
An investigation of the auditory streaming effect using event-related brain potentials

ELYSE SUSSMAN,^{a,b} WALTER RITTER,^a AND HERBERT G. VAUGHAN, JR.^a

^aDepartment of Neuroscience, Albert Einstein College of Medicine, Bronx, New York, USA

^bDepartment of Psychology, City University of New York, USA

Abstract

There is uncertainty concerning the extent to which the auditory streaming effect is a function of attentive or preattentive mechanisms. The mismatch negativity (MMN), which indexes preattentive acoustic processing, was used to probe whether the segregation associated with the streaming effect occurs preattentively. In Experiment 1, alternating high and low tones were presented at fast and slow paces while subjects ignored the stimuli. At the slow pace, tones were heard as alternating high and low pitches, and no MMN was elicited. At the fast pace a streaming effect was induced and an MMN was observed for the low stream, indicating a preattentive locus for the streaming effect. The high deviant did not elicit an MMN. MMNs were obtained to both the high and low deviants when the interval between the across-stream deviance was lengthened to more than 250 ms in Experiment 2, indicating that the MMN system is susceptible to processing constraints.

Descriptors: Auditory stream segregation, Mismatch negativity, Event-related potentials, Streaming effect, Auditory sensory memory

Multiple sources of acoustic energy can impinge on the ear constantly in our everyday experience. It is not uncommon for the sounds of voices and office equipment or the ringing of telephones to occur simultaneously. The task of the brain is to tease apart the cacophony of these sounds, forming meaningful representations of the incoming acoustic information. This process requires a mechanism for segregating the inputs into their original sources. A now classic example, the cocktail party phenomenon, illustrates the ease with which the brain is able to perform this task. Amid the steady din of party sounds (e.g., tinkling of glasses, multiple conversations, music), the brain keeps the sources distinct. Auditory cues such as the location of the sound or the pitch of a speaker's voice help this process of segregating the total stream of sound, which has been called auditory stream segregation (Bregman, 1990).

When tones of a sufficient frequency separation are alternated continuously at a fast enough rate for a period of time, a streaming effect occurs (Bregman, 1978, 1990; Bregman & Campbell, 1971). The streaming effect, an aspect of stream segregation, is a percep-

tual phenomenon governed by both the rate of stimulation and the frequency relationship of a tonal sequence. The perception is that the sets of high and low tones split into separate streams of sound, one formed of the high tones and one formed of the low tones. It sounds somewhat like counterpoint in music, as though the two streams are occurring independently and simultaneously. The purpose of this perceptual segregation is presumably to sort the tones in terms of sound sources, thereby improving the ability to perceive patterns within them. Natural differences of acoustic properties emanating from different sound sources are often reflected in pitch (e.g., the voice of a man vs. the voice of a woman). Therefore, within a mixture of sounds striking the ear, it is likely that sounds in a high frequency range will belong to a source separate from that of the sounds in a low frequency range. If streaming occurs for a sequence of high (H) and low (L) tones (e.g., H1, L1, H2, L2, H3, L3, etc.), the ability to identify the order of the tones within a stream is enhanced (such as H1, H2, H3 and L1, L2, L3), whereas the ability to identify the order of the original sequence of tones is largely impossible (viz., H1, L1, H2, L2, H3, L3; Bregman, 1978, 1990; Bregman & Campbell, 1971). That is, the sequence of tones occurring across streams is less easily perceived than the sequence occurring within the streams. The tones appear to belong to either one or the other stream sequence, but not to both. There is a transformation from the original information about all the sounds (the raw input) to information about what is happening with the sounds from given sources.

The main purpose of this study was to investigate where in the stages of auditory processing streaming occurs. There is uncertainty about the mechanisms responsible for the streaming phenomenon. Two explanations have been offered: one based upon

This research was supported by National Institute of Health grants HD01799, NS30029, and DC00223. This paper was part of a dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

We thank Dr. Albert S. Bregman and Dr. István Winkler for their helpful comments and suggestions on an earlier version of this paper.

Address reprint requests to: Dr. Elyse Sussman, Albert Einstein College of Medicine, Kennedy Center, Rm 915, Department of Neuroscience, 1410 Pelham Parkway South, Bronx, NY 10461. E-mail: esussman@balrog.aecom.yu.edu.

selective attention theories and the other based upon automatic processing theories. Jones and her colleagues (Jones, 1976; Jones, Maser, & Kidd, 1978) used Broadbent's (1958) theory of selective attention to explain streaming as a function of attentional mechanisms. Bregman and his colleagues (Bregman, 1990; Bregman & Campbell, 1971) used Neisser's (1967) theory of preattentive acoustic processing mechanisms as governing the formation of streams. That is, there is disagreement about whether attentive or preattentive processes are responsible for the segregation of the tones into streams.

In Broadbent's (1958) research, a series of digits was presented dichotically at slow and fast paces. When the digits were presented slowly, the subjects had no difficulty reporting the order in which the digits were presented. When presented at a fast pace, however, the subjects could no longer report the order of the digits as they were presented and instead reported by ear, first as they were heard in one ear and then in the other. Broadbent reasoned that report by ear was a breakdown of attention. The subject was unable to switch back and forth between the ears fast enough, so first one ear was attended, and then the other. Along similar lines, Jones and her colleagues (Jones, Kidd, & Wetzel, 1981; Jones et al., 1978), interpret the streaming effect in terms of a perceptual overload, the inability to shift attention quickly enough along a multidimensional acoustic space. Thereby, the frequency ranges of the stimuli serve the same function as the ear of input described by Broadbent. The inability to switch attention across large frequency jumps occurring in quick succession results in the segregation of the stimuli into streams (cf. Dowling, 1973; Norman, 1967).

Bregman and Campbell (1971) presented subjects with sequences of high and low tones (at a fast rate), rather than digits, and asked subjects to report their order. Subjects mostly reported the order of tones by pitch similarity: first reporting the order of the high tones, then the order of low tones (or the reverse). Overall, subjects could not report the order of the tones across streams. Conversely, when the same tones were presented at a slower pace, subjects easily reported the order of the tones as they were presented. Bregman and Campbell (1971) interpret their findings in terms of preattentive auditory mechanisms governing the separation of the tones, suggesting that the segregation occurs automatically, early in auditory processing. Specifically, the authors suggested that the organization of tones into streams occurred prior to any conscious selection criteria. One could then attribute the strategy used by Bregman and Campbell's subjects (organization of tones into high and low streams) as a function of the tones having already been sorted into different sources when they were perceived. Because the information perceived concerns sources and not the original sequence, the reports depict information about sources.

It is important to realize that the difference in interpretation of these experiments may suitably describe the differences in the methods and in the data to which they were applied. That is, in the Broadbent experiment, speech stimuli were presented simultaneously to the left and right ears. A strategy of switching back and forth between the ears would break down when the stimuli were coming at a fast rate (two digits per second). In contrast, the stimuli in the Bregman experiment were alternated (binaurally) at a much faster pace, which induced streaming (10 tones per second). Considering these differences in stimulus presentation, using Broadbent's interpretation to explain stream segregation, although appealing, may not be fitting.

The mismatch negativity (MMN), a component of event-related potentials (ERPs) associated with auditory sensory mem-

ory, can be used to determine whether auditory stream segregation is a preattentive process, because the generation of the MMN component is based on an automatic deviance detection system.

The MMN has been used to investigate automatic processing of acoustic input (for reviews, see Näätänen, 1992; Ritter, Deacon, Gomes, Javitt, & Vaughan, 1995). A common way to elicit an MMN is to present an infrequent stimulus (called the deviant) amid a homogeneous series of tones (called the standard). The MMN is generally thought of as the outcome of a change-detection mechanism. That is, the component is the result of the detection of a change from the immediately preceding acoustic inputs. This mechanism was originally explored using simple acoustic feature changes. Recently, the MMN has been used to investigate more complex aspects of sound processing. For example, MMNs have been reliably elicited to changes in the position of tones within a multiple-tone sequence (Näätänen, Schröger, Karakas, Tervaniemi, & Paavilainen, 1993b; Schröger, Näätänen, & Paavilainen, 1992), and to the repetition of a tone within an alternating pattern of two tones (Alain & Woods, 1997; Nordby, Roth, & Pfefferbaum, 1988). MMNs have also been obtained to changes in the abstract qualities of tone pairs (e.g., rising/falling; Paavilainen, Saarinen, Tervaniemi, & Näätänen, 1995; Saarinen, Paavilainen, Schröger, Tervaniemi, & Näätänen, 1992). The direction of tonal frequency (e.g., rising in pitch) from the first of the tone pair to the second was held constant (the standard), while the absolute tonal frequency of the tones varied from trial to trial. A change in direction between the first and the second tone of the pair (e.g., falling in pitch) produced an MMN.

The detection of the deviant is associated with an MMN, which is maximal over the frontocentral regions of the scalp and usually peaks around 140–220 ms from stimulus onset. The amplitude of the MMN response is related to the magnitude of the stimulus deviance in a direct relationship (the larger the difference between the standard and the deviant, the larger the amplitude of the response).

The MMN indexes early, automatic processing of auditory input, independent of subject attention (i.e., attention is not required to elicit the response; Näätänen, 1992). The MMN component is considered an automatic response to deviations occurring within a repetitive acoustic environment because it has been elicited when subjects' attention is highly focused on other tasks, such as reading a book or doing a demanding visual task (see Näätänen, 1992, for a review). Further, the amplitude of the MMN is similar when subjects attend to stimuli and when they ignore (e.g., Näätänen, Paavilainen, Tiitinen, Jiang, & Alho, 1993a; Novak, Ritter, Vaughan, & Wiznitzer, 1990). However, because attention can modulate the amplitude of the MMN, under certain circumstances (Alain & Woods, 1997; Alho, Woods, Algazi, & Näätänen, 1992; Näätänen et al., 1993a; Woldorff, Hackley, & Hillyard, 1991) generation of the MMN probably is not solely automatic. Although the MMN amplitude can be modulated by attention, the MMN component still reflects automatic, preattentive processing. A process that can be influenced by attention does not preclude that it is preattentive in nature (Schneider & Shiffrin, 1977). An example of how processes affected by attention can still occur automatically is seen in the procedure of learning to drive a car. We have all, as adult drivers, had the experience of looking up and realizing that we have driven for a period of time, making all the appropriate turns or stops, and yet cannot recall the driving experience up to that point because our attention was so totally directed on something other than our driving. The mechanics of driving can be automatic. Suppose then, we have only been driving cars with automatic

transmissions and now decide that it is time to drive a fast sports car that has a standard transmission. Eventually, the new skill required for driving a standard transmission car will also become automatic. That is, with focused attention, this new skill can be learned. Attention can modify a process that was automatic, which can then become automatic again.

Recent evidence that the MMN is based primarily upon preattentive processing mechanisms was provided in a study by Alho and Sinervo (1997). In this study, tonal sequences that consisted of nine successive 50 ms tonal elements of different pitches with no silent intervals between them were presented in a dichotic listening paradigm. The left- and right-ear standard patterns were different sequences, but the same standard sequence of pitches was always delivered to the respective ear. Subjects selectively attended one ear and pressed a key every time they heard either of two deviant nine-tone patterns in that ear. Deviant sequences contained a change of pitch in either the third or seventh position within the nine-tone patterns. The level of difficulty for detecting the deviant patterns was high, as indicated by the 50% average hit rate obtained by subjects. Despite the high level of demand for detecting deviants in the attended channel, MMNs were elicited by the deviant complex tone patterns in the unattended ear, suggesting that processing of the complex sequences was automatic. Even stronger evidence of automatic processing is the MMN reported for the missed deviants on the attended channel. Because subjects did not consciously indicate those stimuli as targets and they were unaccompanied by any attention-related ERP components, the processing associated with the MMN obtained for the missed deviants was most likely based on preattentive mechanisms. Even though subjects did not consciously detect some of the targets, the brain processed them as deviants. These data suggest that even if the amplitude of the MMN can be influenced by attention, the system underlying the generation of the MMN component largely reflects automatic auditory processing, even when subjects attend to the tones (see also Sussman, Ritter, & Vaughan, 1998).

In the current study, a sequence of six different high and low tones was presented at both slow and fast paces. The slow-paced sequence (750 ms stimulus onset asynchrony [SOA]) was heard as alternating high and low pitches. The fast-paced sequence (100 ms SOA) created a streaming effect in which a standard sequence of three tones emerged separately in each stream (e.g., L1, L2, L3), with a deviant sequence of three tones occurring infrequently within each stream (e.g., L3, L2, L1). There is a distinct difference between the order of the tones that enter the system (the alternating sequence of high and low tones; Figure 1a) and the perception of the tones as segregated into separate sequences of high and low tones (see Figure 1b). This difference was exploited to assess where in the system the segregation of the tones occurs.

One purpose of this study was to ascertain the locus of the streaming effect based on the contradictory theories outlined above. If Bregman and his colleagues were correct that the streaming effect occurs preattentively, then an MMN should occur in the fast-paced alternating condition. However, if Jones and her colleagues were correct that the streaming effect occurs postattentively, then no MMN should occur.

We think that the purpose for the mechanisms that underlie the streaming effect is to restore the separate auditory events in the environment and that these events are "assigned by our brains to distinct mental entities" (Bregman, 1990, p. 11). We hypothesized, therefore, that the tones presented at a fast rate would segregate to low- and high-tone sequences preattentively and that the memory underlying the MMN system would maintain the two sequences

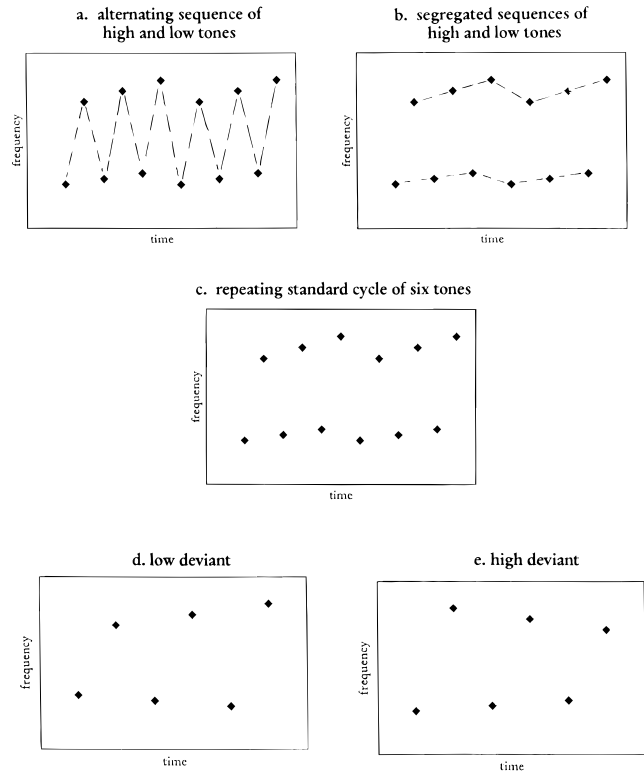


Figure 1. The difference between the order of the tones that enter the system (a) and the segregation of the tones into high and low streams (b); a repeating standard cycle of six tones (c), and deviant stimulus sequences, one cycle of six tones for each (d, e). Time and frequency scales are not precisely depicted.

independently and separately. We therefore expected that two MMNs would be obtained, one generated by the low-tone sequence and one generated by the high-tone sequence. Further, we expected that no MMN would occur when the sequence was presented at a slow pace because the alternation of the high and low tones would interfere with the detection of the standard sequence of tones within the high- and low-pitched sequences. This prediction was based on results of behavioral studies in which slow-paced alternating high and low tones were perceived as a single stream of sound that jumped up and down in pitch (Bregman, 1990).

Two control conditions were run; one for the slow- and one for the fast-paced sequence. One set of tones (the series of low tones) was presented alone in each. The purpose of the fast-paced control was to determine whether isochronous three-tone standard and deviant sequences could elicit an MMN. Because this particular experimental paradigm has not been used before, the absence of an MMN in the fast-paced alternating condition would have been uninterpretable. Therefore, it was necessary to establish that an isochronous three-tone standard and deviant sequence could elicit an MMN (i.e., the low-tone stream presented alone). The purpose of the slow-paced control was to determine whether the 1.5-s pace of the low tones, as they occurred in the slow-paced alternating condition, could elicit an MMN. Otherwise, we could not explain the absence of an MMN for the slow-paced alternating condition. The memory that underlies the MMN system has been estimated to last about 10 s (Cowan, Winkler, Teder, & Näätänen, 1993; Sams, Hari, Rif, & Knuutila, 1993). MMN experiments using frequency

deviants have indicated that at least two standards in a row are necessary to elicit an MMN (Winkler, Cowan, Csépe, Czigler, & Näätänen, 1996). It is not known, however, how many standards in a row are needed when the standard consists of several tones. If three sequences are needed, for example, the time to deliver them in this paradigm would exceed the estimated duration of the underlying memory. Therefore, it was necessary to establish that an MMN could be elicited with the slow-paced sequence.

EXPERIMENT 1

Methods

Subjects

Ten subjects (9 women, 1 man) between the ages of 23 and 42 years with reportedly normal hearing were paid for their participation in the experiment. The control/1.5-s condition was run separately in a later session. Nine subjects (4 women, 5 men) between 24 and 47 years of age participated in the control/1.5-s condition (five of whom participated in the other conditions).

Experimental Procedure

Subjects were seated in a comfortable chair and instructed to ignore the stimuli by reading a book of their choice during the presentation of all conditions. The stimuli were six pure tones (400, 450, 500, 1150, 1250, and 1350 Hz) presented binaurally with insert earphones. Each tone was 50 ms in duration (rise and fall time = 7.50 ms) and had an intensity of 75 dB SPL.

Two experimental conditions (alternating/100 ms and alternating/750 ms) and two control conditions (control/200 ms and control/1.5 s) were used. Two hundred deviants were collected for each stimulus type in eight runs of stimuli presented for each condition. The order of the runs was counterbalanced across subjects, except for the control/1.5-s condition (because it was run at a different time). Subjects took short breaks at approximately one-third and two-thirds of the way through the recording session.

In the alternating/100-ms condition, high (≥ 1150 Hz)- and low (≤ 500 Hz)-frequency tones were alternated at a constant SOA of 100 ms. The standard consisted of a cycle of six tones that, in addition to alternating, rose in frequency from the first to the third tone within each frequency set (e.g., L1, H1, L2, H2, L3, H3, where L1 equals 400 Hz and H1 equals 1150 Hz). Alternating high- and low-frequency tones at the rate of 10 stimuli/s creates a streaming effect (see Bregman, 1990). The stimulus sequence was organized so that when the streams segregated, a three-tone sequence of standards (occurring nonrandomly on 84% of the trials) emerged separately in each stream (e.g., H1, H2, H3; L1, L2, L3). Likewise, a three-tone deviant sequence (occurring nonrandomly on 16% of the trials) occurred nonrandomly in each stream (e.g., L3, L2, L1; H3, H2, H1). Half of the deviants occurred in the low stream and half occurred in the high stream. The occurrence of the deviant sequences in the high or low streams was offset by one cycle so that the low- and high-tone deviants did not occur within the same cycle of six tones. The standard and deviant sequences are presented in Figure 1 (c–e). The low-tone deviant always preceded the high-tone deviant (as shown in Figures 1d and e).

In the alternating/750-ms condition, the same sequence of alternating high and low tones was presented at a slower rate of stimulation (750 ms SOA). When tones are presented at this slower pace, they are heard as alternating high and low pitches.

In the alternating/100-ms condition, the stimulus parameters were selected to produce the strongest streaming effect, as has

been reported from numerous behavioral studies (Bregman, 1990). In the alternating/750-ms condition, the repetition rate of the tones was slowed from the streaming pace to a pace that would be heard as a single stream of sound that alternated in pitch.

For both control conditions, the low tones only (≤ 500 Hz) were used. The order of the tones was kept the same as that of the low-frequency tones in the experimental conditions (see Figure 1, bottom). In the alternating/100-ms condition, a low-frequency tone occurred every 200 ms. Therefore, a 200-ms SOA was used in the control/200-ms condition to replicate the stimulus rate of the low-frequency tones of the alternating/100-ms condition. In the alternating/750-ms condition, a low-frequency tone occurred every 1.5 s. Therefore, a 1.5-s SOA was used in the control/1.5-s condition to replicate the stimulus rate of the low-frequency tones of the alternating/750-ms condition. The standard sequence was L1-L2-L3. The deviant sequence was L3-L2-L1. The probability that a deviant pattern would occur was 16%.

A separate set of data collected in our laboratory as pilot data for another experiment is pertinent to the current study. Eight adult subjects (6 women, 2 men) with normal hearing and between the ages of 25 and 40 years participated in the pilot study. These data were collected using the same stimulus parameters as used in the current experiment except the pace was 1 tone/s in the control condition and 1 tone/500 ms in the alternating condition. Subjects ignored the stimuli and read material of their choice during the recording of all the runs. The data from the control condition will be referred to here as the control/1-s condition.

Recording

The electrical brain activity was recorded using DC-coupled amplifiers, with a low-pass filter setting of 40 Hz. The digitization rate was 400 Hz. An epoch duration of 600 ms was used, which includes a 100-ms prestimulus baseline. Electrode recordings were obtained at the following sites: Fpz, Fz, Cz, Pz, Oz, Fp1, Fp2, F3, F4, F7, F8, FC1, FC2, P3, P4, LM, and RM (left and right mastoids, respectively). All recordings were referenced to the nose. Vertical ocular potentials were monitored with a bipolar electrode configuration using Fp1 and an external electrode placed below the left eye. Horizontal eye movements were monitored using electrodes F7 and F8. Trials on which electrical activity exceeded $\pm 100 \mu\text{V}$ were automatically rejected. The remaining averaged ERPs were examined for residual artifact.

Data Analysis

ERPs elicited by the first tone of each standard sequence were averaged together across the eight runs for each subject, separately for the high and low tones, in each condition. Likewise, the ERPs elicited by the first tone of the deviant sequence were averaged together across the eight runs separately for each subject and each set of tones in each condition.

The grand mean ERPs were used for the purposes of display. Grand mean difference waveforms were calculated by subtracting the ERPs from the standard from those of the deviant, separately for each set of high and low tones and each condition. The peak latency of the MMN was selected in the grand mean difference waveforms. A latency window from 25 ms before to 25 ms after the peak latency of the MMN in the grand means was used to measure the amplitude of the ERPs elicited by the standards and deviants for each subject and each condition. Peak latency was selected as 174 ms in the alternating/100-ms condition, 188 ms in the control/200-ms condition, and 260 ms in the control/1.5-s condition. There was no evidence of an MMN in the alternating/

750-ms condition to either the high- or low-frequency tones nor in the alternating/100-ms condition to the high-frequency tones.

Where there was evidence of an MMN, the data were statistically analyzed using a two-way analysis of variance (ANOVA) for repeated measures with factors of stimulus type and electrode. The MMN component was measured relative to the 100-ms prestimulus activity. The mean voltages in the 50-ms window around the peak of the MMN were used to determine whether the ERPs to the standard and deviant differed significantly at Fz, FC1, FC2, LM, and RM. Because there was no evidence of an MMN in the high tones of the alternating/100-ms condition, the mean voltages in the 50-ms latency window around the peak of the MMN as obtained in the grand mean for the low tones was used to measure the difference between the standard and deviant ERPs. Tukey HSD (honestly significant difference) post hoc comparisons were then used to determine statistical significance at individual electrode sites.

In the control/1-s and control/1.5-s conditions, presence of the MMN was measured for individual subjects. Separate two-way repeated-measures ANOVAs with factors of stimulus type and electrode were calculated for each subject using the number of runs as entities, comparing the standard and deviant ERPs measured in the latency range of the MMN.

To compare scalp distributions of the MMNs obtained in the control and experimental conditions, the data were scaled (McCarthy & Wood, 1985) and then evaluated using condition and electrode in a two-way repeated-measures ANOVA.

To assess whether the amplitude of the MMN varied as a function of SOA, a one-way repeated measures ANOVA, with a main factor of condition, was conducted on all of the subjects that participated in the three control conditions. In each of the three conditions, the MMN component was measured relative to the 100-ms prestimulus activity, and the mean voltages were calculated in the 50-ms window taken around the peak of the MMN. The peak was determined separately in each condition as the most negative peak in the grand average, with all subjects included (188 ms for the control/200-ms condition, 200 ms for the control/1-s condition, and 260 ms for the control/1.5-s condition).

Greenhouse–Geisser procedures were used as appropriate. An alpha level of .05 was used.

Results

Subject Report

At the end of the recording session, subjects were asked their subjective experience of the tones occurring in the alternating conditions. All subjects reported hearing two different paces of tones, one fast and one slow. The slow-paced tones were reported as alternating high and low pitches. The fast-paced sequence was reported as two parallel melodies. Some of the subjects experienced the sequences as dichotic. That is, the streaming effect was so strong it sounded as if the low-tone melody was presented to one ear and the high-tone melody was presented to the other.

Control/200-ms Condition

The low-frequency tones were presented at a rate of one tone per 200 ms. Figure 2 (upper left) presents the across-subjects averages of the ERPs to the standard and deviant stimuli at seven recording sites. The N1-P2 components can be seen, although they are not distinctive because the interval between the tones was short (Javitt, Doneshka, Zylberman, Ritter, & Vaughan, 1993). The N1-P2 components elicited by subsequent stimuli can also be seen because the

epoch extends 500 ms after stimulus onset. A broad negative deflection separating the deviant from the standard beginning about 140 ms represents the MMN. Figure 3 (upper left) presents the difference waveforms obtained by subtracting the ERPs to the standard from the ERPs to the deviant. The first negative peak, seen at Fz, Cz, FC1, and FC2, delineates the MMN component. The presence of the MMN was established by an overall significant difference between the standard and deviant waveforms revealed in a two-way repeated-measures ANOVA, $F(1,9) = 14.9$, $p = .004$. Tukey HSD post hoc comparisons confirmed the difference of stimulus type at the frontal electrodes and the right mastoid at the .01 level but not at the left mastoid site. Table 1 presents the grand mean amplitudes of the standard and deviant ERPs measured in the latency range of the MMN. The mean amplitudes of the MMN component obtained in the same latency window are also provided. These data show that an MMN can be obtained using an isochronous three-tone standard.

Alternating/100-ms Condition

In this condition, the high and low tones were alternated at a rate of one every 100 ms, inducing a streaming effect. The ERPs elicited by the high and low tones were analyzed separately and are presented in Figure 4 (left column). The N1-P2 components from subsequent stimuli can also be seen in this epoch period. In the low stream, a negative deflection separating the deviant from the standard waveforms can be seen beginning about 140 ms and represents the MMN. No MMN was elicited by the deviant within the high-tone sequence, $F(1,9) = 0.45$, $p = .52$.

Figure 4 also presents the grand mean difference waveforms (right column) obtained by subtracting the ERPs to the standard from the ERPs to the deviant separately for each stream. An MMN was elicited for the low stream, seen most prominently in the frontal electrodes. The negative deflection peaking at about 175 ms delineates the MMN component. The MMN obtained to the low deviant in the alternating/100-ms condition is similar in amplitude and latency to the MMN elicited in the control/200-ms condition (see Figure 3). However, the inversion at the mastoids, typically accompanying the MMN component, is attenuated in the alternating/100-ms condition compared with the inversion obtained in the control/200-ms condition. The presence of the MMN was established by the overall significant difference found between the ERPs elicited by the standard and deviant waveforms (measured in the latency range of the MMN) in a two-way repeated measures ANOVA, $F(1,9) = 22.5$, $p = .001$. Tukey HSD post hoc comparisons confirmed the difference at the frontal electrode sites, where the MMN is most prominently seen, but not at the mastoid sites. Table 1 presents the grand mean amplitudes of the standard and deviant ERPs obtained in the latency range of the MMN. MMN mean amplitudes are also provided in Table 1. These data show that an MMN can be elicited when two streams are presented, but only in one stream.

To ascertain whether topographical differences existed between the control/200-ms and the alternating/100-ms conditions, potentially accounting for the differential effects found at the mastoids, the data were scaled and a repeated-measures ANOVA with variables of condition and electrode was conducted on the scaled data. The results of this analysis revealed a significant interaction between condition and electrode, $F(2, 18) = 4.8$, $p = .03$, $\epsilon = 0.7659$.

Control/1.5-s Condition

As can be seen in Figure 2 (upper right), the ERPs elicited by the standard tones contained a negative component (N1) that peaked at

Control Conditions

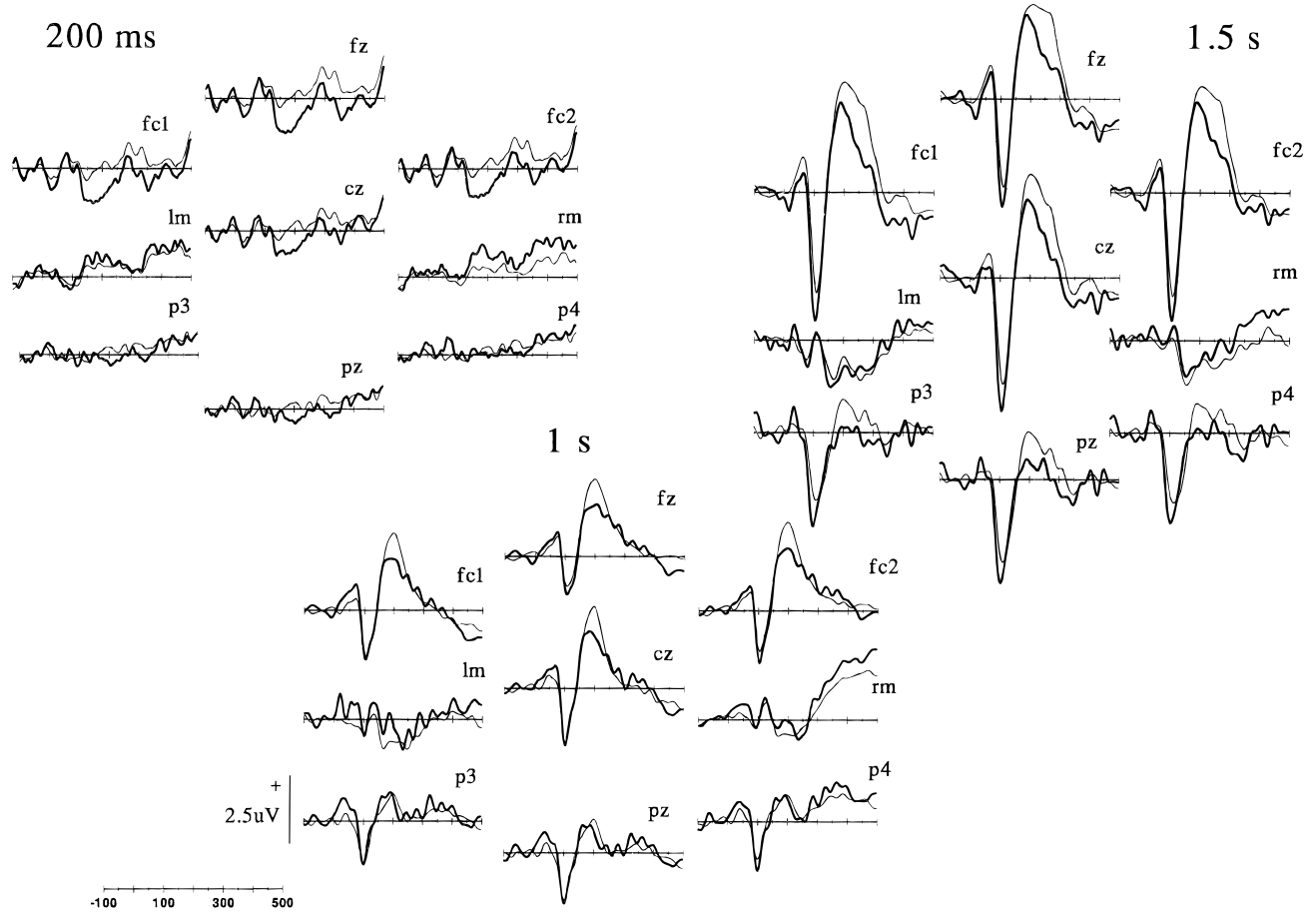


Figure 2. Standard (thin line) and deviant (thick line) ERPs at Fz, Cz, Pz, FC1, FC2, P3, P4, LM, and RM in three conditions when low tones were presented alone (control/200 ms, control/1.5 s, and control/1 s). For the control/1.5-s and the control/1-s conditions, only the four subjects in each condition that had MMNs were included in the display.

100 ms and was largest at Cz. A positive component (P2) peaked around 200 ms and was also largest at Cz. The N1-P2 response was prominent because the interstimulus interval was sufficient (1.5 s).

The difference between the amplitude of the standard and that of the deviant ERPs in the region of the MMN, when all nine subjects were included, was not significant. When each subject was examined individually, four of the nine subjects had MMNs. Two-way repeated-measures ANOVAs revealed the presence of the MMN for each individual where an MMN was visibly detected: Subject 1 (S1): $F(1,9) = 7.1, p = .026$; S2: $F(1,9) = 7.2, p = .025$; S3: $F(1,8) = 33.1, p < .001$; S4: $F(1,9) = 8.2, p = .019$. Figure 3 (upper right) shows the difference waveforms only for the group of four subjects who had MMNs.

Alternating/750-ms Condition

In this condition, the ERPs elicited by the standard tones contained a negative component (N1) that peaked at 100 ms and was largest at Cz. A positive component (P2) peaked around 200 ms and was also largest at Cz. The N1-P2 response was prominent because the interstimulus interval was sufficient (750 ms). As can be seen in Figure 5, for the low tones, the response to the standard tones was separated from that to the deviant tones in a negative direction (left

column), creating a positive peak in the subtraction waveforms (right column). Likewise, for the high tones, there was no indication of an MMN.

Similar to the control/1.5-s condition, half of the eight subjects in the control/1-s condition showed MMNs to the deviant sequence and half did not. Figures 2 and 3 (center bottom) display only the four subjects who had MMNs. The difference between the average amplitude of the standard and that of the deviant ERPs in the region of the MMN when all eight subjects were included was not significant. However, the individual analyses conducted on the four subjects established that the ERPs to the standard were significantly different from the ERPs to the deviant, in the range of the MMN: S1: $F(1,6) = 8.7, p < .03$; S2: $F(1,6) = 25.4, p < .02$; S3: $F(1,6) = 31.2, p < .01$; S4: $F(1,6) = 8.3, p < .03$. None of the subjects who obtained MMNs in the control had MMNs when the tones were alternating high and low tones, as was also demonstrated when the tones were presented at the slower pace of the main experiment.

An ANOVA for repeated measures with variables of condition and electrode was conducted on the scaled data of the three control conditions (control/200 ms, control 1.5 s, and control 1 s). The absence of an interaction, $F(10,30) < 1, p = .572, \epsilon = 0.144$, in

Control Conditions

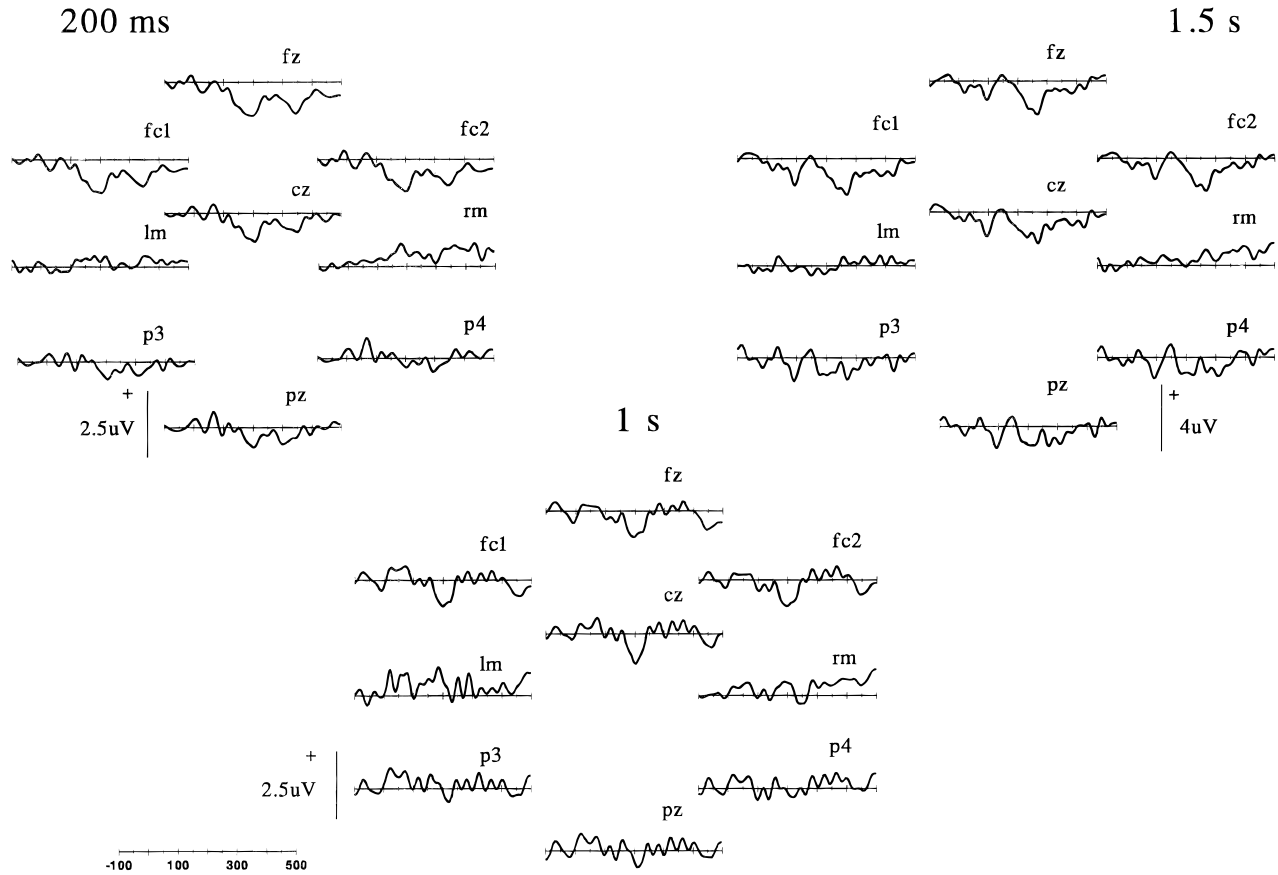


Figure 3. Difference waves at Fz, Cz, Pz, FC1, FC2, P3, P4, LM, and RM in three conditions when low tones were presented alone (control/200 ms, control/1.5 s, and control 1 s). For the control/1.5-s and the control/1-s conditions, only the four subjects in each of the conditions that had MMNs were included in the display. Note the difference in scale for the control/1.5-s condition.

this analysis established that there were no significant differences in scalp distribution among the MMNs obtained in each of the three conditions.

When the amplitude of the MMN was compared across the three control conditions, amplitude did not vary as a function of SOA, $F(2,18) < 1$, $p > .35$, $\epsilon = 0.782$.

Discussion

An MMN can be obtained to a three-tone standard and deviant sequence presented isochronously (the low tones only in the control/200-ms condition). This result extends the findings of Schröger, Paavilainen, and Näätänen (1994), who obtained MMNs to changes

Table 1. Amplitude of the Difference Waveform and the Standard and Deviant ERPs Measured on the Grand Means in the Latency Window of the MMN for Experiment 1

Condition	Electrode	Standard (μV)	Deviant (μV)	Difference (μV)	p
Alternating/100 ms (low tones)	Fz	0.50 (0.57)	-0.54 (0.78)	-1.04 (0.59)	**
	FC1	0.38 (0.74)	0.48 (0.91)	-0.86 (0.76)	**
	FC2	0.55 (0.61)	-0.33 (0.82)	-0.88 (0.74)	**
	LM	0.48 (0.66)	0.63 (0.52)	0.14 (0.41)	
	RM	0.43 (0.73)	0.24 (0.74)	-0.16 (0.66)	
Control/200 ms	Fz	0.07 (0.54)	-1.13 (0.85)	-1.20 (1.14)	**
	FC1	0.04 (0.58)	-1.14 (0.76)	-1.20 (1.07)	**
	FC2	0.16 (0.53)	-0.92 (0.82)	-1.08 (1.06)	**
	LM	0.40 (0.77)	0.70 (0.86)	0.30 (0.80)	
	RM	0.32 (0.62)	0.98 (0.63)	0.65 (0.51)	**

Note: Value given as mean (SD).

* $p < .05$. ** $p < .01$.

Alternating/100 ms Condition

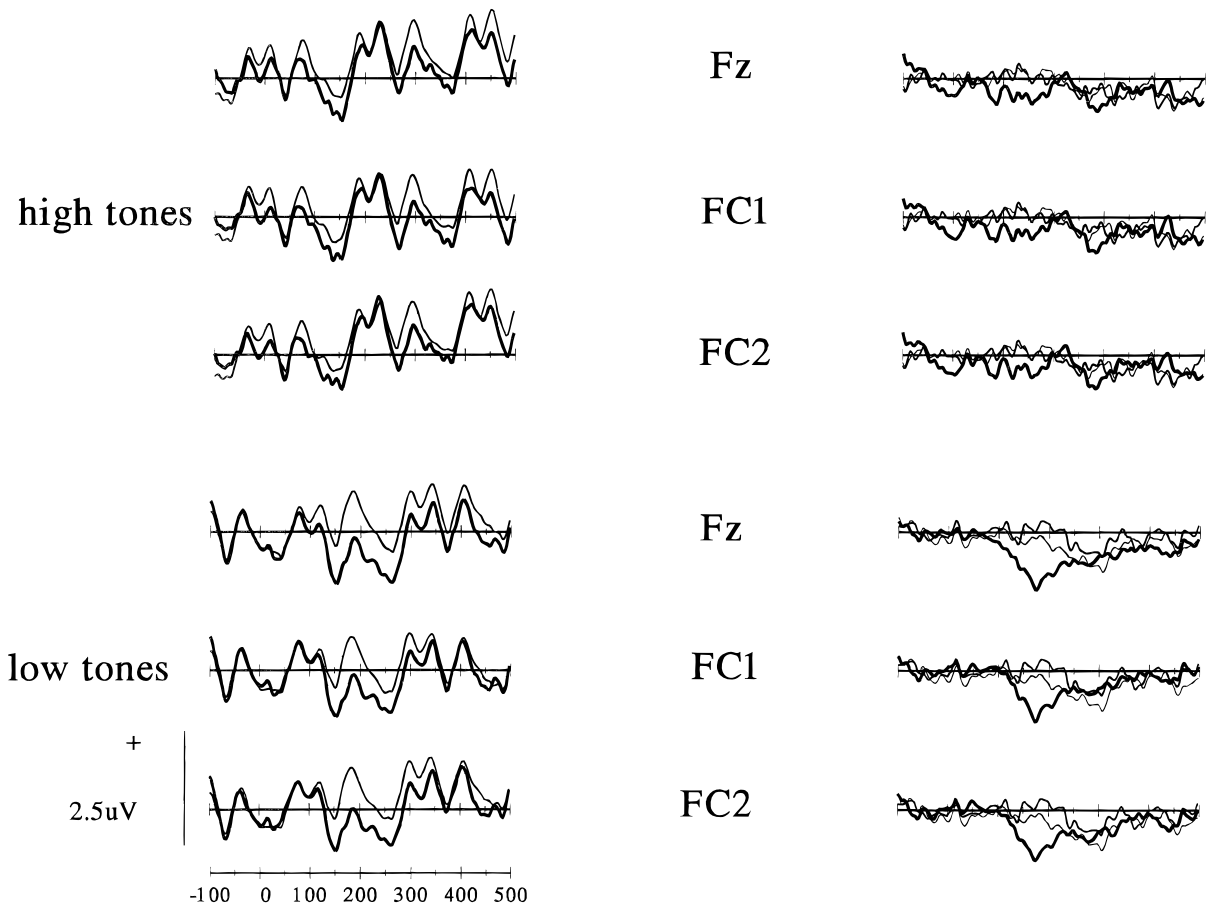


Figure 4. Alternating/100-ms condition of Experiment 1. Standard (thin line) and deviant (thick line) ERPs at Fz, FC1, and FC2 (left column) for the high (top) and low (bottom) tones. Difference waves (thick line) with ERPs recorded at the mastoids overlain (thin lines) at Fz, FC1, and FC2 (right column) for the high (top) and low (bottom) tones.

occurring within a repeating tone pattern (consisting of five different frequency elements) with no silent intervals between them (see also Winkler & Schröger, 1995). In contrast, in other MMN studies using similar stimulus sequence designs, tonal patterns were separated with intertrain (Schröger, 1994) or interpair (Paavilainen et al., 1995; Saarinen et al., 1992) intervals. The offsetting interval may facilitate the detection of the pattern of tones, which cannot be detected in an isochronous rhythm.

When the high tones were embedded in the low tones in the alternating/750-ms condition, no MMN was elicited. One possible explanation is that the high tones interfered with the emergence of the three-tone sequence within the low (or high) tones when alternated at the rate of one tone per 0.75 s. Another possibility is that the duration of the memory underlying the MMN was exceeded in this case; the low tones (or high tones) occurred once every 1.5 s.

The control/1.5-s condition was run separately to determine whether the three-tone standard sequence exceeded the limits of the memory underlying the MMN system. An MMN was obtained in four of the nine subjects, which indicates that the memory can persist long enough in some, but not in all, subjects to elicit an

MMN. The duration of the memory probably is variable among individuals. Evidence for this viewpoint can be found in a study designed to investigate whether the duration of the memory underlying MMN generation is shorter in children than in adults (for detecting a change in tonal frequency; Gomes et al., in press). In addition to their main finding of age-related differences across the subject groups (there was a positive correlation between age and duration of the memory), the data indicated that there were individual memory differences within the groups.

None of the subjects that obtained MMNs when the low tone pattern was presented alone obtained MMNs when the high tone pattern was embedded within it (i.e., when the high and low tones alternated). The same pattern of results occurred even when the pace of the tones was increased. In the control/1-s condition, half of the subjects obtained MMNs and half did not, and no MMNs occurred when the low and high tones alternated. Despite any individual subject variability, those subjects who had MMNs in the control conditions did not have MMNs when the tones alternated.

Taken together, these data show that those subjects who had significant MMNs when the low tones were presented alone at a slow pace did not have MMNs to the low tones when high tones

Alternating/750 ms Condition

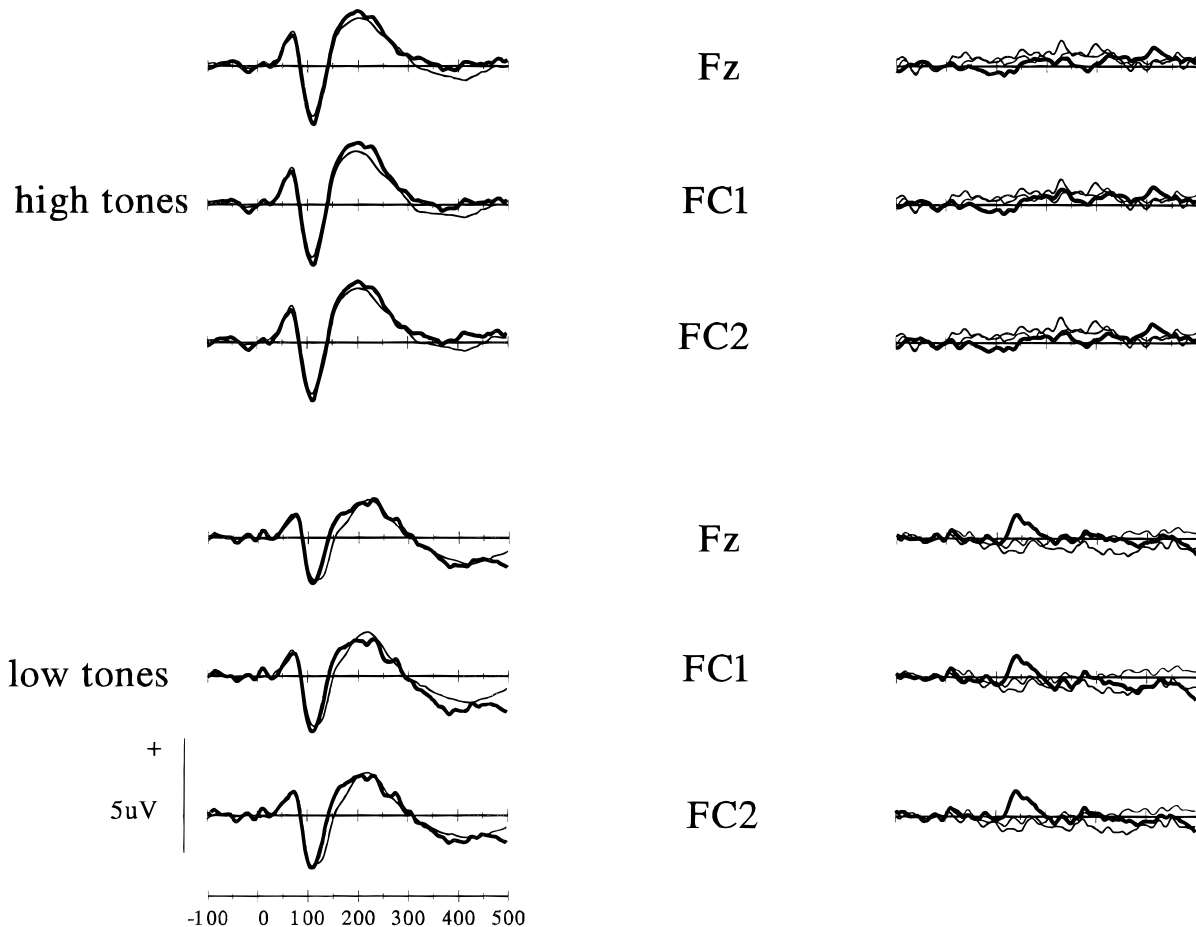


Figure 5. Alternating/750-ms condition of Experiment 1. Standard (thin line) and deviant (thick line) ERPs at Fz, FC1, and FC2 (left column) for the high (top) and low (bottom) tones. Difference waves (thick line) with ERPs recorded at the mastoids overlain (thin lines) at Fz, FC1, and FC2 (right column) for the high (top) and low (bottom) tones.

were embedded in between them. Therefore, the lack of an MMN may be attributed to interference of the within-stream sequences by the alternation of the high and low tones and not to constraints of the memory, at least for some subjects. Because the absence of an MMN in this condition was not likely due to limits in the duration of the memory needed for this paradigm, we concluded that the alternation of the high and low tones interfered with the emergence of the high- and low-pitched tonal sequences in the alternating/750-ms condition.

When the same alternating sequence was run at a faster pace, an MMN was obtained, presumably because of the sorting associated with streaming. That is, when the tones were alternated at this pace they did not interfere with one another because high- and low-tone streams emerged preattentively. The within stream patterns appear to have emerged prior to or at the level of the MMN system. These results demonstrate that the streaming effect is governed by preattentive mechanisms of the auditory system.

An alternative explanation is that the MMN in the 100-ms but not the 750-ms alternating condition could be due to an SOA effect on automatic pattern formation rather than to a segregation effect.

Although many studies have demonstrated that when the MMN is elicited its amplitude does not change as a function of SOA for features of a tone (Böttcher-Gandor & Ullsperger, 1992; Czigler, Csibra, & Csontos, 1992; Gomes et al., in press; Näätänen, Paavilainen, Alho, Reinikainen, & Sams, 1987; Sams et al., 1993; Schröger, 1996), Alain, Woods, and Ogawa (1994) reported that MMN amplitude is affected by SOA for pattern deviants (the break in a regularly alternating pitch of two tones). Alain et al. (1994) reported that a decrease in SOA produces an increase in MMN amplitude independently of the frequency range of the alternating tones presented to subjects. In this study, tones regularly alternated in frequency between three different semitone separations (1, 6, or 12) at three different SOAs (150, 400, and 900 ms). They reported an effect of SOA on the MMN amplitude when collapsed across semitone conditions. However, visual inspection of the waveforms shows that the amplitude difference was brought about by the 6- and 12-semitone conditions and not by the 1-semitone condition. That is, the amplitude remained the same across the three SOA conditions for the one semitone data. On the basis of this result, there is no expectation that the within stream pattern deviation in

the current study should be accounted for by an SOA change, because the frequency proximity of the tones was within two semitones.

Second, at the 6- and 12-semitone conditions at the faster SOA conditions a segregation effect may have occurred. Two main factors influence the manifestation of a segregation effect (when high and low tones comprise the auditory input). Decreasing the SOA and increasing the frequency separation between the high and low tones increases the probability that a segregation effect will occur. These two factors together, SOA and frequency separation, influence whether a segregation effect will occur. That is, if the pace of the tones is speeded up, a smaller frequency separation between the high and low tones will cause a streaming effect, and vice versa. Therefore, in the study of Alain et al. (1994) in which they simultaneously increased the frequency separation (6 and 12 semitones) and speeded up the pace of the high and low tones (e.g., 150 and 400 ms), they may have induced a segregation effect. If this occurred, then the MMN would have been elicited on the basis of a decrease in SOA when the break in alternation occurred and not on the basis of the change in pattern. Specifically, if the tones separated into high and low streams at the faster SOAs with the larger frequency separations, when two low tones occurred there would have been a shorter SOA between them than between the low-frequency standards. The reported change in amplitude could have been confounded by a segregation effect because they would be measuring a pattern MMN in some conditions and an SOA MMN in others. In all, it seems unlikely that the MMN obtained in the fast-paced alternating condition is due to an SOA effect.

The MMN obtained in the control/200-ms condition had a more typical scalp distribution than the MMN obtained in the alternating/100-ms condition. The grand mean difference waves obtained in the control/200-ms and alternating/100-ms conditions were similar in both amplitude and latency. However, the inversion in polarity at the mastoid sites differed between these conditions. The MMN in the control condition revealed a typical inversion at the mastoids, whereas there was no significant reversal in polarity in the mastoids in the alternating/100-ms condition. The notable distinction between the two conditions, possibly contributing to this difference, is that an MMN was elicited within a single stream of sound (a low tone stream) in the control/200-ms condition and an MMN was elicited within the context of two streams of sound (a high-tone stream and a low-tone stream) in the alternating/100-ms condition. The significant interaction of the ANOVA conducted on the scaled data provides supporting evidence in this study that different neural generators may have subserved the MMN process in the control and the experimental conditions.

In the alternating/100-ms condition, an MMN was expected in both the low-tone and high tone streams. The low-tone deviant always preceded the high tones and an MMN was obtained to the low-tone deviant but not to the high-tone deviant. A possible explanation for why no MMN occurred for the high-tone deviant in Experiment 1 is that the low-tone deviant disrupted the pattern of tones across the streams. This explanation assumes that the deviant continues for the following high-tone deviant, a total of 12 tones from the onset of the low-tone deviant, and further supposes that the break in the global pattern of the high and low tones induced an MMN on the premise that streaming did not occur. We used parameters along both the temporal and frequency domain that, based on the collection of studies conducted by Bregman and his colleagues, unambiguously induce a streaming effect. In addition to the physical parameters used in the study, subjects reported experiencing a streaming effect. Because this explanation is based

on the premise that streaming was not induced, it cannot account for the results of this study.

Because only 250 ms separated the offset of the three-tone deviant occurring in the low tones from the onset of the three-tone deviant occurring in the high tones, we speculated that the MMN system may have been subject to processing constraints. A second experiment was conducted specifically to address the issue of whether only one MMN occurred because there was not enough time for the MMN system to process both deviants.

EXPERIMENT 2

The purpose of Experiment 2 was to test whether two MMNs were not obtained in Experiment 1 because the interval between the across-stream deviance was too small for the MMN system to process both. If the global properties of a tonal sequence influence the MMN process such that an MMN can only be elicited by the first of two deviant events that occur in a row across two streams of sound, then widening the time between across-stream deviance may make an MMN appear to the second deviant as well. We hypothesized that if the same stimulus sequence were used as in Experiment 1 (alternating/100 ms) with a wider interval between the low-tone deviant and high-tone deviant, two MMNs would be obtained. This would show that streaming sorts the tones from separate sources preattentively so the MMN system can detect the deviance in both streams.

Methods

Eleven subjects (9 women, 2 men) 21–45 years of age participated in the experiment. The same procedures used for data collection and analysis in Experiment 1 were also used for Experiment 2. The stimulus sequence employed in Experiment 1 (alternating/100-ms condition) was used in Experiment 2 with the following modification. The time between the occurrence of the across-stream deviants, the low deviant and then the high deviant (see Figures 1d and e), was increased by separating the six tone cycle containing the low deviant from the six tone cycle containing the high deviant with a six-tone cycle of standards (see Figure 1c).

Results

Figure 6 displays the grand averages of the ERPs elicited by the standard and deviant waveforms, separately for the low and the high streams (left). The difference waveforms obtained by subtracting the standard ERPs from the deviant ERPs separately for each stream are also displayed (right). A negative deflection can be seen in the difference waves, peaking at about 175 ms in the low stream and about 135 ms in the high stream (Figure 6, right column), delineating the MMN components.

An MMN was obtained in both streams, established in a two-way ANOVA for repeated measures with variables of stimulus type and electrode, calculated on the low and high tones separately in the latency window of the MMN, $F(1,10) = 7.1$, $p = .024$; $F(1,10) = 6.3$, $p = .031$, respectively. Tukey HSD post hoc comparisons were conducted to confirm the presence of the MMN at the same electrodes as reported in Experiment 1. Table 2 presents the grand mean amplitudes and standard deviations of the standard and deviant ERPs measured in the latency range of the MMN. The mean amplitudes of the difference waves are also reported.

Experiment 2

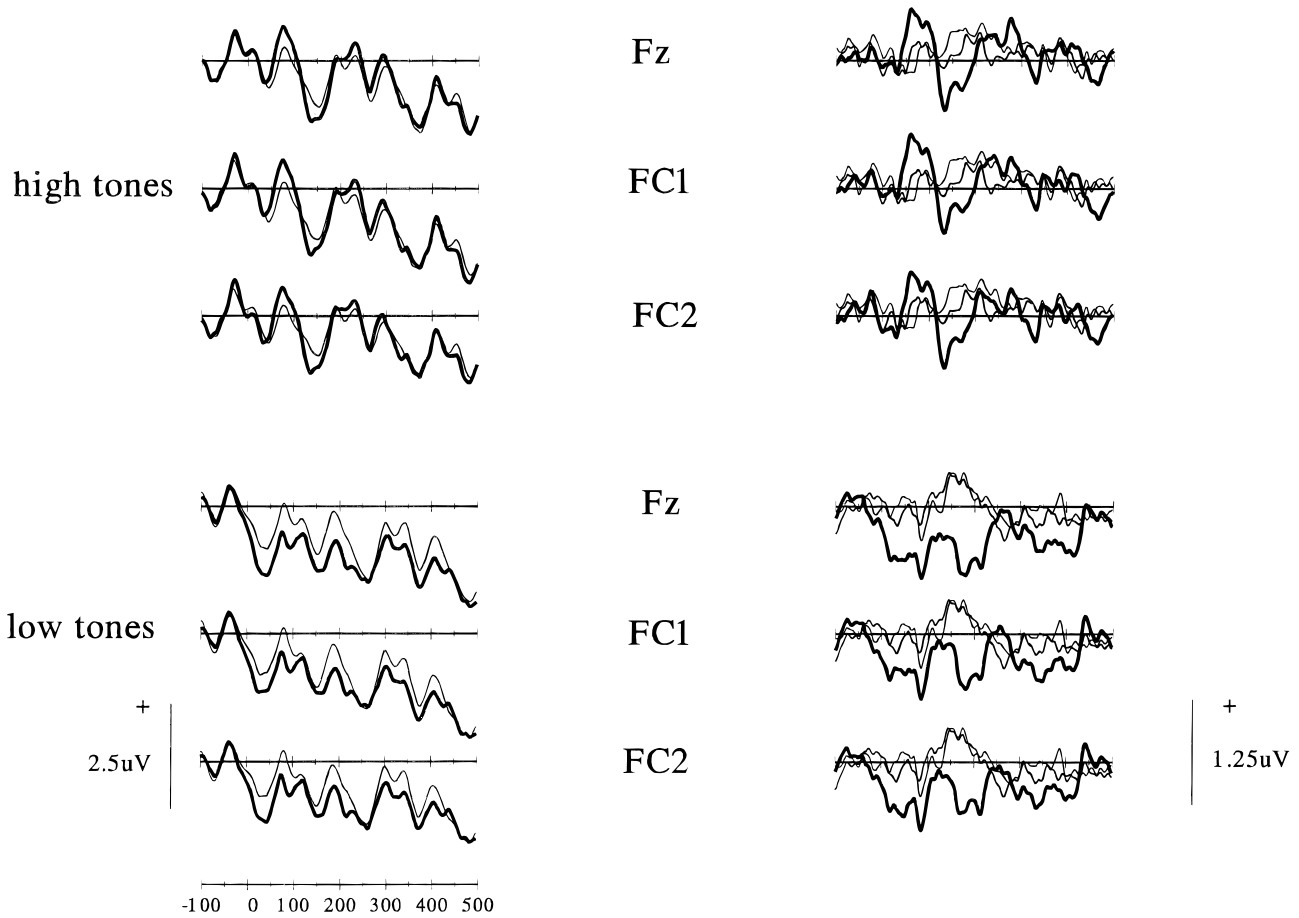


Figure 6. Experiment 2. Standard (thin line) and deviant (thick line) ERPs elicited by the high tones (left column, top) and difference waves (thick line) with ERPs recorded at the mastoids overlain (thin line; right column, top) at Fz, FC1, and FC2. Standard (thin line) and deviant (thick line) ERPs elicited by the low tones (left column, bottom) and difference waves (thick line) with ERPs recorded at the mastoids overlain (thin line; right column, bottom) at Fz, FC1, and FC2.

Discussion

The results of Experiment 1 were somewhat puzzling because we expected that if the streams segregated preattentively two MMNs would be obtained. Because only one MMN was obtained, Experiment 2 was conducted to try to clarify the results. The streaming condition from Experiment 1 was run with an additional six-tone standard cycle separating the low-tone deviant from the high-tone deviant. An MMN was obtained in both the low- and high-tone streams. These results extend our finding in Experiment 1 that the sorting of streams from separate sources occurs preattentively when alternating high and low tones are presented at a fast enough rate to induce the streaming effect.

Although the ratio differences among the three high tones were slightly smaller than those among the three low tones, the presence of the MMN in both streams in Experiment 2 argues against an explanation that no MMN occurred in Experiment 1 because the discrimination was more difficult for the high tones. If the discrimination were more difficult for the high than for the low tones, then no MMN should have been obtained for the high tones in

Experiment 2, because the frequency separations were exactly the same between Experiments 1 and 2.

The two MMNs obtained in Experiment 2 seem to explain why only one MMN was obtained in Experiment 1. Apparently, there was not enough time between the occurrence of the three-tone deviants, across streams, for the MMN system to process both of them. The results of Experiment 2, therefore, indicate that the MMN system may be subject to processing constraints in the context of streaming. Another consideration is that it may take more time to process a three-tone standard than to process a single-tone standard. We reasoned, accordingly, that both aspects of the paradigm used in Experiment 1 (the three-tone standard and the two streams of sound) influenced how the deviants were processed with respect to time.

SUMMARY

When alternating high and low tones occur at a rapid pace, the tones are sorted to separate streams of sound. This sorting process facilitates the ability to identify the order of the within-stream

Table 2. Amplitude of the Difference Waveform and the Standard and Deviant ERPs Measured on the Grand Means in the Latency Window of the MMN for Experiment 2

Tones	Electrode	Standard (μV)	Deviant (μV)	Difference (μV)	p
Low	Fz	-0.46 (0.44)	-1.16 (0.92)	-0.70 (0.92)	**
	Cz	-0.43 (0.53)	-0.58 (0.60)	-0.16 (0.74)	**
	FC1	-0.59 (0.39)	-1.12 (0.78)	-0.52 (0.70)	**
	FC2	-0.40 (0.53)	-0.90 (0.85)	-0.51 (0.95)	**
	LM	-0.43 (0.48)	-0.17 (0.55)	0.25 (0.40)	**
	RM	-0.17 (0.92)	0.01 (1.03)	0.25 (0.58)	**
High	Fz	-0.86 (0.82)	-1.24 (0.98)	-0.38 (0.69)	**
	Cz	-0.68 (0.87)	-1.10 (0.92)	-0.42 (0.91)	**
	FC1	-1.00 (0.75)	-1.30 (0.86)	-0.30 (0.66)	**
	FC2	-0.72 (0.87)	-1.11 (0.95)	-0.38 (0.65)	**
	LM	-0.15 (0.53)	0.14 (0.76)	0.29 (0.44)	**
	RM	0.06 (0.81)	0.11 (0.86)	0.05 (0.40)	*

Note: Values given as mean (SD).
* $p < .05$. ** $p < .01$.

sequences but impedes the ability to identify the order of across stream sequences (Bregman, 1990). When Bregman and Campbell (1971) presented high and low tones fast enough to produce the streaming effect, subjects mostly reported the order of the tones occurring within streams rather than across streams. In this study, the MMNs obtained in the fast-paced alternating conditions show

that the ability to detect within-stream sequences over across stream sequences was facilitated. The three-tone deviant sequences could be detected as different than the three-tone standard sequences as a function of the streaming effect. When the high and low tones alternated at a slower pace, the three-tone within stream standard sequences did not emerge.

REFERENCES

- Alain, C., & Woods, D. L. (1997). Attention modulates auditory pattern memory as indexed by event-related brain potentials. *Psychophysiology*, 34, 534–546.
- Alain, C., Woods, D. L., & Ogawa, K. H. (1994). Brain indices of automatic pattern processing. *NeuroReport*, 6, 140–144.
- Alho, K., & Sinervo, N. (1997). Preattentive processing of complex sounds in the human brain. *Neuroscience Letters*, 233, 33–36.
- Alho, K., Woods, D. L., Algazi, A., & Näätänen, R. (1992). Intermodal selective attention. II. Effects of attentional load on processing of auditory and visual stimuli in central space. *Electroencephalography and Clinical Neurophysiology*, 82, 356–368.
- Böttcher-Gandor, C., & Ullsperger, P. (1992). Mismatch negativity in event-related potentials to auditory stimuli as a function of varying interstimulus interval. *Psychophysiology*, 29, 546–550.
- Bregman, A. S. (1978). Auditory streaming is cumulative. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 380–387.
- Bregman, A. S. (1990). *Auditory scene analysis*. Cambridge, MA: MIT Press.
- Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, 89, 244–249.
- Broadbent, D. E. (1958). *Perception and communication*. New York: Pergamon.
- Cowan, N., Winkler, I., Teder, W., & Näätänen, R. (1993). Memory prerequisites of the mismatch negativity in the auditory event-related potential (ERP). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 909–921.
- Czigler, I., Csibra, G., & Csontos, A. (1992). Age and inter-stimulus interval effects on event-related potentials to frequent and infrequent auditory stimuli. *Biological Psychology*, 33, 195–206.
- Dowling, W.J. (1973). The perception of interleaved melodies. *Cognitive Psychology*, 5, 322–337.
- Gomes, H., Sussman, E., Ritter, W., Kurtzberg, D., Cowan, N., & Vaughan, H.G., Jr. (in press). Electrophysiological evidence of developmental changes in the duration of auditory sensory memory. *Developmental Psychology*.
- Javitt, D., Doneshka, P., Zylberman, I., Ritter, W., & Vaughan, H.G., Jr. (1993). Impairment of early cortical processing in schizophrenia: An event-related potential confirmation study. *Biological Psychiatry*, 33, 513–519.
- Jones, M.R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 83, 323–355.
- Jones, M.R., Kidd, G., & Wetzel, R. (1981). Evidence for rhythmic attention. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 1059–1073.
- Jones, M.R., Maser, D.J., & Kidd, G.R. (1978). Rate and structure in memory for auditory patterns. *Memory and Cognition*, 6, 246, 258.
- McCarthy, G., & Wood, C.C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology*, 62, 203–208.
- Näätänen, R. (1992). *Attention and brain function*. Hillsdale, NJ: Erlbaum.
- Näätänen, R., Paavilainen, K., Alho, K., Reinikainen, & Sams, M. (1987). Interstimulus interval and the mismatch negativity. In C. Barber & T. Blum (Eds.), *Evoked potentials III* (pp. 392–397). London: Butterworths.
- Näätänen, R., Paavilainen, P., Tiitinen, H., Jiang, D., & Alho, K. (1993a). Attention and mismatch negativity. *Psychophysiology*, 30, 436–450.
- Näätänen, R., Schröger, E., Karakas, S., Tervaniemi, M., & Paavilainen, P. (1993b). Development of a memory trace for a complex sound in the human brain. *NeuroReport*, 4, 503–506.
- Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.
- Nordby, H., Roth, W.T., & Pfefferbaum, A. (1988). Event-related potentials to breaks in sequences of alternating pitches or interstimulus intervals. *Psychophysiology*, 25, 262–268.
- Norman, D.A. (1967). Temporal confusion and limited capacity processors. *Acta Psychologica*, 27, 293–297.
- Novak, G., Ritter, W., & Vaughan, H.G., Jr., & Wiznitzer, M.L. (1990). Differentiation of negative event-related potentials in an auditory discrimination task. *Electroencephalography and Clinical Neurophysiology*, 75, 255–275.
- Paavilainen, P., Saarinen, J., Tervaniemi, M., & Näätänen, R. (1995). Mismatch negativity to changes in abstract sound features during dichotic listening. *Journal of Psychophysiology*, 9, 243–249.

- 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 & Hearing, 16*, 52–67.
- Saarinen, J., Paavilainen, P., Schröger, E., Tervaniemi, & Näätänen, R. (1992). Representation of abstract attributes of auditory stimuli in the human brain. *NeuroReport, 3*, 1149–1151.
- Sams, M., Hari, R., Rif, J., & Knuutila, J. (1993). The human auditory sensory memory trace persists about 10 s: Neuromagnetic evidence. *Journal of Cognitive Neuroscience, 5*, 363–370.
- Schneider, W., & Shiffrin, R.M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*, 1–66.
- Schröger, E. (1994). An event-related potential study of sensory representation of unfamiliar tonal patterns. *Psychophysiology, 31*, 175–181.
- Schröger, E. (1996). The influence of stimulus intensity and inter-stimulus interval on the detection of pitch and loudness changes. *Electroencephalography and Clinical Neurophysiology, 100*, 517–526.
- Schröger, E., Näätänen, R., & Paavilainen, P. (1992). Event-related brain potentials reveal how non-attended complex sound patterns are represented by the human brain. *Neuroscience Letters, 146*, 183–186.
- Schröger, E., Paavilainen, P., & Näätänen, R. (1994). Mismatch negativity to changes in a continuous tone with regularly varying frequencies. *Electroencephalography and Clinical Neurophysiology, 92*, 140–147.
- Sussman, E., Ritter, W., & Vaughan, H.G., Jr. (1998). Attention affects the organization of auditory input associated with the mismatch negativity system. *Brain Research, 789*, 130–138.
- Winkler, I., Cowan, N., Csépe, V., Czigler, I., & Näätänen, R. (1996). Interactions between transient and long-term auditory memory as reflected by the mismatch negativity. *Journal of Cognitive Neuroscience, 8*, 403–415.
- Winkler, I., & Schröger, E. (1995). Neural representation for the temporal structure of sound patterns. *NeuroReport, 6*, 690–694.
- 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.

(RECEIVED July 23, 1997; ACCEPTED May 19, 1998)