Published in final edited form as:

Brain Res. 2006 February 23; 1075(1): 165–174. doi:10.1016/j.brainres.2005.12.074.

Neurophysiological evidence for context-dependent encoding of sensory input in human auditory cortex

Elyse Sussman^{a,b,*} and Mitchell Steinschneider^{a,c}

^a Department of Neuroscience, Albert Einstein College of Medicine, 1410 Pelham Parkway South, NY 10461, USA

^b Department of Otorhinolaryngology-Head and Neck Surgery, Albert Einstein College of Medicine, 1410 Pelham Parkway South, NY 10461, USA

^c Department of Neurology, Albert Einstein College of Medicine, 1410 Pelham Parkway South, NY 10461, USA

Abstract

Attention biases the way in which sound information is stored in auditory memory. Little is known, however, about the contribution of stimulus-driven processes in forming and storing coherent sound events. An electrophysiological index of cortical auditory change detection (mismatch negativity [MMN]) was used to assess whether sensory memory representations could be biased toward one organization over another (one or two auditory streams) without attentional control. Results revealed that sound representations held in sensory memory biased the organization of subsequent auditory input. The results demonstrate that context-dependent sound representations modulate stimulus-dependent neural encoding at early stages of auditory cortical processing.

Keywords

Auditory cortex; Event-related potentials (ERPs); Mismatch negativity (MMN); Priming; Sensory memory

1. Introduction

The relative contribution of stimulus-driven and attentional processes in organizing the memory representation of the acoustic environment is controversial. There is clear evidence that attention biases response properties of neurons within sensory cortices, indicating that the neural mechanisms of attention modulate the stimulus-driven processes (Bushnell et al., 1981; Chapin and Woodward, 1981; Chelazzi et al., 1993; Desimone, 1998; Hocherman et al., 1976; Hubel et al., 1959; Jancke et al., 1999; Luck et al., 1997; Moran and Desimone, 1985; Nelson et al., 1991; Petkov et al., 2004; Woodruff et al., 1996). In humans, attended sounds evoke enhanced scalp-recorded electrical responses when compared to the responses elicited by unattended sounds (Hillyard et al., 1973). This enhancement begins as early as 20–50 ms after stimulus onset (Hillyard et al., 1973, 1998; Woldorff and Hillyard, 1991; Woldorff et al., 1987), indicating that actively selecting a subset of sensory information can modify responses evoked by sensory input at an early processing stage. While there is convincing evidence that attention modifies neural activity evoked by sounds, little is known about modulatory processes

*Corresponding author. Department of Neuroscience, Albert Einstein College of Medicine, 1410 Pelham Parkway South, Bronx, NY 10461, USA. Fax: +1 718 430 8821. esussman@aecom.yu.edu (E. Sussman).

reflecting context-dependence of stimulus-driven neural activity. Context dependence refers to the modulation of neural activity evoked by identical stimuli in different stimulus environments. It is not known, for example, whether existing auditory memory representations affect the encoding of incoming information without attentional control. In the current study, we test the hypothesis that sensory memory representations formed on the basis of stimulus-driven processes can bias subsequent memory representations toward one sound context over another (one or two auditory streams).

Event-related brain potentials (ERPs) provide a noninvasive electrophysiological measure of cortical auditory processing. The mismatch negativity (MMN) component of event-related brain potentials (ERPs) reflects the outcome of a change detection process that is based upon the memory of sound regularities (often called the “standard”) in ongoing auditory input (Näätänen et al., 2001; Näätänen and Winkler, 1999; Picton et al., 2000; Sussman and Gumenyuk, 2005; Sussman et al., 1998b, 1999, 2002a; Winkler et al., 1996). Incoming sounds that deviate from the neural representation of the standard sound elicit MMN. Infrequent changes in simple auditory features such as frequency, intensity, tone duration, or spatial location, as well as changes in more complex features such as sequential tone patterns, elicit MMN (e.g., Sussman et al., 1999, for review, see Näätänen et al., 2001). MMN is typically observed with a peak latency between 100 and 200 ms from the time that the deviation from the regularity is detected. Thus, MMN represents an early process of deviance detection based upon a memory of the previous sound stimulation.

Support for a role of MMN in indexing memory processes is provided by the finding that MMN generation is suppressed by blockage of *N*-methyl-D-aspartate (NMDA) receptors without disturbing elicitation of the prior obligatory cortical response to sound onsets (Javitt et al., 1996). NMDA-mediated activity is associated with cellular mechanisms involved in memory processes (Compte et al., 2000; Kauer et al., 1988; Pulvirenti, 1992), suggesting a link between memory processes and MMN.

The main neural generators of the MMN are located bilaterally in the supratemporal plane, as determined by dipole-modeling of electric (Scherg et al., 1989) and magnetic responses (Sams and Hari, 1991; Woldorff et al., 1998), scalp-current density maps of scalp-recorded ERPs (Giard et al., 1990), functional magnetic resonance imaging (fMRI) (Opitz et al., 1999, 2005), and intracortical ERP recordings (Halgren et al., 1995; Kropotov et al., 1995, 2000). The location of the generators in auditory cortex accounts for the observed scalp topography of the waveform, which is maximally negative over the frontocentral scalp locations and often inverts in polarity below the Sylvian fissure. Thus, MMN primarily represents the outcome of memory-related processes represented within auditory cortices.

Several characteristics of this deviance detection process make the MMN an especially useful index for the purposes of the present study. Firstly, MMN is elicited by the violation of detected auditory regularities (the “standard”), and thus can be used to determine which regularities (individual features or pattern of sounds) are represented in memory at the time the “deviant” occurs. Secondly, MMN is elicited even when the sounds have no relevance to ongoing behavior, as its elicitation does not require participants to actively detect the deviant sounds (Rinne et al., 2001; Ritter et al., 1999; Sussman et al., 2002b, 2003a). Moreover, focused attention in and of itself does not affect MMN elicitation (Sussman et al., 2003b). Additionally, Sussman et al. (2005a) found no significant difference in the amplitude of the MMN elicited in conditions of highly focused attention to a visual task compared to a low load visual task (watching a video) (see also Winkler et al., 2005). Thirdly, Näätänen and his coworkers established that the process generating the MMN component is based on an auditory sensory-memory representation of the standard and not by differences in the refractory state of the neural elements activated by the standard and the deviant stimuli (Näätänen, 1992; Näätänen

and Alho, 1997; Näätänen and Winkler, 1999); see also Jacobsen et al., 2003). Finally, the MMN shows good correspondence with the perception of the same sounds in performance of active tasks, including, but not limited to, the ability to index the amount of change (Tiitinen et al., 1994), to index the distinction between one and two auditory streams (Sussman et al., 2001), and to index learning effects (Näätänen et al., 1993). These properties of the MMN-generating process allow for the assessment of memory formation of the sound input when attention is not focused on the sounds and no task is required of the participants.

In the current study, deviants were used as ‘probes’ of the neural representations extracted from the ongoing sound input. This is possible because the response to a particular sound is based upon the memory of the previous sounds. As an example, in the well-known auditory oddball paradigm, an “oddball” or infrequently occurring sound is presented randomly among instances of a frequently repeating sound. The *oddball* elicits MMN because it is detected as deviating in some sound feature (e.g., frequency, intensity, duration, spatial location) from the frequently repeating sound. However, when the same ratio of frequent to infrequent sounds is presented in a fixed order (e.g., the ‘oddball’ tone, represented by an ‘O’, occurs every fifth tone: XXXXOXXXXO...) rather than being presented (20% of the time) in a randomized order, if the brain detects the regularity of the sequence, then no MMN will be elicited by the “oddball” tone (Sussman and Gumenyuk, 2005; Sussman et al., 1998b, 2002a). It is important to note from this that MMN is *not* necessarily elicited simply when there is a situation of a frequent and an infrequent tone presented in a sound sequence. What is crucial to its elicitation is what is detected as the regularity in the sound sequence and stored in auditory memory. Thus, MMN elicitation is highly dependent upon the context of the stimuli. Because MMN is context-based, it can be used to assess features of how the acoustic information is structured and stored in memory. These considerations served as the logical basis for the current study, in that MMN elicitation would indicate which regularity was extracted and stored in memory at the time the probe tones occurred.

In the current study, we test the hypothesis that sensory memory representations can be biased toward one sound organization over another (the representation of one vs. two auditory streams) without direct attentional control. Subjects were presented with two tones of different frequencies (called ‘X’ and ‘O’) in a fixed sequential pattern (XXXXOXXXXO...) that was perceptually ambiguous for hearing one or two sound streams (see Stimuli and procedures for further details). An additional set of sounds (the “prime”) was presented prior to the “test” sounds (see Fig. 1) to encourage the ambiguity to be resolved as either one stream (*One-stream-prime* condition) or two streams (*Two-streams-prime* condition). To assess how sound regularities were stored in memory, a “probe tone” (an infrequent higher intensity ‘X’ tone) was embedded within the test sounds but not within the prime. The paradigm was designed so that a regularity of tone intensity among the ‘X’ tones (they all had the same tone intensity) could only be detected when the ‘X’ and ‘O’ tones segregated into separate streams. The higher intensity probe tones would thus elicit MMN, detected as deviant with respect to the intensity of the ‘X’ tones, only when the overall sequence was represented as two streams. If the ‘X’ and ‘O’ tones were detected as part of the same stream, in contrast, the intensity variation of the ‘O’ tones (see Experimental procedures for details) would prevent the emergence of any regularity in intensity, and thus the brain’s response to the probe tones would not differ from the standard ‘X’ tones. Thus, in the present study, MMN elicitation indexes whether the ‘X’ and ‘O’ sounds were represented as segregated or integrated in auditory memory. If the priming sounds can influence the neural representation of the test sounds, then MMN should be elicited in the *Two-streams-prime* condition (segregated organization), and not in the *One-stream-prime* condition (integrated organization), thus demonstrating context-dependence of stimulus-driven processes.

2. Results

2.1. ERP analysis

Fig. 2A displays the grand-mean ERP waveforms elicited by the probe and control sounds for each condition. The obligatory P1–N1 responses can be seen clearly within the 500 ms poststimulus epoch. P1 latency was $M = 50$ ms [SD = 18], 45 ms [SD = 15], and 51 ms [SD = 17], in the *Two-streams-prime*, *One-stream-prime*, and *No-prime* conditions, respectively. N1 latency was also, respectively, $M = 107$ ms [SD = 19], 97 ms [SD = 20], and 107 ms [SD = 21]. P1–N1 components elicited by successive stimuli can also be seen within the epoch, occurring regularly at the 180-ms stimulus onset rate of the sounds.

The responses evoked by the frequent control stimuli did not differ as a function of condition ($F(2,18) < 1$, $\varepsilon = 0.99$, $P > 0.8$). That is, there was no statistically significant difference in the mean amplitude of the ERP responses elicited by the control tones (measured from peak-to-trough) among the three conditions (Two-streams-prime condition: $M = -1.30$, SD = 0.76; One-stream-prime condition: $M = -1.22$, SD = 0.90; and No-prime condition: $M = -1.35$, SD = 0.82).

The key result of this study was that only when the *Two-streams-prime* preceded the test tones was a significant MMN component detected within the ERPs elicited by the probe tone (see Fig. 2A). A repeated measures ANOVA with factors of condition (Two-streams-prime vs. No-prime vs. One-stream-prime), stimulus type (control vs. deviant), and electrode (Fz, F3, F4, FC1, FC2) revealed a main effect of condition ($F(2,18) = 10.42$, $\varepsilon = 0.8$, $P < 0.003$), a main effect of stimulus type ($F(1,9) = 7.45$, $P < 0.03$), a main effect of electrode ($F(4,36) = 5.46$, $\varepsilon = 0.4$, $P < 0.03$), and a significant interaction between condition and stimulus type ($F(2,18) = 10.37$, $\varepsilon = 0.9$, $P < 0.002$) with no other interactions. The post hoc analysis of the interaction showed that the deviant stimuli were significantly more negative than the control stimuli *only* for the *Two-streams-prime* condition (Two-streams-prime: $P < 0.0025$; No-prime: $P > 0.51$; and One-stream-prime: $P > 0.62$). Thus, the probe tone elicited a detectable MMN only when the test sounds were primed for a segregated organization. Fig. 2B (top row) also shows the voltage maps, which are characterized by negative fields that are maximum at frontocentral sites in the condition where MMN was elicited (Two-streams-prime condition), with positive field potentials at sites below the Sylvian fissure.

Attention-related components, which can be elicited by detected deviations of sound regularities when attention is focused on the sounds (e.g., N2b/P3b) or by the saliency of the detected violations drawing attention involuntarily to the sounds (e.g., P3a; see Friedman et al., 2001 for review), were not observed in the current set of data. The absence of any attention-related ERP components in the current data suggests that participants were not overtly (e.g., no N2b/P3b) or covertly (e.g., no P3a) listening to the sounds.

2.2. Scalp current density

Fig. 2B (bottom row) also displays the SCD maps for all three conditions. The peak latency of the MMN component, elicited in the *Two-streams-prime* condition, was observed in the grand-averaged probe tone-minus-control difference ERP waveforms at 152 ms from stimulus onset, consistent with the MMN latency observed in similar paradigms (Sussman et al., 2001). This latency was thus used to calculate the voltage and Laplacian waveforms for all three conditions. The SCD maps show the frontocentral negative field in the *Two-streams-prime* condition as having bilateral current sinks at FC1 and FC2 electrodes in the condition where MMN was elicited. This is typical distribution for MMN, consistent with bilateral generators in auditory cortices (Alho, 1995). It should be noted that this scalp distribution was not observed for either

the *One-stream-prime* or the *No-prime* conditions, where there was no significant difference between the ERP response elicited by the probe and control tones.

2.3. Analysis of the control condition ERPs

While the present findings indicate that MMN was elicited only in the *Two-streams-prime* condition, it is possible that the ERP response represented an enhanced N1 component evoked by the louder probe tone in an environment where other 'X' tones were of softer intensity. This interpretation can be challenged, at least in part, by demonstrating an MMN to an intensity decrement. We performed this analysis by examining the difference waveforms from the control conditions for the three main conditions (Fig. 3). While these control conditions were originally intended to obtain a comparison waveform for delineating the MMN in the main conditions (i.e., to elicit an ERP response to a physically identical stimulus, see Control for delineating the MMN component for further details), they provide an excellent opportunity to examine the waveforms in a situation in which an intensity decrement rather than an intensity increment was the MMN-eliciting probe tone. The ERP response to the "standard" 'X' tones from the main experiments (intensity 71 dB) was subtracted from the ERP responses to the intensity decrement probe tone (71 dB) of the control conditions. To identify the presence of the MMN component, the mean amplitude was measured in a 40-ms window around the peak negativity (228 ms) observed in the grand mean waveforms for each individual and then subjected to statistical analysis. It should be noted that the peak was longer in latency than the main conditions (228 ms vs. 152 ms), which may have been due to the MMN being generated by an intensity decrement vs. an intensity increment, or to the absence of the possible N1 enhancement. A one-sample *t* test was used to test whether the negativity observed in the waveforms was significantly greater than zero, indicating the presence of MMN. Despite the fact that a smaller number of probe tones were obtained, this analysis revealed a significant MMN in the *Two-streams-prime* control condition ($t_9 = 2.44$, $P < 0.038$) but not in the *One-stream-prime* ($t_9 = 1.51$, $P > 0.163$) or the *No-prime* ($t_9 = 0.866$, $P > 0.409$) control conditions.

3. Discussion

The results of this study indicate that the organization of sound input in auditory cortex is influenced by the preceding stimulus patterns. Even though the immediate acoustic environment surrounding the probe tone was identical in all three experimental conditions, the probe tone elicited MMN only when the test sounds were primed for a segregated organization and not when they were primed for an integrated organization. These results suggest that (1) the test sounds were maintained in memory as either one integrated stream or two separate streams according to the stimulus pattern that preceded them; and (2) that previous sounds bias the way the neural trace encodes incoming sensory information in auditory memory.

A common concern when using a passive paradigm, such as was used in the current study, is whether attention was focused on the sound sequences even though participants were requested to ignore the auditory input and watch a silent video. In a previous study (Winkler et al., 2005), no difference was found in the amplitude or latency of the MMN elicited when participants watched a silent video compared to when participants ignored the test sounds and performed a demanding noise change detection task. This result showed that the passive viewing condition provides a reasonable testing situation for assessing the outcome of preattentive processes as indexed by MMN (see also Sussman et al., 2003b). Moreover, electrophysiological indices (e.g., the attention-related ERP components) can be used to ascertain whether or not participants attended to the sounds or noticed the structure of the sounds. In the current set of data, no N2b/P3b components (an index of voluntary attention), and no P3a components (an index of involuntary attention) were observed. The absence of attention-related ERP components supports the claim that participants ignored the acoustic

input, and provides further evidence that the results of the current study were not determined by attentional mechanisms.

However, despite all evidence that attention was not focused on the sounds, the possibility that some subtle, covert attentional modulation may at least in part be responsible for the results cannot be ruled out. For example, there might have been slight attention shifts when the prime blocks were changed. Although we do not think that the effects were due to attentional shifts, this possibility needs to be acknowledged and explored in greater detail in future studies.

Another concern, when using a louder intensity sound for eliciting MMN, is that the response may reflect an N1 enhancement rather than a true MMN. Although this is a possibility, multiple considerations argue against N1 enhancement as an explanation for the current set of results. First, results of animal studies indicate that at the suprathreshold intensities used in the current study, the firing rates of neurons change very little in response to intensity level if they have monotonic rate intensity functions, and can even decrease for those cells that have nonmonotonic intensity functions (e.g., Schreiner et al., 1992). Moreover, in humans, at the intensity and frequency levels used in the current study, only a minimal difference in N1 amplitude has been reported (Biermann and Heil, 2000; Schröger, 1994). Second, the comparison response used to delineate the MMN component had the same physical parameters as the deviant sound (i.e., both the probe and the comparison tones had the same intensity value). Third, there was no difference in the stimulus parameters *surrounding* the probe tone. If there had been an effect of the probe tone being louder, then all three conditions should have had an N1 enhancement because the acoustic environment immediately surrounding the probe tone, for an average time of 7 seconds, was identical in all three conditions. Fourth, the strongest evidence is the analysis of the control condition data (see Fig. 3), in which an intensity *decrement* was the probe tone. This analysis revealed the same pattern of results as in the main experiment: a statistically significant negativity was present only for the Two-streams-prime condition. This negativity elicited by the softer probe tones only in the Two-streams-prime condition cannot be explained by an enhancement of the N1. Although further testing is warranted to confirm these control condition results because of the fewer number of probe tones obtained per subject, the evidence argues against an N1 enhancement as an explanation for the current results.

3.1. Context-dependent encoding of auditory information in auditory cortex

The current data support the hypothesis that what is held in auditory memory is based not only upon the acoustic characteristics of the sounds themselves but also upon the relationship between successive sound elements of the previous input. This means that auditory processes, including the cortical level where MMN is generated, rely on contextual information and preserve the relationship between successive sounds. The current memory representation would thus form the basis for evaluating new, incoming sensory inputs, and is indicative of a context-dependent neural process. Contextual influences on response properties of neurons in auditory cortex have been found in animal models. The context, or environment surrounding a probe tone, modulates the amplitude of both single- and multi-unit neural activity. Contextual factors such as temporal sound density (Blake and Merzinich, 2002; Brosch and Schreiner, 2000), temporal onset time (Steinschneider et al., 2005), interaural phase disparity (Malone et al., 2002), or forward (Brosch and Schreiner, 1997; Fishman et al., 2004) and backward (Brosch et al., 1998) masking, arising from properties of adjacent sounds, have been shown to influence the way that auditory cortical neurons respond to the probe tone. What is unique in the current study is that the cortically generated response to the probe tone was modified by sounds not in its immediate proximity. We can rule out any simple effects of a tone on its neighbor (such as masking) because the sound environment around the probe tone (the immediately adjacent sounds) was identical in all three conditions. Thus, the current results specify a role for higher-

order context-dependent neural processes in auditory cortex not previously observed in animal studies.

3.2. Biasing not dependent upon attention

Attention clearly biases the way successive stimuli are encoded in auditory cortex (Sussman et al., 1998a, 2002a). When a portion of auditory information is selected for further evaluation (e.g., selecting a pattern of high sounds and ignoring lower-pitched sounds), a reorganization of the neural memory representations used in the auditory deviance detection process occurs (Sussman et al., 1998a). This biasing is contingent upon the listener's task. In contrast, the current results indicate that there is also a stimulus-driven modulation of neural activity that is not solely dependent upon attentional mechanisms. As such, sensory input was "primed" simply by the context of the sound information currently held in sensory memory, a process occurring early in the representation of sound at the cortical level.

4. Experimental procedures

4.1. Subjects

Ten subjects (four males) between the ages of 23 and 45 years (mean age 34 years) were paid to participate in the study. All participants passed a hearing test and reported no history of hearing or neurological problems. Participants gave informed consent after the procedures were explained to them, in accordance with the human subjects research protocol approved by the Committee for Clinical Investigations at the Albert Einstein College of Medicine.

4.2. Stimuli and procedures

Three 50 ms pure tones (5 ms rise and fall times) were created with Neuroscan STIM software and presented binaurally with insert earphones. The "test" sounds were presented in a fixed pattern –XOOOXOOOXOOO... – at a fixed onset-to-onset interval of 180 ms (see Fig. 1). The frequency of the 'X' tones was 1046 Hz and the frequency of the 'O' tones was 1397 Hz (a 5 semitone difference). 'X' tones had two different intensity values. The intensity of the frequent 'X' tones (90%) was 71 dB SPL and the intensity of the infrequent 'X' tones (10%) was 83 dB SPL (tone intensity was calibrated with a Brüel and Kjær sound-level meter). "O" tones were randomly and equiprobably distributed around the 'X' tones at intensities of 67 dB, 75 dB, 79 dB, and 87 dB. The infrequently occurring 'X' tones with a higher intensity than the frequently occurring 'X' tones were the "probe tones" (see Fig. 1). Note that intensity values of the frequent and infrequent 'X' tones were neither the highest nor lowest value of the range of intensities presented in the sequence. Thus, the standard intensity could not be distinguished as the softest sounds and the deviant intensity could not be distinguished as the loudest sounds in the overall sequence.

The sound sequence was perceptually ambiguous for hearing one or two sound streams, as determined in a separate behavioral study that used the same type of stimulus paradigm as in the current study (Sussman et al., 2005b). In the behavioral study, listener's heard two distinct streams approximately 50% of the time at a 5-semitone frequency separation. Similar ambiguity in behavioral results has been found using different paradigms but similar frequency distances between sounds (Bey and McAdams, 2003; Carlyon et al., 2001). The sounds may be heard as a repeating four-tone pattern of sounds within one stream (e.g., XOOOXOOOXOOO...), or they may be heard as two distinct streams, one stream of the higher frequency sound (–OOO–OOO–OOO ...) and another stream of the lower frequency sound (X–X–X–...). Perception of segregation (i.e., two streams) changes the overall perceived rhythm of the sounds (Bregman, 1990; Van Noorden, 1975). When the X tones separate out from the O tones, the infrequent intensity increments occurring in the X–X–X sequence can be detected as violating the standard intensity of the frequent X tones and should elicit MMN.

However, if the intervening 'O' tones with varying intensity are not segregated from the 'X' tones, then the regularity of intensity among the X tones will not be detected and no MMN will be elicited.

There were two different types of "prime" sound sequences that preceded the test sounds (see Fig. 1). In one condition, the prime was intended to bias the organization of the memory representation to two streams (*Two-streams-prime* condition) and in a second condition the prime was intended to bias the organization of the memory representation to one stream (*One-stream-prime* condition). A third condition was conducted in which no prime was presented at all (*No-prime* condition) to show the effect of having no additional sensory influences on the test sounds. In the *Two-streams-prime* condition, the 'prime' sound sequence consisted of 10 'X' tones presented at a 720 ms stimulus onset asynchrony (SOA, onset-to-onset; note that one 'X' tone occurred every 720 ms in the test sounds sequence). The 'X' tones were presented alone to bias the segregation of the X tones to a separate stream. In the *One-stream-prime* condition, the prime sound sequence was similar to the test sequence, except that the frequency of the 'O' sound was 1245 Hz. The frequency difference between 'X' and 'O' tones was decreased to three semitones in the prime to strengthen the integration into one stream. In both conditions, the prime was presented for 7.2 s, and then the test sounds were presented for 14.4 s, alternating in a continuous presentation loop throughout each run with no silent break or demarcation between the prime and the test sounds. In the *No-prime* condition, the test sounds were presented continuously without priming stimuli. Three control conditions were also run to control for stimulus-specific ERP effects that would have resulted from simply comparing tones of two different intensities (see Section 4.4 for further description). The three additional control conditions corresponded to each of the main conditions (*One-stream-prime* condition, *Two-streams-prime* condition, and a *No-prime* condition) with the exception that the deviant and standard tone intensities were reversed. For each of the three main conditions, 9000 stimuli were presented. This yielded (per condition) 225 deviants and 1074 standards (the 'X' tones; of which the priming stimuli and the two standards following each deviant were excluded). For each of the three control conditions, 3000 stimuli were presented, which yielded (per control condition) 75 deviants and 358 standards. The six conditions (three main and three controls) were presented in a randomized order and counterbalanced across subjects in one recording session that took about 2.5 h (including electrode cap placement and breaks).

Participants were comfortably seated in a sound-attenuated booth and instructed to ignore the sounds and watch a silent captioned video of their choice. The experimenter monitored the ongoing electroencephalography (EEG) and electrooculogram (EOG) for eye movement artifacts produced by eye saccades that indicated that participants were actively reading the captions of the silent movie. Regularity of the eye saccade movements throughout the recording sessions, associated with reading the captions, was observed for all participants by the experimenter.

4.3. Data recording

EEG recordings were measured with an electrode cap containing 32 electrodes that approximated the International 10–20 system. Additional electrodes were placed at the left (LM) and right (RM) mastoids. The reference electrode was placed at the tip of the nose (Vaughan and Ritter, 1970). Horizontal eye movements were monitored using bipolar configurations between F7 and F8 and vertical eye movements were monitored with Fp1 and an electrode placed below the left eye. EEG and EOG were digitized at a rate of 500 Hz (.05–100 Hz bandpass) and then filtered off-line (bandpass 1.5–15 Hz). Epochs of 600 ms, which included a 100-ms prestimulus period, were extracted from the continuous EEG recording for each stimulus. Epochs with a voltage change exceeding 75 μ V were assumed to be artifactual and were rejected from further analysis.

4.4. Control for delineating the MMN component

To control for stimulus-specific ERP effects that would have resulted from simply comparing tones of two different intensities, three additional conditions were presented to participants that corresponded to each of the main conditions (*One-stream-prime* condition, *Two-streams-prime* condition, and a *No-prime* condition). In these control conditions, the intensity of the frequent and infrequent 'X' tones was reversed so that the frequent 'X' tone was 83 dB and the infrequent 'X' tones were 71 dB. All other parameters were identical to the corresponding experimental condition. The purpose was to obtain a comparison tone that had the exact same tone intensity when it was a "deviant" in the sequence to that when it was a "standard" in the sequence (83 dB). That is, the probe tone and comparison tone were physically identical. This type of comparison, which leaves only the response related to the role the stimulus played in the sequence (e.g., deviant or standard), results in a reasonable estimation of the intensity MMN (i.e., not derived from refractoriness of the neurons) based upon the memory comparison of intensity within the context that the sounds occur (Jacobsen et al., 2003).

4.5. Data analysis

ERP responses were averaged separately for the infrequent 'X' tone (the probe tone) in each condition, and for the frequent 'X' tone (the control) obtained from each of the corresponding control blocks. Responses evoked by the priming stimuli were not included in any of the analyses. On average, 202 deviant and 308 control responses were retained per subject, per condition. ERPs were re-referenced to the averaged mastoids to maximize the MMN response by including its polarity-inverted manifestation. The average amplitude of the responses to the probe tone and control stimuli was measured separately for each subject in a 40-ms interval centered on the peak MMN (152 ms), identified in the group-average difference waveforms at Fz of the *Two-streams-prime* condition (where MMN was elicited). This interval was then used to measure the MMN in each condition. Repeated measures analysis of variance (ANOVA) with factors of condition (*Two-streams-prime* vs. *No-prime* vs. *One-stream-prime*), stimulus type (control vs. deviant), and electrode (Fz, F3, F4, FC1, FC2) was used to determine whether the mean voltage elicited by the control and probe tone ERPs were significantly different from each other in the interval of the expected MMN latency. These frontocentral electrodes were used to assess MMN because they provide scalp responses with the best signal-to-noise ratio for this component.

Additionally, the mean amplitude of the ERPs elicited by the control tones was measured in each condition, from peak-to-trough, in the individual grand mean waveforms at the Cz electrode (the site of the best signal-to-noise-ratio for auditory obligatory responses). One-way repeated measures ANOVA was used to compare across conditions.

A Greenhouse–Geisser correction was applied to correct for violations of sphericity and the *P* values reported. Post hoc analyses were made with Tukey HSD tests.

4.6. Scalp current density

A reference-free measure of the scalp current density (SCD) was also used to depict the MMN. Maps showing scalp voltage topography and SCD were computed from the mean amplitude difference waveforms for each condition, corresponding to the peak latency of the MMN (152 ms, see Fig. 2). The SCD analysis, an estimate of the second spatial derivative of the voltage potential (the Laplacian), was performed using *BESA 2000 FOCUS* software to sharpen the differences in the scalp fields and obtain additional information about the cortical generators. The SCD maps, expressed in $\mu\text{V}/\text{cm}^2$, show the scalp areas where the current either emerges (sources) from the brain into the scalp or enters (sinks) from the scalp into the brain. This facilitates interpretation of the spatiotemporal contributions from multiple overlapping sources.

Acknowledgments

This research was supported by the National Institutes of Health (R01 DC 04263). We thank Judith Kreuzer for data collection and Fakhar Khan for his assistance with manuscript preparation.

References

- Alho K. Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear Hear* 1995;16 (1):38–51. [PubMed: 7774768]
- Bey C, McAdams S. Postrecognition of interleaved melodies as an indirect measure of auditory stream formation. *J Exp Psychol Hum Percept Perform* 2003;29 (2):267–279. [PubMed: 12760614]
- Biermann S, Heil P. Parallels between timing of onset responses of single neurons in cat and of evoked magnetic fields in human auditory cortex. *J Physiol* 2000;84 (5):2426–2439.
- Blake DT, Merzinich MM. Changes of AI receptive fields with sound density. *J Neurophys* 2002;88:3409–3420.
- Bregman, AS. *Auditory Scene Analysis*. MIT Press; MA: 1990.
- Brosch M, Schreiner CE. Time course of forward masking tuning curves in cat primary auditory cortex. *J Neurophys* 1997;77:923–943.
- Brosch M, Schreiner CE. Sequence sensitivity of neurons in cat primary auditory cortex. *Cereb Cortex* 2000;10:1155–1167. [PubMed: 11073865]
- Brosch M, Schulz A, Scheich H. Neuronal mechanisms of auditory backward recognition masking in macaque auditory cortex. *NeuroReport* 1998;9 (11):2551–2555. [PubMed: 9721931]
- Bushnell MC, Goldberg ME, Robinson DL. Behavioral enhancement of visual responses in monkey Cerebral cortex: I. Modulation in posterior parietal cortex related to selective visual attention. *J Neurophys* 1981;4:755–772.
- Carlyon RP, Cusack R, Foxton JM, Robertson IH. Effects of attention and unilateral neglect on auditory stream segregation. *J Exp Psychol Hum Percept Perform* 2001;27 (1):115–127. [PubMed: 11248927]
- Chapin JK, Woodward DJ. Modulation of sensory responsiveness of single somatosensory cortical cells during movement and arousal behaviours. *Exp Neurol* 1981;72 (1):164–178. [PubMed: 7202621]
- Chelazzi L, Miller EK, Duncan J, Desimone R. A neural basis for visual search in inferior temporal cortex. *Nature* 1993;363:345–347. [PubMed: 8497317]
- Compte A, Brunel N, Goldman-Rakic PS, Wang XJ. Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb Cortex* 2000;10 (9):910–923. [PubMed: 10982751]
- Desimone R. Visual attention mediated by biased competition in extrastriate visual cortex. *Philos Trans R Soc Lond, B Biol Sci* 1998;353:1245–1255. [PubMed: 9770219]
- Fishman YI, Arezzo JC, Steinschneider M. Auditory stream segregation in monkey auditory cortex: effects of frequency separation presentation rate and tone duration. *J Acoust Soc Am* 2004;116:1656–1670. [PubMed: 15478432]
- Friedman D, Cycowicz YM, Gaeta H. The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neurosci Biobehav Rev* 2001;25:355–373. [PubMed: 11445140]
- Giard MH, Perrin F, Pernier J, Bouchet P. Brain generators implicated in processing of auditory stimulus deviance: a topographic event-related potential study. *Psychophysiology* 1990;27:627–640. [PubMed: 2100348]
- Halgren E, Baudena P, Clarke JM, Heit G, Liegeois C, Chauvel P, Musolino A. IntraCerebral potentials to rare target and distractor auditory and visual stimuli: I. Superior temporal plane and parietal lobe. *Electroencephalogr Clin Neurophysiol* 1995;94:191–220. [PubMed: 7536154]
- Hillyard SA, Hink RF, Schwent VL, Picton T. Electrical signs of selective attention in the human brain. *Science* 1973;182:177–180. [PubMed: 4730062]
- Hillyard SA, Vogel EK, Luck SJ. Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philos Trans R Soc Lond, B Biol Sci* 1998;353:1257–1270. [PubMed: 9770220]
- Hocherman S, Benson DA, Goldstein MH, Heffner HE, Hienz RD. Evoked unit activity in auditory cortex of monkeys performing a selective attention task. *Brain Res* 1976;117:51–68. [PubMed: 825193]

- Hubel DH, Henson CO, Rupert A, Galambos R. Attention units in the auditory cortex. *Science* 1959;129 (3358):1279–1280. [PubMed: 13658956]
- Jacobsen T, Horenkamp T, Schröger E. Preattentive memory-based comparison of sound intensity. *Audiol Neurotol* 2003;8:338–346. [PubMed: 14566104]
- Jancke L, Mirzazade S, Shah NJ. Attention modulates activity in the primary and the secondary auditory cortex: a functional magnetic resonance imaging study in human subjects. *Neurosci Lett* 1999;266 (2):125–128. [PubMed: 10353343]
- Javitt DC, Steinschneider M, Schroeder CE, Arezzo JC. Role of cortical *N*-methyl-D-aspartate receptors in auditory sensory memory and mismatch negativity generation: implications for schizophrenia. *Proc Natl Acad Sci U S A* 1996;93:11962–11967. [PubMed: 8876245]
- Kauer JA, Malenka RC, Nicoll RA. NMDA application potentiates synaptic transmission in the hippocampus. *Nature* 1988;334:250–252. [PubMed: 2840582]
- Kropotov JD, Näätänen R, Sevostianov AV, Alho K, Reinkainen K, Kropotova OV. Mismatch negativity to auditory stimulus change recorded directly from the human temporal cortex. *Psychophysiology* 1995;32:418–422. [PubMed: 7652119]
- Kropotov JD, Alho K, Näätänen R, Ponomarev VA, Kropotova OV, Anichkov AD, Nechaev VB. Human auditory-cortex mechanisms of preattentive sound discrimination. *Neurosci Lett* 2000;280:87–90. [PubMed: 10686384]
- Luck SJ, Chelazzi L, Hillyard SA, Desimone R. Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of Macaque visual cortex. *J Neurophysiol* 1997;77:24–42. [PubMed: 9120566]
- Malone BJ, Scott BH, Semple MN. Context-dependent adaptive coding of interaural phase disparity in the auditory cortex of awake macaques. *J Neurosci* 2002;22:4625–4638. [PubMed: 12040069]
- Moran J, Desimone R. Selective attention gates visual processing in the extrastriate cortex. *Science* 1985;229:782–784. [PubMed: 4023713]
- Näätänen R. *Attention and Brain Function*. Erlbaum; Hillsdale, NJ: 1992. p. 494
- Näätänen R, Alho K. Mismatch negativity (MMN)—The measure for central sound representation accuracy. *Audiol Neurotol* 1997;2:341–353. [PubMed: 9390839]
- Näätänen R, Winkler I. The concept of auditory stimulus representation in cognitive neuroscience. *Psychol Bull* 1999;125:826–859. [PubMed: 10589304]
- Näätänen R, Schröger E, Karakas S, Tervaniemi M, Paavilainen P. Development of a memory trace for a complex sound in the human brain. *NeuroReport* 1993;4:503–506. [PubMed: 8513127]
- Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I. Pre-attentive cognitive processing (“primitive intelligence”) in the auditory cortex as revealed by the mismatch negativity (MMN). *Trends Neurosci* 2001;24:283–288. [PubMed: 11311381]
- Nelson RJ, Li B, Douglas VD. Sensory response enhancement and suppression of monkey primary somatosensory cortical neurons. *Brain Res Bull* 1991;27 (5):751–757. [PubMed: 1756394]
- Opitz B, Mecklinger A, von Cramon DY, Kruggel F. Combining electrophysiological and hemodynamic measures of the auditory oddball. *Psychophysiology* 1999;36 (1):142–147. [PubMed: 10098390]
- Opitz B, Schröger E, von Cramon DY. Sensory and cognitive mechanisms for preattentive change detection in auditory cortex. *Eur J Neurosci* 2005;21 (2):531–535. [PubMed: 15673452]
- Petkov CI, Kang X, Alho K, Bertrand O, Yund EW, Woods DL. Attentional modulation of human auditory cortex. *Nat Neurosci* 2004;7 (6):653–658.
- Picton T, Alain C, Otten L, Ritter W, Achim A. Mismatch negativity: different water in the same river. *Audiol Neurotol* 2000;5:111–139. [PubMed: 10859408]
- Pulvirenti L. Neural plasticity and memory: towards an integrated view. *Funct Neurol* 1992;7:481–490. [PubMed: 1338432]
- Rinne T, Antila S, Winkler I. MMN is unaffected by top-down predictive information. *NeuroReport* 2001;12:2209–2213. [PubMed: 11447336]
- Ritter W, Sussman E, Deacon D, Cowan N, Vaughan HG Jr. Two cognitive systems simultaneously prepared for opposite events. *Psychophysiology* 1999;36:835–838. [PubMed: 10554596]
- Sams M, Hari R. Magnetoencephalography in the study of human auditory information processing. *Ann N Y Acad Sci* 1991;620:102–117. [PubMed: 2035937]

- Scherg M, Vajsa J, Picton TW. A source analysis of the late human auditory evoked potentials. *J Cogn Neurosci* 1989;1:336–355.
- Schreiner CE, Mendelson JR, Sutter ML. Functional topography of cat primary auditory cortex: representation of tone intensity. *Exp Brain Res* 1992;92 (1):105–122. [PubMed: 1486946]
- Schröger E. Automatic detection of frequency change is invariant over large intensity range. *NeuroReport* 1994;5 (7):825–828. [PubMed: 8018858]
- Steinschneider M, Volkov IO, Fishman YI, Oya H, Arezzo JC, Howard MA III. Intracortical responses in human and monkey primary auditory cortex support a temporal processing mechanism for encoding of the voice onset time phonetic parameter. *Cereb Cortex* 2005;15:170–186. [PubMed: 15238437]
- Sussman E, Gumenyuk V. Organization of sequential sounds in auditory memory. *NeuroReport* 2005;16 (13):1519–1523. [PubMed: 16110282]
- Sussman E, Ritter W, Vaughan HG Jr. Attention affects the organization of auditory input associated with the mismatch negativity system. *Brain Res* 1998a;789:130–138. [PubMed: 9602095]
- Sussman E, Ritter W, Vaughan HG Jr. Predictability of stimulus deviance and the mismatch negativity system. *NeuroReport* 1998b;9:4167–4170. [PubMed: 9926868]
- 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, Čeponienė R, Shestakova A, Näätänen R, Winkler I. Auditory stream segregation processes operate similarly in school-aged children as adults. *Hear Res* 2001;153:108–114. [PubMed: 11223301]
- Sussman E, Winkler I, Huutilainen M, Ritter W, Näätänen R. Top-down effects on the initially stimulus-driven auditory organization. *Cogn Brain Res* 2002a;13:393–405.
- Sussman E, Winkler I, Kreuzer J, Saher M, Näätänen R, Ritter W. Temporal integration: intentional sound discrimination does not modify stimulus-driven processes in auditory event synthesis. *Clin Neurophysiol* 2002b;113:909–920.
- Sussman E, Winkler I, Schröger E. Top-down control over involuntary attention-switching in the auditory modality. *Psychon Bull Rev* 2003a;10:630–637. [PubMed: 14620357]
- Sussman E, Winkler I, Wang W. MMN and attention: competition for deviance detection. *Psychophysiology* 2003b;40:430–435. [PubMed: 12946116]
- Sussman E, Bregman AS, Wang WJ, Khan FJ. Attentional modulation of electrophysiological activity in auditory cortex for unattended sounds in multi-stream auditory environments. *Cogn Affect Behav Neurosci* 2005a;5 (1):93–110. [PubMed: 15913011]
- Sussman E, Wong R, Horváth J, Winkler I, Wang W. The development of the perceptual organization of sound by frequency separation in 5–11 year-old children. 2005b manuscript under review.
- Tiitinen H, May P, Reinikainen K, Näätänen R. Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature* 1994;372:901–902.
- Van Noorden, LPAS. Unpublished doctoral dissertation. Eindhoven University of Technology; 1975. Temporal coherence in the perception of tone sequences.
- Vaughan HG Jr, Ritter W. The sources of auditory evoked responses recorded from the human scalp. *Electroencephalogr Clin Neurophysiol* 1970;28:360–367.
- Winkler I, Karmos G, Näätänen 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, Czigler I, Sussman E, Horváth J, Balázs L. Pre-attentive binding of auditory and visual stimulus features. *J Cogn Neurosci* 2005;17 (2):320–339. [PubMed: 15811243]
- Woldorff M, Hillyard S. Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalogr Clin Neurophysiol* 1991;79:170–191. [PubMed: 1714809]
- Woldorff, M.; Hansen, JC.; Hillyard, S. Evidence for effects of selective attention in the mid-latency range of the human auditory event-related potential. In: Johnson, R., Jr; Rohrbaugh, JW.; Parasuraman, R., editors. *Current Trends in Event-Related Potential Research*. Elsevier Science; Amsterdam: 1987. p. 146-154.

- Woldorff M, Hillyard S, Gallen CC, Hampson SR, Bloom FE. Magnetoencephalographic recordings demonstrate attentional modulation of mismatch-related neural activity in human auditory cortex. *Psychophysiology* 1998;35:283–292. [PubMed: 9564748]
- Woodruff PW, Benson RR, Bandettini PA, Kwong KK, Howard RJ, Talavage T, Belliveau J, Rosen BR. Modulation of auditory and visual cortex by selective attention is modality-dependent. *NeuroReport* 1996;7 (12):1909–1913. [PubMed: 8905690]

Fig. 1.

Schematic diagram of the stimulus paradigm. The prime-types are displayed in the left panel, preceding the test sounds. The dotted line is for display purposes only. Sound sequences were continuous with no demarcation between the prime and test sounds. The prime and test sequences were alternated in a blocked fashion; each condition was presented separately. The gray scale denotes the intensity value of the tones ('X' tones were 71 dB; probe tones were 83 dB; 'O' tones were 67 dB, 75 dB, 79 dB, and 87 dB. Notice that the 'X' tones had neither the highest nor lowest intensity value within the range presented). The ordinate shows sound frequency in hertz (Hz). The abscissa indicates time in milliseconds. The probe tone is shown with an arrow. Notice that the infrequently occurring probe tone has a higher intensity than the other frequently occurring 'X' tones.

Fig. 2.

ERP responses and Scalp current density (SCD) maps. (A) The grand-averaged ERP responses to the probe tones (thick line) and control tones (thin line) are displayed for Fz. The area of significant difference is shaded in gray and denotes the mismatch negativity (MMN) component (difference between the responses elicited by the probe and control tones). MMN was only elicited when the Two-streams-prime preceded the test sounds, demonstrating that the memory underlying MMN generation was biased by the priming stimuli. (B) Grand mean scalp potential (top row) and SCD maps (bottom row) of the difference waveforms are shown for the three conditions. Negative voltage and current sinks are shown in red, and positive voltage and currents sources are shown in blue (stippled). Circles represent the scalp locations

of the 32 electrodes. Typical MMN topography is seen only in the *Two-streams-prime* condition (peak latency 152 ms). The arrow points to the Fz electrode, the site of the ERP traces shown in panel A.

Difference Waveforms

Intensity decrement control

Fig. 3.

Difference waveforms of the (intensity decrement) control conditions at the Fz electrode site. The difference waves for the three reverse-block control conditions (Two-streams-prime [solid line], One-stream-prime [dashed line], and No-prime [dotted line] conditions) were obtained by subtracting the ERP response elicited by the frequent 'X' tones (71 dB) of the main conditions from their respective ERP responses elicited by the intensity decrement probe tones (71 dB) in the control conditions. An MMN (peak latency 228 ms) was present only in the Two-streams-prime condition (indicated with an arrow).