deviant tone that differed only in duration from the standards. In Sussman *et al.* (2003), for example, one standard tone was 50 ms in duration and the other standard tone was 300 ms in duration. The deviant tone was 150 ms in duration. Hence, this was an oddball paradigm that used two standards and one deviant. The deviant duration tone elicited two MMNs with appropriate latencies with respect to the durations of the two standards. In contrast, Ritter *et al.* (2000, 2006) presented one set of tones to one ear with a given standard frequency and duration, and another set of tones to the other ear with a different standard frequency and duration.

The results of the present study agree with another MMN study that employed a rather different approach (Shinozaki *et al.*, 2000). These authors presented a sequence of alternating high- and low-frequency tones to the left ear at a rate (2 Hz) where subjects reported no streaming percept between the two tone sets. One set of tones was composed of 500 Hz standard tones (85%) with 750 Hz deviants (15%). The second set was composed of 1500 Hz (85%) standard tones with 2000 Hz (15%) deviants. Not surprisingly, a clear MMN was obtained to the deviant of each set. They then increased the probability of the deviant in the high-frequency tone set and found that the MMN decreased in size for that deviant. However, the MMN to the low-

frequency deviant remained unchanged. Put simply, the different probabilities of the deviants in the high- and low-frequency tone sets were processed independently, indicating that the two tone sets were pre-attentively grouped. They went on to show that, if the separation in frequency between the two sets was reduced such that their ranges now overlapped, the MMN in the lower frequency tone set became strongly influenced by the higher frequency tone set, suggesting that independent grouping of the tone sets was only maintained when the frequency separation between them was sufficient.

Throughout this study the stimulus presentation rate has been referred to with regard to SOA. Bregman *et al.* (2000) have shown that streaming is determined by the within-stimulus interval, rather than SOA. The within-stimulus interval refers to the interstimulus interval (ISI), i.e. the interval from stimulus offset to stimulus onset, of the tones within a given stream. The longest within-stream ISI used in the study of Bregman *et al.* (2000) was 150 ms. Our tones were 100 ms in duration. In the condition where the SOA was 1000 ms, the ISI between two identical tones in a row was 900 ms. However, as there were one, two, three or four tones of identical frequency in a row followed by a switch to one, two, three or four tones in a row of the other frequency, and these switches continued throughout a run, the overall mean ISI between tones of the same frequency (the within-group ISI) was 2.6 s. This is many times longer than the longest within-stream ISI used by Bergman *et al.* (2000). In the condition where the SOA was 370 ms, the mean within-group ISI was 899 ms. Thus, our data pertain to grouping that occurs within an entirely different ISI range than that previously used.

In conclusion, the separation of sensory input into groups of meaningful acoustic events is a fundamental aspect of auditory scene segregation. We show that pre-attentive grouping of tones outside the stimulus rate used for streaming occurs when the frequency difference that separates them is large, regardless of the stimulus rate that we used. For small frequency differences, inputs are only treated as separate groups when the stimulus rate is relatively fast. Such basic segregation processes are likely to be key in auditory object perception and speech comprehension. An understanding of the factors that drive segregation is important for testing the integrity of these processes in clinical groups that exhibit deficits in basic sensory-cognitive processes. For example, there is already good evidence that the analysis of basic auditory information might be altered in patients with schizophrenia (e.g. Leitman *et al.*, 2007) and autism (Dunn *et al.*, 2008). As such, the present design may be applicable for investigating deficits in basic auditory scene analysis in clinical populations.

Acknowledgments

References

Atienza M, Cantero JL, Gomez CM. The mismatch negativity component reveals the sensory memory during REM sleep in humans. Neurosci Lett 1997;237:21–24. [PubMed: 9406870]

Brandeis D, Lehmann D. Segments of event-related potential map series reveal landscape changes with visual attention and subjective contours. Electroencephalogr Clin Neurophysiol 1989;73:507–519. [PubMed: 2480885]

Bregman, A. Auditory Scene Analysis. MIT Press; Cambridge, MA: 1990.

Bregman AS, Ahad PA, Crum PA, O'Reilly J. Effects of time intervals and tone durations on auditory stream segregation. Percept Psychophys 2000;62:626–636. [PubMed: 10909253]

Butler RA. Effect of changes in stimulus frequency and intensity on habituation of the human vertex potential. J Acoust Soc Am 1968;44:945–950. [PubMed: 5683660]

Dunn MA, Gomes H, Gravel J. Mismatch negativity in children with autism and typical development. J Autism Dev Disord 2008;38:52–71. [PubMed: 17624605]

Foxe JJ, Murray MM, Javitt DC. Filling-in in schizophrenia: a high-density electrical mapping and source-analysis investigation of illusory contour processing. Cereb Cortex 2005;15:1914–1927. [PubMed: 15772373]

Gaeta H, Friedman D, Ritter W, Cheng J. The effect of perceptual grouping on the mismatch negativity. Psychophysiology 2001;38:316–324. [PubMed: 11347876]

Giese-Davis JE, Miller GA, Knight RA. Memory template comparison processes in anhedonia and dysthymia. Psychophysiology 1993;30:646–656. [PubMed: 8248456]

Lehmann D, Skrandies W. Reference-free identification of components of checkerboard-evoked multichannel potential fields. Electroencephalogr Clin Neurophysiol 1980;48:609–621. [PubMed: 6155251]

Leitman DI, Hoptman MJ, Foxe JJ, Saccente E, Wylie GR, Nierenberg J, Jalbrzikowski M, Lim KO, Javitt DC. The neural substrates of impaired prosodic detection in schizophrenia and its sensorial antecedents. Am J Psychiat 2007;164:474–482. [PubMed: 17329473]

Manly, BF. Randomization and the Monte Carlo Methods in Biology. Chapman & Hall; London, UK: 1991.

Michel CM, de Grave PR, Lantz G, Gonzalez AS, Spinelli L, Blanke O, Landis T, Seeck M. Spatiotemporal EEG analysis and distributed source estimation in presurgical epilepsy evaluation. J Clin Neurophysiol 1999;16:239–266. [PubMed: 10426407]

Molholm S, Martinez A, Ritter W, Javitt DC, Foxe JJ. The neural circuitry of pre-attentive auditory change-detection: an fMRI study of pitch and duration mismatch negativity generators. Cereb Cortex 2005;15:545–551. [PubMed: 15342438]

Näätänen, R. Attention and Brain Function. Erlbaum; Hillsdale, NJ: 1992.

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

Paavilainen P, Karlsson ML, Reinikainen K, Naatanen R. Mismatch negativity to change in spatial location of an auditory stimulus. Electroencephalogr Clin Neurophysiol 1989;73:129–141. [PubMed: 2473880]

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

Pascual-Marqui RD, Michel CM, Lehmann D. Segmentation of brain electrical activity into microstates: model estimation and validation. IEEE Trans Biomed Eng 1995;42:658–665. [PubMed: 7622149]

Picton TW, Alain C, Otten L, Ritter W, Achim A. Mismatch negativity: different water in the same river. Audiol Neurootol 2000;5:111–139. [PubMed: 10859408]

Ritter W, Sussman E, Molholm S. Evidence that the mismatch negativity system works on the basis of objects. Neuroreport 2000;11:61–63. [PubMed: 10683830]

Ritter W, Sussman E, Molholm S, Foxe JJ. Memory reactivation or reinstatement and the mismatch negativity. Psychophysiology 2002;39:158–165. [PubMed: 12212664]

Ritter W, De Sanctis P, Molholm S, Javitt DC, Foxe JJ. Preattentively grouped tones do not elicit MMN with respect to each other. Psychophysiology 2006;43:423–430. [PubMed: 16965603]

Sams M, Alho K, Naatanen R. Sequential effects on the ERP in discriminating two stimuli. Biol Psychol 1983;17:41–58. [PubMed: 6626636]

Shinozaki N, Yabe H, Sato Y, Sutoh T, Hiruma T, Nashida T, Kaneko S. Mismatch negativity (MMN) reveals sound grouping in the human brain. Neuroreport 2000;11:1597–1601. [PubMed: 10852208]

Sussman E, Ritter W, Vaughan HG Jr. An investigation of the auditory streaming effect using event-related brain potentials. Psychophysiology 1999;36:22–34. [PubMed: 10098377]

Sussman E, Sheridan K, Kreuzer J, Winkler I. Representation of the standard: stimulus context effects on the process generating the mismatch negativity component of event-related brain potentials. Psychophysiology 2003;40:465–471. [PubMed: 12946119]

Winkler I, Karmos G, Naatanen R. Adaptive modeling of the unattended acoustic environment reflected in the mismatch negativity event-related potential. Brain Res 1996;742:239–252. [PubMed: 9117400]

Winkler I, Schroger E, Cowan N. The role of large-scale memory organization in the mismatch negativity event-related brain potential. J Cogn Neurosci 2001;13:59–71. [PubMed: 11224909]

Winkler I, Kushnerenko E, Horvath J, Ceponiene R, Fellman V, Huotilainen M, Naatanen R, Sussman E. Newborn infants can organize the auditory world. Proc Natl Acad Sci U S A 2003a;100:11 812–11 815.

Winkler I, Sussman E, Tervaniemi M, Horvath J, Ritter W, Naatanen R. Preattentive auditory context effects. Cogn Affect Behav Neurosci 2003b;3:57–77. [PubMed: 12822599]

Yabe H, Winkler I, Czigler I, Koyama S, Kakigi R, Sutoh T, Hiruma T, Kaneko S. Organizing sound sequences in the human brain: the interplay of auditory streaming and temporal integration. Brain Res 2001;897:222–227. [PubMed: 11282382]

Abbreviations

ERP

event-related potential

ISI

interstimulus interval

MMN

mismatch negativity

SOA

stimulus onset asynchrony

TANOVA

topographic ANOVA
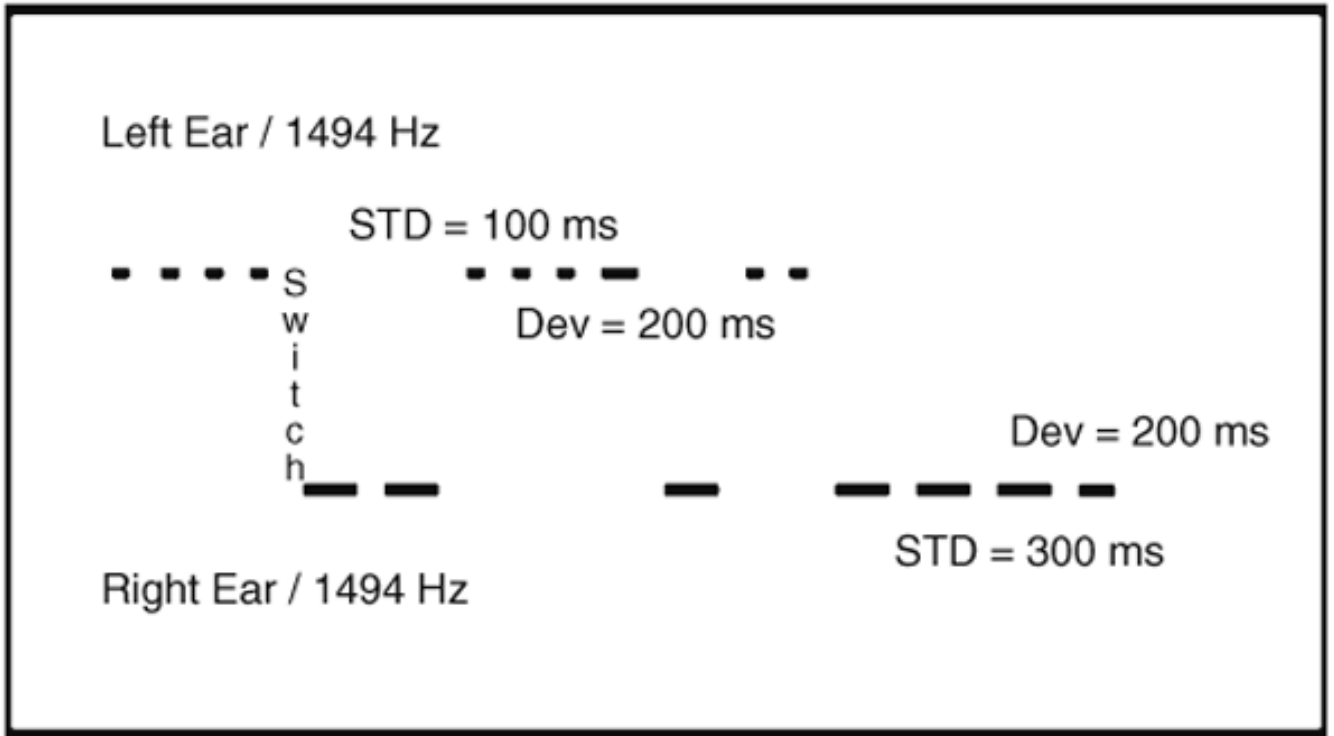
Fig. 1.
Schematic representation of the stimulus design used in Ritter *et al.* (2006). Tones were presented in a quasi-random manner with one to four tones of one set followed by one to four tones of the alternative set. Tones presented to the left ear were 1494 Hz and 100 ms [standard tone (STD)] or 200 ms [duration deviant (Dev)] in duration. Tones presented to the right ear were 440 Hz and 300 ms (STD) or 200 ms (Dev) in duration. A switch deviant was defined as a change of tone delivered to one ear that was preceded by four tones presented to the other ear.
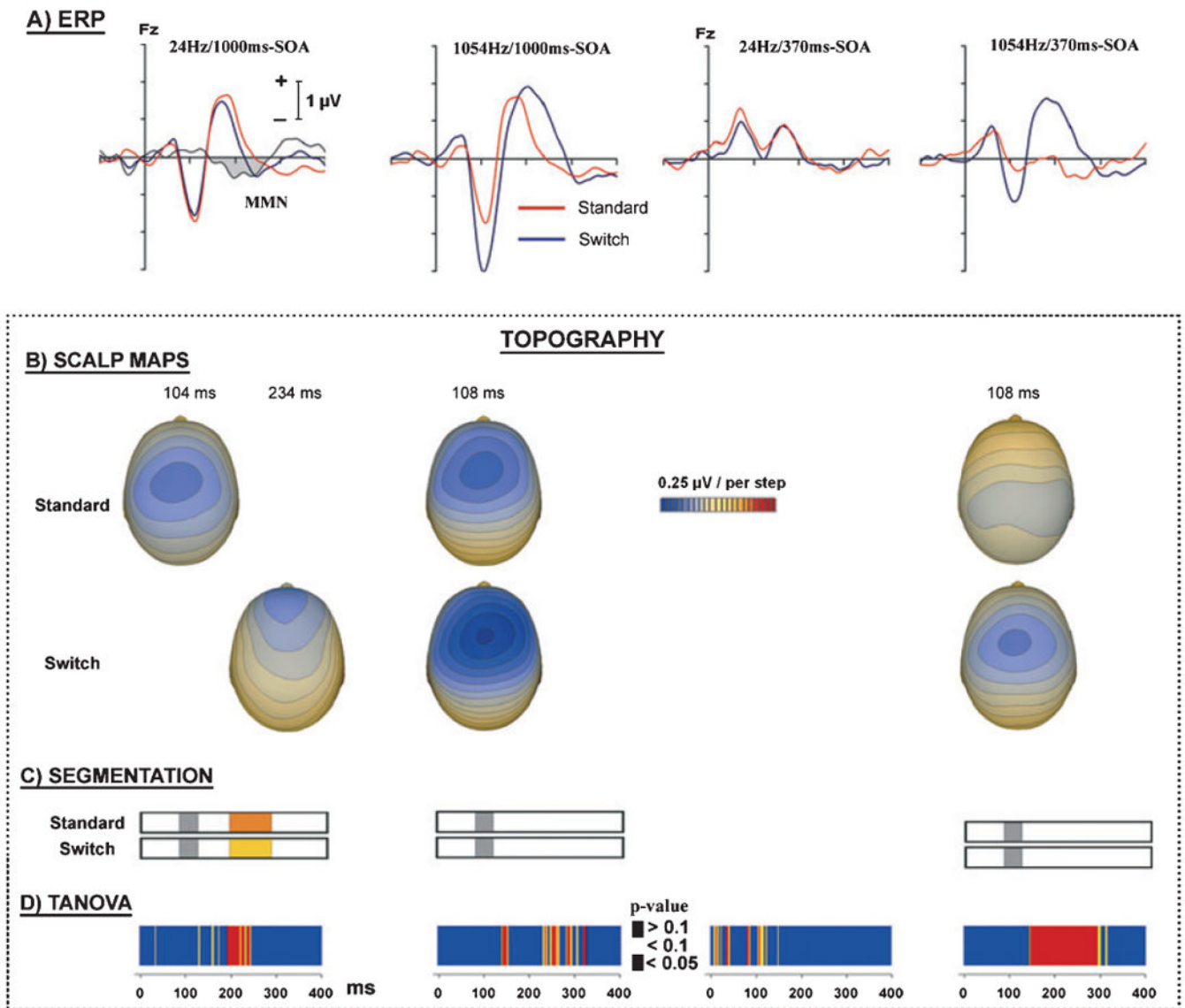
Fig. 2.
(A) Grand-average (*n* = 13) ERPs at electrode Fz elicited by the switch tones (blue lines) and standard tones (red lines). The far left column shows the difference waveform obtained by subtracting the standard from the switch ERPs (gray line). The two columns on the left show ERPs when stimuli were presented at 1000 ms SOA for the small and large frequency separations. The two columns on the right show ERPs when stimuli were presented at 370 ms SOA for small and large frequency separations. (B) Isopotential maps for the standard and switch trials in the conditions that showed enhanced negativities. For the far left, the map for the standard trials is based on the latency of N1 and the map for the switch trials is based on the latency of the MMN. For the other columns, the maps for the standard and switch trials are based on the latency of N1. (C) Segmentation analysis shows the assignment of the most prominent topography to switch and standard tones, which is limited to conditions and time periods where enhanced negativities for switch tones were observed. Only in the condition with small frequency separation and SOA of 1000 ms, in the latency window of the MMN, does the analysis assign two different topographies as the most prominent. (D) TANOVA tested

for topographical differences at each time point with significance marked by the color codes blue/yellow/red.