



Published in final edited form as:

Eur J Neurosci. 2009 January ; 29(1): 205–211. doi:10.1111/j.1460-9568.2008.06561.x.

Neural representations of auditory input accommodate to the context in a dynamically changing acoustic environment

Torsten Rahne^{1,2,3} and Elyse Sussman³

¹ Department of Experimental Audiology and Medical Physics, Otto-von-Guericke-University, Magdeburg, Germany

² Department of Otolaryngology, Martin-Luther-University, Halle-Wittenberg, Germany

³ Department of Neuroscience, Albert Einstein College of Medicine, Bronx, NY, USA

Abstract

The auditory scene is dynamic, changing from 1 min to the next as sound sources enter and leave our space. How does the brain resolve the problem of maintaining neural representations of the distinct yet changing sound sources? We used an auditory streaming paradigm to test the dynamics of multiple sound source representation, when switching between integrated and segregated sound streams. The mismatch negativity (MMN) component of event-related potentials was used as index of change detection to observe stimulus-driven modulation of the ongoing sound organization. Probe tones were presented randomly within ambiguously organized sound sequences to reveal whether the neurophysiological representation of the sounds was integrated (no MMN) or segregated (MMN). The pattern of results demonstrated context-dependent responses to a single tone that was modulated in dynamic fashion as the auditory environment rapidly changed from integrated to segregated sounds. This suggests a rapid form of auditory plasticity in which the longer-term sound context influences the current state of neural activity when it is ambiguous. These results demonstrate stimulus-driven modulation of neural activity that accommodates to the dynamically changing acoustic environment.

Keywords

auditory stream segregation; event-related potentials; mismatch negativity; priming

Introduction

Imagine standing on a busy city street corner. The auditory scene is dynamically changing from 1 min to the next as cars, buses, car horns and voices pass and enter your space. These concurrent sound sources contribute to the acoustic signal that enters the ears. How does the brain resolve the problem of maintaining neural representations of distinct yet changing sound sources? A key function of the auditory system involves the ability to integrate sequentially presented sound elements that belong together (integration) and segregate those that come from different sound sources (segregation). Previous research has studied processes related to the auditory scene analysis problem separately, addressing how the focus of attention influences change detection (Sussman *et al.*, 1998, 2005; Jones *et al.*, 1999; Carlyon *et al.*, 2001) or by determining what parametrical factors influence the sound object detection in the absence of

Correspondence: Dr T. Rahne, E-mail: torsten.rahne@medizin.uni-halle.de.

*Present address: Universitätsklinikum Halle (Saale), Klinik für HNO-Heilkunde, Magdeburger Straße 12, 06112 Halle (Saale), Germany.

attention (Sussman *et al.*, 1999; Ritter *et al.*, 2000; Sussman, 2005; Snyder *et al.*, 2006; Sussman & Steinschneider, 2006; Rahne *et al.*, 2007; De Sanctis *et al.*, 2008). Here, we tested the dynamics of context-change detection (switching the context from integrated to segregated) by examining the electrophysiological response to a single sound occurring within the ongoing background of a changing sequence of sounds. The current study thus investigated how neural representations of auditory input accommodate the changing multi-stream acoustic environment that is common in everyday life.

Auditory change detection is a useful method for testing the dynamics of context-dependent modulation because determining whether a sound represents a change depends on the previous context of the sounds, not on the specific parameters of any particular sound (Sussman & Steinschneider, 2006; for review, see Sussman, 2007). The mismatch negativity (MMN) component of event-related brain potentials (ERPs) reflects the outcome of a change detection process (Picton *et al.*, 2000; Näätänen *et al.*, 2001; Sussman, 2007). MMN is elicited by the violation of detected auditory regularities, and can thus be used to determine which regularities (individual features or pattern of sounds) are represented at the time a sound occurs. In this way, MMN reveals effects of the auditory context on change detection (Sussman *et al.*, 1999, 2007; Sussman & Winkler, 2001; Winkler *et al.*, 2001; Atienza *et al.*, 2003; Müller *et al.*, 2005; Sussman & Steinschneider, 2006; Rahne *et al.*, 2007). A specific tone can be used as a 'probe' of the neural representation that is extracted from the ongoing sound input in different contexts because the standard represented in memory determines what is deviant in the incoming signal (Sussman *et al.*, 2007). Further, these context-dependent changes can be assessed without active perception of the sound sequence because MMN is elicited even when the sounds have no relevance to ongoing behaviour. Its elicitation does not require participants to actively detect the deviant sounds (Ritter *et al.*, 1999; Rinne *et al.*, 2001; Sussman *et al.*, 2003).

The study had two main goals, to determine whether priming ambiguous input would occur in an ongoing, dynamic fashion, and to test the time course of the switch from integrated to segregated sound streams. To test the dynamics of sound organization, subjects were presented with two tones of different frequencies (called 'X' and 'O') in a fixed sequential pattern (XOOOXOOOXOOO...) that was perceptually ambiguous for hearing one or two sound streams (Fig. 1). Additional sets of sounds (the 'prime') were presented prior to the 'test' sounds (Fig. 2) to bias the ambiguity to be resolved as either one stream (integrated organization) or two streams (segregated organization). 'Probe' tones (an infrequent higher intensity 'X' tone) were embedded within the test and prime blocks. The paradigm was designed so that higher intensity probe ('X') tones could only be detected as 'deviant' when the 'X' and 'O' tones segregated into separate frequency streams. The higher intensity 'X' tones would elicit MMN. In contrast, if the 'X' and 'O' tones were detected as part of the same stream, the probe tones occurred equiprobably among the six different tone intensities (see Materials and methods for details) and would not elicit MMN. Thus, in the present study, MMN elicitation indicates that the 'X' and 'O' tones were neurophysiologically segregated.

Materials and methods

Subjects

Thirteen adults (four males) between the ages of 19 and 45 years (mean = 29 years) were paid to participate in the study. All participants passed a hearing screen (threshold of 20 dB HL or better from 250 to 4000 Hz), and had no reported history of hearing or neurological problems. Participants gave informed consent after the procedures were explained to them, in accordance with the human subject research protocol approved by the Committee for Clinical Investigations at the Albert Einstein College of Medicine, where the study was conducted. The procedures conform to the Code of Ethics of the World Medical Association (Declaration of

Helsinki). The data from one subject were excluded from analysis due to excessive eye artefacts.

Stimuli and procedures

Three sinusoidal tones of 50 ms duration (5 ms rise/fall times) were created with Neuroscan STIM software™ (Compumedics, El Paso, TX, USA) and presented binaurally through insert earphones (E-a-rtone® 3A, Aearo, IN, USA). There was one 'test' sequence and two 'priming' sequences. For the test sequence, two tones were presented in a fixed pattern (XOOOXOOOXOOO...), where 'X' represents a tone of 1046 Hz and 'O' represents a tone of 1397 Hz [a 5-semitone (ST) frequency difference (Δf) between them]. The onset-to-onset pace was 180 ms. The intensity level of the 'X' tones was 71 dB (87.2%). The probe 'X' tone was 83 dB SPL (12.8%). The intensity value of the intervening 'O' tones varied above and below the 'X' tone intensity values in four steps (67 dB, 75 dB, 79 dB and 87 dB SPL), and were equiprobably distributed (Fig. 1). Thus, the 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. The standard intensity could therefore not be distinguished as the softest sounds and the deviant intensity could not be distinguished as the loudest sound in the overall sequence. All sounds were calibrated with a Brüel and Kjaer sound-level meter (2209) with an artificial ear.

The test sequence [T] was perceptually ambiguous for hearing one or two sound streams. That is, the sounds may be heard as a repeating four-tone pattern of sounds within one stream (e.g. XOOOXOOOX-OOO...), 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-...; Fig. 1). This was determined in a separate behavioural study, in which listeners heard two distinct streams approximately 50% of the time at a 5-ST frequency separation (Sussman *et al.*, 2007). Similar ambiguity in the perception of a sound sequence, comprised of one integrated or two segregated streams, was found with different paradigms but similar Δf between sounds (Carlyon *et al.*, 2001; Bey & McAdams, 2003). The ambiguous sequence allows us to test whether the priming sequences can influence the test sounds to be physiologically represented as one- or two-streams. This is accomplished by randomly presenting probe tones in the 'X' tone stream. When the 'X' tones physiologically separate out from the 'O' tones, the intensity of the 'X' tones is the same (71 dB), allowing for the infrequent probe tones with intensity increments (83 dB) to be detected as 'deviant' from the standard intensity of the 'X' tones. However, if the intervening 'O' tones, which vary in intensity above and below the 'X' tone values, are not segregated from the 'X' tones, then there is no intensity regularity in the integrated sequence. Thus, the probe tones will elicit MMN only when the 'X' and 'O' tones are physiologically segregated.

The prime sequence that was intended to bias the test sequence toward segregation [PrimeSeg] presented the 'X' tones alone, at the rhythm they occurred within the test sequence (i.e. once every 720 ms; Fig. 2). The prime sequence that was intended to bias the test sequence toward integration [PrimeInt] presented the XOOOXOOO pattern of sounds at 180 ms, except that the frequency value of the 'O' tones was 1245 Hz, making the Δf between 'X' and 'O' tones smaller (3 ST) than in the test sequence (Fig. 2). A smaller Δf was used to strengthen integration of tones to the XOOOXOOO pattern, heard when the 'X' and 'O' tones are integrated into one stream.

The probe tones in the test block were presented randomly amongst all the 'X' tones. The probe tones in the two priming blocks were presented randomly as one of the first three tones and one of the last three 'X' tones. The test blocks were presented for 14.4 s and the prime blocks were presented for 7.2 s.

The prime and test blocks were presented in an eight-block design: PrimeInt-Test-PrimeInt-Test-PrimeSeg-Test-PrimeSeg-Test (Fig. 2). For the purpose of discussion, each of the eight segments will be labelled according to their role as prime or test [e.g. (P) or (T)], along with their intended bias [e.g. toward integration (1) or segregation (2)] and position of the type of block [i.e. (a) for the first test following prime; (b) for the second test following the same type of prime]. Thus, each part of the eight-block design had a unique label as follows: P1a – T1a – P1b – T1b – P2a – T2a – P2b – T2b (Fig. 2). Six repetitions of the eight-block configuration, without any silent breaks between them, were presented in one run.

To control for stimulus-specific effects of increased tone intensity for the probe tone, additional control runs were presented in which the intensities of the standard (71 dB SPL) and probe (83 dB SPL) 'X' tones were reversed and all other parameters kept the same. This was done so that the MMN could be delineated using the same physical stimulus, subtracting the ERP when it was a standard (83 dB) from when it was a deviant (83 dB). In total, 15 runs (13 experimental and two controls), yielding 312 deviants (probe tones) of each prime block type (P1a, P2a, P1b, P2b) and 400 deviants of each test block type (T1a, T2a, T1b, T2b), were presented in a counterbalanced order. The total session time, including breaks and electrode cap placement, was approximately 3 h.

Participants were seated in a sound-attenuated booth, instructed to ignore the sounds and watch a silent, captioned video of their choice. The experimenter monitored the electroencephalogram (EEG) for regular eye saccades to ensure that participants were watching the movies and reading the captions.

Data recording

EEG were continuously recorded with Neuroscan Synamps (Compumedics) AC-coupled amplifiers (0.05–100 Hz bandwidth; sampling rate: 500 Hz) using a 32-channel electrode cap (Electrocap) placed according to the modified International 10–20 System (Fpz, Fz, Cz, Pz, Oz, Fp1, Fp2, F7, F8, F3, F4, Fc5, Fc6, Fc1, Fc2, T7, T8, C3, C4, Cp5, Cp6, Cp1, Cp2, P7, P8, P3, P4, O1, O2) plus the left and right mastoids. The reference electrode was placed on the tip of the nose. Impedances were kept below 5 k Ω . Vertical electrooculogram (EOG) was recorded with a bipolar electrode configuration using Fp2 and an external electrode placed below the right eye. Horizontal EOG was recorded using the F7 and F8 electrodes.

Data analysis

EEG was filtered off-line (bandpass 1–15 Hz), using epochs of 600 ms, which included a 100-ms prestimulus period that was extracted from the continuous EEG. Epochs in which voltage changes exceeded 75 μ V were rejected from further analysis. On average, the rejection rate due to artefacts was 10.5%. The remaining epochs were averaged according to stimulus type (probe and matching intensity-control tones), within each block type separately. After baseline correction, the average amplitudes of the responses were measured separately for each subject in a 30-ms interval centred on the MMN peak, which were identified in the group-average difference waveforms at Fz. To statistically assess the presence of the MMN in the test blocks that followed the integrated prime blocks (P1b, T1b), in which no MMN was visually present, the time interval from the similar block in which MMN was visually present was used (P1a, T1a). The peak latencies were 132 ms (P1a, T1a, P1b, and T1b), 162 ms (P2a), 186 ms (T2a), 152 ms (P2b) and 188 ms (T2b).

A repeated-measures analysis of variance (anova) with factors of stimulus type (control vs. probe), electrode (Fz, Cz, F3, F4, FC1, FC2, C3, C4) and block type (P1a, P1b, T1a, T1b, P2a, P2b, T2a, T2b) 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 MMN or

expected MMN latency. The fronto-central electrodes were used to provide the best signal-to-noise ratio for MMN. The difference waveforms were also calculated (subtracting the standard ERP elicited by control tones from the deviant ERP elicited by probe tones) and the MMN amplitude statistically compared in repeated-measures anova with factors of block (prime vs. test) and time ('a' vs. 'b').

The amplitude of the obligatory ERP response was also assessed to determine any effect of block type (prime or test). The mean voltage in a 30-ms window centred on the peak latency of the N1 of the control stimuli for each block type was measured and then analysed in a repeated-measures ANOVA with factors of time ('a' vs. 'b'), block type (prime vs. test) and organization (segregated vs. integrated). Greenhouse–Geisser correction was applied to correct for violations of sphericity. *Post hoc* tests were made with Tukey HSD.

Results

P1 and N1 waveforms elicited by both probe and control tones in each of the eight block types were clearly observed (Fig. 3). The amplitude of the N1 evoked by the control tones did not significantly differ among them (block type: $F_{7,77} = 0.74$, $P > 0.1$; electrode: $F_{7,77} = 2.6$, $P > 0.1$; and no interaction: $F_{49,539} = 0.94$, $P > 0.1$).

Comparison of the mean amplitudes of the ERPs in the observed or expected MMN peak latency revealed main effects of the stimulus type ($F_{1,11} = 26.14$, $P < 0.001$), block type ($F_{7,77} = 5.57$, $P < 0.01$), and a significant interaction between stimulus type and block type ($F_{7,77} = 2.38$, $P < 0.05$). No other interactions were significant. *Post hoc* analyses of the interaction revealed that the ERPs elicited by the probe tones were significantly more negative than the control in all conditions except for the second presentation of the PrimeInt (P1b) and Test (T1b) blocks. Thus, significant MMN components were elicited by the probe tones in the two PrimeSeg and their respective following Test blocks, the first PrimeInt and its following Test block. MMNs were not significantly present in the second PrimeInt and its following Test block (Fig. 4).

The amplitude of the difference waveforms, delineated by subtracting the ERP elicited by the control tones from the ERP elicited by the probe tones, in the second blocks of the PrimeInt and Test (P2b/T2b) were significantly smaller than those in all the other blocks (consistent with no MMN elicited in those blocks). MMN amplitude was larger in the first blocks of PrimeSeg and Test (P2a/T2a) than the second (P2b/T2b) (significant interaction between time [a vs. b] and organization [integrated vs. segregated]: $F_{1,11} = 5.15$, $P < 0.05$). In the blocks in which MMN was present, the amplitude did not significantly differ ($P > 0.1$).

Discussion

The results of this study demonstrate that neurophysiological indices of deviance detection accommodate to the dynamically changing acoustic environment. When presented with ambiguous sound input, previous sounds influence the organization of new sounds, maintaining the organization that was current. Even though the immediate environment of the probe tones was the same in all of the test blocks, the prime tones modulated the response to the probe tones within the test blocks toward integration or segregation. In a previous study, we demonstrated that previous sounds could modulate ambiguous input (Sussman & Steinschneider, 2006; see also Snyder *et al.*, 2008). The current results extend the Sussman and Steinschneider findings by showing that priming effects on ambiguous input occur dynamically, changing from moment to moment according to the changing sound context. However, the effects were not symmetrical.

Another goal of the current study was to examine the time course of the modulation to determine when the switching occurred from integrated to segregated (or from segregated to integrated). Therefore, probes were also included in the priming stimuli to examine the time course of the switch from MMN elicitation (neurophysiologically indicating segregated streams) to no MMN elicitation (indicating one integrated stream). We expected a somewhat symmetrical switch, which would occur within the duration (7 s) of the priming blocks. However, we found that the segregated organization was maintained through the first presentation of the PrimeInt block and the following Test block (Fig. 2, #1–2, P1a-T1a), evidenced by probe tones eliciting significant MMNs in these blocks. Not until the second presentation of the PrimeInt block (Fig. 2, #3, P1b) was no MMN elicited by the probes, indicating that the organization had then switched to an integrated stream. In contrast, the switch from the integrated to the segregated organization occurred immediately after the first presentation of the PrimeSeg block: MMN was elicited by the probe tones in the following Test block (Fig. 2, #6, T2a). Thus, our results demonstrate an asymmetrical switching from two- to one-stream, with the neurophysiological representation showing a rapid switch going from integrated to segregated (one to two streams) and a more gradual switch from segregated to integrated (two to one stream). Thus, the segregated organization was held onto memory longer than the integrated organization.

This asymmetrical switch may be due to asymmetry in the strength of the priming blocks. The PrimeInt blocks may have been effectively more subtle because there was only a slight decrease in the frequency separation between the priming tones and the test tones (i.e. still within the ‘ambiguous’ range rather than strongly integrated). In contrast, the PrimeSeg blocks presented only one stream by itself, which may have been more effective in segregating out the ‘X’ tone stream for distinct representation from within the ambiguously organized test block.

Another possibility is that a 5-ST difference is still relatively large for integration to occur automatically. This is suggested by similar asymmetries that have been reported in the psychophysical literature. A commonly used paradigm asks listeners to hold onto an integrated ‘galloping’ rhythm across frequency spans, showing that it is easier to segregate sounds than it is to hold onto the integration of sounds when they are in the ambiguous region (van Noorden, 1975; Bregman, 1990; Carlyon *et al.*, 2001). The Δf for hearing integration when the listener is trying to hear integration is smaller than the Δf that is required to hear segregated streams when the listener is trying to segregate sounds. That is, segregation can be held onto at much smaller frequency separations than integration can be held onto at larger frequency separations. Our neurophysiological evidence of an asymmetry is consistent with this behavioural data, even though no task was being performed with the sounds. Thus, our results point toward an additional explanation that is related to changes in the underlying neural activity contributing to the observed asymmetry. The indication is that neurophysiological changes, driven by the ongoing dynamic changes in stimulus context, interact with or occur in conjunction with the listener’s attentional set to bias the global organization.

A recent psychophysical study manipulated the listener’s attention to sounds and found context effects on auditory stream segregation (Snyder *et al.*, 2008). Snyder *et al.* asked subjects to judge, trial to trial, whether they heard one or two streams. When the current trial was in the ambiguous range (e.g. 3 or 6 ST), current judgements were influenced by the previous one. When subjects reported ‘two streams’ on a previous trial, they were more likely to judge 3 or 6 ST as two streams, whereas when they reported ‘one stream’ on a previous trial, they were more likely to judge the current trial (3 or 6 ST) as one stream (Experiment 3). This bias only occurred for intermediate range of Δf . When the current trial was at one of the extremes (e.g. 0 ST or 12 ST), the Δf of previous trials had no effect on perception of the sounds as one or two streams (i.e. 0 ST was always heard as one stream and 12 ST was always heard as streaming). These results are consistent with our results in that there was a general propensity to hold onto the previous stream organization when judging ambiguously organized incoming

sounds. However, their results also differed from ours in that they found a ‘contrast effect’, in which subjects reported hearing 7 ST as two streams less often when preceded by 12 ST, and more often as two streams when preceded by 3 ST (Experiments 1 and 2). Thus, the judgement of the current trial was influenced in contrast to the previous one as opposed to concordant with it. One aspect of their experimental design may account for this contrast effect, namely that they had a silent period between trials, which we did not have. Silence would cause the buildup of stream segregation to start anew for each trial (Cusack *et al.*, 2004; Sussman *et al.*, 2007). Thus, the influence of the previous trial was not immediate on the current trial, as each trial started from an ‘integrated’ percept. Second, the silence could allow for a ‘same-different’ approach in judging each trial. When the previous trial was clearly segregated, a smaller Δf on the current trial could seem less segregated in contrast, and vice versa. Thus, other cognitive strategies for judging the current trial may have influenced the type of context effects that were found in Snyder *et al.* Further studies are needed to determine how different attentional strategies interact with stimulus-driven factors to alter the neurophysiological response to sounds.

The current results, showing a dynamic process of adapting the neural response to the current state, may indicate that there are stimulus-driven feedback mechanisms that modulate neural activity of new input. Such feedback models have been demonstrated during attentional modulation of receptive fields in auditory cortex of animals (Weinberger, 2004; Fritz *et al.*, 2007; Rahne *et al.*, 2008). The current results in humans suggest that these feedback mechanisms may not necessarily be initiated by attention, but can also be initiated by stimulus-driven factors.

Overall, we demonstrated context-dependent responses to a single tone that was modulated in a dynamic fashion as the auditory environment rapidly changed. The current memory representation formed the basis for evaluating new, incoming sensory inputs, indicative of a context-dependent neural process that biases the way neural traces encode incoming sensory information in auditory memory. This suggests a rapid form of auditory plasticity in which the longer-term sound context influences the current state of neural activity when it is ambiguous. This flexible, adaptive process would be needed in everyday situations to maintain stable auditory representations of the environment.

Acknowledgements

This research was supported by the National Institutes of Health (DC 004263) and by the Deutsche Forschungsgemeinschaft (DFG) (SFB/TRR 31 ‘The active auditory system’).

Abbreviations

EEG	electroencephalogram
EOG	electrooculogram
ERP	event-related potentials
MMN	mismatch negativity
ST	semitone

References

- Atienza M, Cantero JL, Grau C, Gomez C, Dominguez-Marin E, Escera C. Effects of temporal encoding on auditory object formation: a mismatch negativity study. *Brain Res Cogn Brain Res* 2003;16:359–371. [PubMed: 12706216]
- Bey C, McAdams S. Postrecognition of interleaved melodies as an indirect measure of auditory stream formation. *J Exp Psychol* 2003;29:267–279.
- Bregman, AS. *Auditory Scene Analysis: The Perceptual Organization of Sounds*. MIT Press; Cambridge, MA: 1990.
- 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;27:115–127. [PubMed: 11248927]
- Cusack R, Deeks J, Aikman G, Carlyon RP. Effects of location, frequency region, and time course of selective attention on auditory scene analysis. *J Exp Psychol Hum Percept Perform* 2004;30:643–656. [PubMed: 15301615]
- De Sanctis P, Ritter W, Molholm S, Kelly SP, Foxe JJ. Auditory scene analysis: the interaction of stimulation rate and frequency separation on pre-attentive grouping. *Eur J Neurosci* 2008;27:1271–1276. [PubMed: 18364041]
- Fritz JB, Elhilali M, David SV, Shamma SA. Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in A1? *Hear Res* 2007;229:186–203. [PubMed: 17329048]
- Jones D, Alford D, Bridges A, Tremblay S, Macken B. Organizational factors in selective attention: the interplay of acoustic distinctiveness and auditory streaming in the irrelevant sound effect. *J Exp Psychol Learn Mem Cogn* 1999;25:464–473.
- Müller D, Widmann A, Schröger E. Auditory streaming affects the processing of successive deviant and standard sounds. *Psychophysiology* 2005;42:668–676. [PubMed: 16364062]
- Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I. ‘Primitive intelligence’ in the auditory cortex. *Trends Neurosci* 2001;24:283–288. [PubMed: 11311381]
- van Noorden, LPAS. *Temporal Coherence in the Perception of Tone Sequences*. Eindhoven University of Technology; Eindhoven, The Netherlands: 1975. Unpublished Dissertation
- Picton TW, Alain C, Achim A, Otten L, Ritter W. Mismatch negativity: different water in the same river. *Audiol Neurootol* 2000;5:111–139. [PubMed: 10859408]
- Rahne T, Böckmann M, von Specht H, Sussman ES. Visual cues can modulate integration and segregation of objects in auditory scene analysis. *Brain Res* 2007;1144:127–135. [PubMed: 17306232]
- Rahne T, Deike S, Selezneva E, Brosch M, König R, Scheich H, Böckmann M, Brechmann A. A multilevel and cross-modal approach towards neuronal mechanisms of auditory streaming. *Brain Res* 2008;1220:118–131. [PubMed: 17765207]
- Rinne T, Antila S, Winkler I. Mismatch negativity is unaffected by top-down predictive information. *Neuroreport* 2001;12:2209–2213. [PubMed: 11447336]
- Ritter W, Sussman ES, Deacon D, Cowan N, Vaughan HG Jr. Two cognitive systems simultaneously prepared for opposite events. *Psychophysiology* 1999;36:835–838. [PubMed: 10554596]
- Ritter W, Sussman ES, Molholm S. Evidence that the mismatch negativity system works on the basis of objects. *Neuroreport* 2000;11:61–63. [PubMed: 10683830]
- Snyder JS, Alain C, Picton TW. Effects of attention on neuroelectric correlates of auditory stream segregation. *J Cogn Neurosci* 2006;18:1–13. [PubMed: 16417678]
- Snyder JS, Carter OL, Lee SK, Hannon EE, Alain C. Effects of context on auditory stream segregation. *J Exp Psychol Hum Percept Perform* 2008;34:1007–1016. [PubMed: 18665741]
- Sussman ES. Integration and segregation in auditory scene analysis. *J Acoust Soc Am* 2005;117:1285–1298. [PubMed: 15807017]
- Sussman E. A new view on the MMN and attention debate: Auditory context effects. *J Psychophysiology* 2007;21:164–175.
- Sussman E, Steinschneider M. Neurophysiological evidence for context-dependent encoding of sensory input in human auditory cortex. *Brain Res* 2006;1075:165–174. [PubMed: 16460703]
- Sussman E, Winkler I. Dynamic sensory updating in the auditory system. *Brain Res Cogn Brain Res* 2001;12:431–439. [PubMed: 11689303]

- Sussman E, Ritter W, Vaughan HG Jr. Attention affects the organization of auditory input associated with the mismatch negativity system. *Brain Res* 1998;789:130–138. [PubMed: 9602095]
- Sussman ES, 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 ES, Winkler I, Wang W. MMN and attention: competition for deviance detection. *Psychophysiology* 2003;40:430–435. [PubMed: 12946116]
- Sussman ES, Bregman AS, Wang WJ, Khan FJ. Attentional modulation of electrophysiological activity in auditory cortex for unattended sounds within multistream auditory environments. *Cogn Affect Behav Neurosci* 2005;5:93–110. [PubMed: 15913011]
- Sussman ES, Horváth J, Winkler I, Orr M. The role of attention in the formation of auditory streams. *Percept Psychophys* 2007;69:136–152. [PubMed: 17515223]
- Weinberger NM. Specific long-term memory traces in primary auditory cortex. *Nat Rev Neurosci* 2004;5:279–290. [PubMed: 15034553]
- Winkler I, Schröger E, Cowan N. The role of large-scale memory organization in the mismatch negativity event-related brain potential. *J Cogn Neurosci* 2001;13:59–71. [PubMed: 11224909]

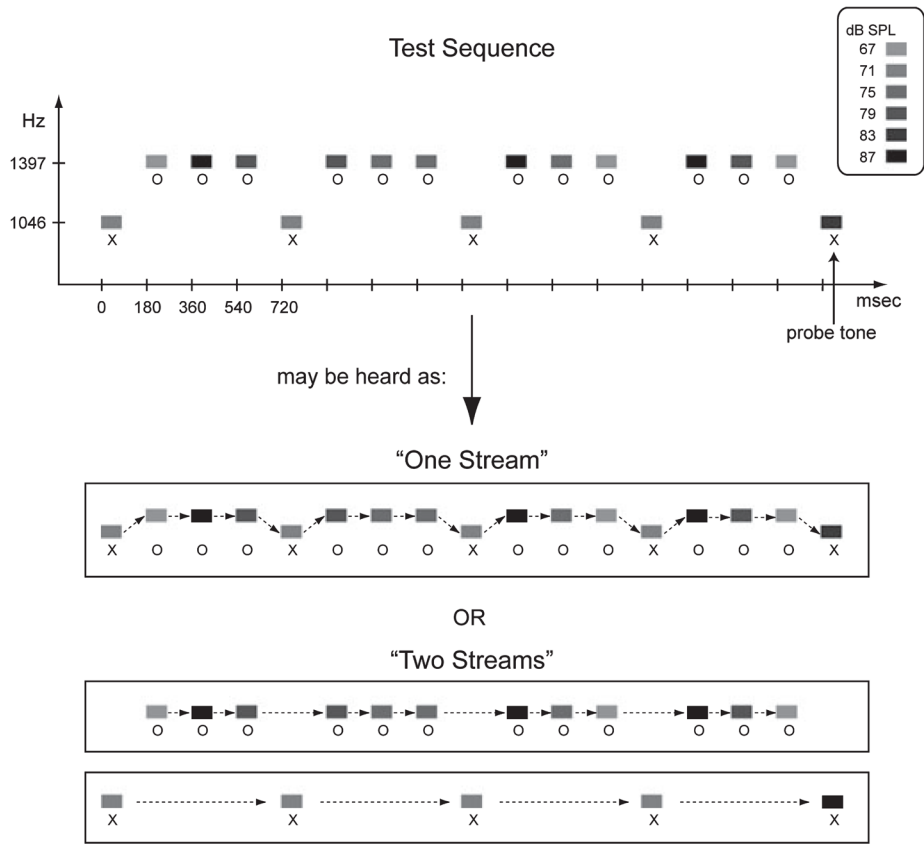


Fig. 1. Schematic diagram of the ambiguous test sequence (top). The y-axis represents Hz and the x-axis represents time in ms. The ticks show the onset pace. Rectangles represent two different tones ('X' and 'O'), denoting the two different frequency values. Grey-scale denotes intensity values, which range from 67 to 87 dB in 4-dB steps. The probe tone (indicated with an arrow) is an 'X' tone of 83 dB. The Δf between the 'X' and 'O' tones is ambiguous for hearing integration or segregation. The sound sequence is perceived equally often as one (integrated) stream (XOOOXOOO...) or two (segregated) streams (one made up of the X tones and the other made up of the O tones, X—X—X— and -OOO—OOO—OOO) (Sussman *et al.*, 2007). Only when the 'X' tones are segregated to separate streams is the louder intensity probe tone detected as deviant. When the sounds are integrated, the louder 'X' tone is absorbed into a single stream, i.e. as one of six different intensities that is not a deviant. Thus, the response to the probe tone indicates whether the ambiguous sounds are integrated or segregated.

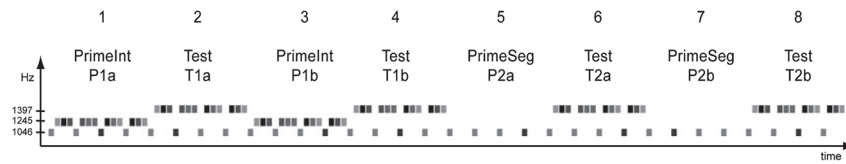


Fig. 2.

Schematic diagram of the stimulus paradigm showing the order (labelled #1–8, top-most row) of the Integrated (1) and Segregated (2) priming (P) blocks that are alternated with the test (T) blocks in the eight-block sequence. Prime blocks are named to indicate whether they intend to bias the test block toward integration or segregation (PrimeInt and Prime Seg). Each prime-then-test block was presented twice (a, b) to assess the time course of the responses to probe tones within. Thus, e.g. T1b (fourth block, second row) denotes the second test block that was primed toward an integrated organization. Grey-scale indicates intensity variation (see also Fig. 1). The y-axis shows the tone frequency in Hz. The x-axis is time.

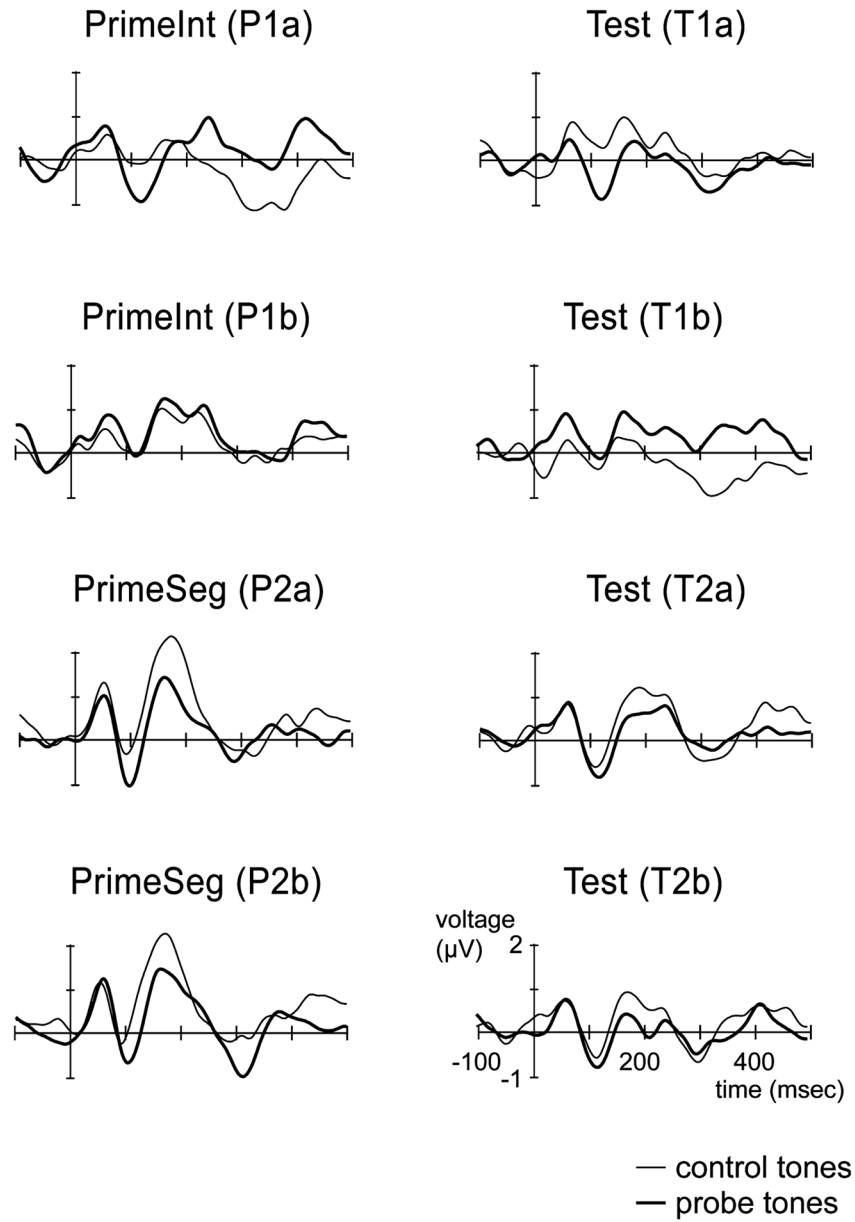


Fig. 3. Grand-mean ERP responses to the control tones (thin line) and probe tones (thick line) of the Prime (left column) and Test (right column) blocks are depicted at electrode Fz. P1 and N1 ERP components elicited by both probe and control tones in each of the eight block types are clearly observed.

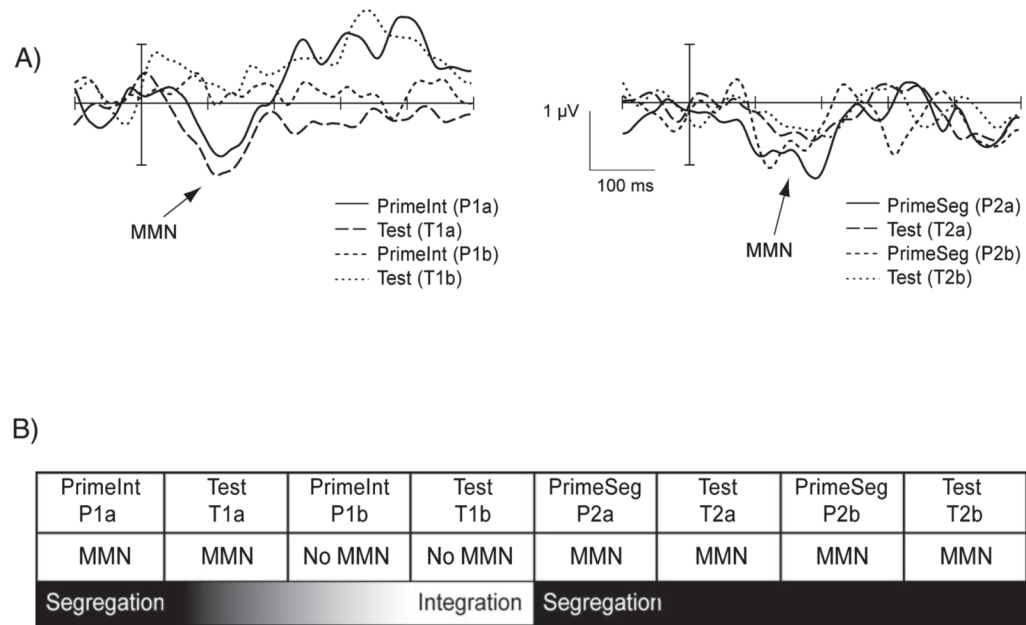


Fig. 4. (A) Grand-mean difference waveforms (delineated by subtracting the ERP elicited by the control tones from the ERP elicited by the probe tones) are displayed for all block types for the Fz electrode. Mismatch negativities (MMNs; indicated with arrows) were elicited by probes in the PrimeSeg blocks (P2a, P2b, right side, thick and thin solid lines) and their Test blocks (T2a, T2b, right side, thick and thin dashed lines). MMNs were also elicited by probe tones in the first PrimeInt (P1a) and Test (T1a) blocks (left side, thin solid and dashed lines), but not in the second presentation (P1b, T1b, left side, thick solid and dashed lines). (B) The asymmetry of the transitions is schematically displayed. There was a rapid transition from integrated to segregated, and a slower transition from segregated back to integrated (see Discussion).