Published in final edited form as:

Neuropsychologia. 2014 November ; 64: 218–229. doi:10.1016/j.neuropsychologia.2014.09.039.

Effects of task-switching on neural representations of ambiguous sound input

Elyse S. Sussman^{a,*}, Albert S. Bregman^b, and Wei-Wei Lee^a

^aAlbert Einstein College of Medicine, Bronx, NY, USA

^bMcGill University, Montreal, Quebec, Canada

Abstract

The ability to perceive discrete sound streams in the presence of competing sound sources relies on multiple mechanisms that organize the mixture of the auditory input entering the ears. Many studies have focused on mechanisms that contribute to integrating sounds that belong together into one perceptual stream (integration) and segregating those that come from different sound sources (segregation). However, little is known about mechanisms that allow us to perceive individual sound sources within a dynamically changing auditory scene, when the input may be ambiguous, and heard as either integrated or segregated. This study tested the question of whether focusing on one of two possible sound organizations suppressed representation of the alternative organization. We presented listeners with ambiguous input and cued them to switch between tasks that used either the integrated or the segregated percept. Electrophysiological measures indicated which organization was currently maintained in memory. If mutual exclusivity at the neural level was the rule, attention to one of two possible organizations would preclude neural representation of the other. However, significant MMNs were elicited to both the target organization and the unattended, alternative organization, along with the target-related P3b component elicited only to the designated target organization. Results thus indicate that both organizations (integrated and segregated) were simultaneously maintained in memory regardless of which task was performed. Focusing attention to one aspect of the sounds did not abolish the alternative, unattended organization when the stimulus input was ambiguous. In noisy environments, such as walking on a city street, rapid and flexible adaptive processes are needed to help facilitate rapid switching to different sound sources in the environment. Having multiple representations available to the attentive system would allow for such flexibility, needed in everyday situations to maintain stable auditory percepts, and to allow rapid scanning of interesting events in a busy environment.

Keywords

Auditory scene analysis; Stream segregation; Mismatch negativity (MMN); Auditory perception; Task-switching

*Correspondence to: Departments of Neuroscience and Otorhinolaryngology, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461 USA. Tel.: +1 718 430 3313. elyse.sussman@einstein.yu.edu (E.S. Sussman).

1. Introduction

In natural situations, the auditory environment is dynamically changing, with multiple sound sources overlapping in time. Imagine yourself walking on a busy city street as noises from the environment constantly change: a jet flies overhead; a jackhammer makes repeated sound bursts; cars drive by; people are talking as they walk past you. The neural representation or model of the auditory environment must account for the dynamics of such situations as the population of sounds in the auditory scene can change from moment to moment. Most studies investigating the neurobiological basis of auditory scene analysis (the process that allows us to hear distinct auditory sound events in noisy environments) have focused on how the brain disentangles two fixed sets of sound that differ in frequency (Bregman, 1978; Brochard et al., 1999; Carlyon et al., 2001; Cusack et al., 2004; Micheyl et al., 2005; Müller et al., 2005; Rahne et al., 2007; Rahne and Sussman, 2009; Shamma et al., 2011; Sussman, 2005; Sussman et al., 2007a, 2007b; Sussman-Fort and Sussman, 2014; Sussman et al., 1999; Sussman and Steinschneider, 2006; Szalárdy et al., 2013). However, most of the input to our ears overlaps dynamically in time and is often not clearly disambiguated. The current study investigated the problem of how ambiguous input is resolved by the auditory system, and what systems mediate perception of individual sound events when there are competing sound streams. We tested the hypothesis that the attended sound organization ‘wins’ out for neural representation when the input is ambiguous for hearing a single integrated stream or two segregated streams. Thus, we asked the question of whether focusing on one of two possible organizations (integrated or segregated) results in suppression of the alternative organization, or instead whether both organizations can be represented even when only one organization appears in perception.

Previously, we have used the mismatch negativity component (MMN) of event-related brain potentials (ERPs) as an index of sound organization (Sussman et al., 1998; Sussman et al., 1999; Sussman, 2013; Sussman, 2007). The MMN is elicited by discriminability of a deviation from a detected sound regularity maintained in the neural trace of the input (Näätänen et al., 2001; Näätänen et al., 2014). Any deviant can elicit MMN as long as the memory trace of the regularity is represented (Sussman, 2007). Our previous paradigms were set up such that within-stream regularities would only emerge when the sounds were neurophysiologically segregated, permitting detection of within-stream deviants (e.g., intensity or pattern deviants) (Rahne and Sussman, 2009; Sussman and Steinschneider, 2006; Sussman et al., 2005; Sussman, 2005). These studies allowed us to index sound organization, but the results could not distinguish whether both possible organizations were present simultaneously. This was because the MMN was elicited by deviant tones only when the sounds were segregated into distinct streams. We had no direct assessment of the integrated organization. For the current study, we modified a previous paradigm (Sussman and Steinschneider, 2006) to provide two distinct indices, one evoked only by integration and the other only by segregation (Fig. 1). Attention was cued to switch between performing one of two different tasks: a loudness detection task that required segregation of the higher from the lower tones (Fig. 1B), and a pattern identification task that required integration of the entire sequence (Fig. 1C). Thus, the demands of the task required either segregation of the sounds to perceive an oddball in one of the streams or integration of the sounds to

perceive patterns of the stimuli. The task dictated which of the two possible organizations was to be active in perception. The MMN response was used to index whether a loudness change (e.g., intensity deviant) and/or a pattern deviant were detected based on different regularity representations from the same input.

The question of how ambiguous input is represented in auditory cortex has been of interest in the last decade (Sterzer et al., 2009; Nelken, 2004; Pressnitzer and Hupé, 2006; Sussman and Steinschneider, 2006; Rahne and Sussman, 2009; Denham et al., 2014; Winkler et al., 2009; Sussman, 2010). Much of the work has studied the temporal dynamics of a phenomenon called auditory bi-stability, in which spontaneous switching between two or more perceptual states occurs when the input is ambiguous. This is akin to the visual face-vase illusion (or binocular rivalry) in which perception of one excludes perception of the other at any given time even though switching back and forth between the two arises spontaneously. With long presentations of sounds, participants record the points at which they hear integrated, segregated, or some other organization of the input (e.g., Pressnitzer and Hupé, 2006; Denham et al., 2013; Denham et al., 2014; Winkler et al., 2012). The measures of the spontaneous switching between perceptual states across time indicate that whereas multiple alternative organizations of the input are proposed to be simultaneously maintained, one dominates in perception for a period of time (before switching to another) depending on the stimulus-driven characteristics of the input (Denham et al., 2013; Denham et al., 2014; Sterzer et al., 2009; Pressnitzer and Hupé, 2006). These results suggest that multiple sound organizations are available when the input is ambiguous and that spontaneous perceptual switching occurs as a result of low-level competition between the potential sound organizations (Denham et al., 2013; Horváth et al., 2001; Nelken, 2004; Pressnitzer and Hupé, 2006; Sterzer et al., 2009; Winkler et al., 2012). Although most of these studies report perception arising from spontaneous switching behavior, when participants were instructed to volitionally group (integrate) or segregate the sounds, it did not increase the duration of the designated perceptual state across trials but rather decreased the proportion of reported trials for the irrelevant organization (Pressnitzer and Hupé, 2006). Thus, it is not fully clear how attention modulates the ambiguous case to facilitate task performance. The results of the Pressnitzer and Hupé (2006) study, for example, suggest this may occur by reducing the intrusion of the unattended sound organization rather than by strengthening the attended one.

The current study tested effects of top-down processing on the storage of neural representations for ambiguous input. Attention was manipulated to control which organization was required to perform the task, while the elicitation of the MMN component was used as an index of which organization(s) was present during task performance. Thus, we assessed, using unique triggers for integration and segregation, whether one or both of the potential organizations were neurophysiologically represented while participants actively integrated or actively segregated the sound input to perform a task. This study thus asks the question of whether task performance modulates neural representations of ambiguous sound input toward the organization that is used to perform the task, and minimizes the irrelevant organization. What happens to the alternative, unattended organization when attention is focused on one set of the sounds?

If mutual exclusivity at the neural level was the rule, attention to one of two possible organizations would preclude neural representation of the other. This would predict that when attention biased the neurophysiological response toward the organization needed to perform the task, MMN would be elicited only by the deviants produced by the organization that was induced by the task, suppressing the alternative, unattended organization. Alternatively, if both organizations were represented, regardless of the task performed, then MMNs would be elicited by unattended non-target deviants of the organization that was not required by the task. Thus, we determine if both organizations are represented in the neural trace or if attention ‘resolves’ the ambiguity toward the organization that is used to perform the task.

2. Experiment 1: ambiguous (5 ST) condition

2.1. Methods and materials

2.1.1. Participants—Fifteen healthy young adults (10 male) aged 18–33 years old ($M = 26$ years, $SD = 4$) participated in the study. All participants passed a hearing screen (20 dB HL for pure tones at 500, 1000, 2000, and 4000 Hz, bilaterally). Participants gave written consent and were paid for their participation. Three participants could not perform at least one of the tasks. Their data were excluded from analyses. The remaining 12 participants’ data are included in this report.

2.1.2. Stimuli—Stimuli were created using Neuroscan software (STIM, Compumedics, Corp., Charlotte, NC). Two pure tones (50 ms. duration, 5 ms. rise/fall time) were presented bilaterally through insert earphones (E-A-R-tone® 3A, Indianapolis, IN). Sounds were calibrated for peak-to-peak equivalent sound pressure level (ppe SPL) using a Brüel & Kjær sound level meter (2209) with an artificial ear (4152). There was a 5 semitone [ST] distance between tones: the higher frequency tone (H) was 1397 Hz and the lower frequency tone (L) was 1046 Hz (Fig. 1A). Tones were presented in three patterns: HLHH (Pattern1), HHLH (Pattern2), and HHHL (Pattern3), randomly distributed, with Pattern 1 occurring 40%, Pattern 2 occurring 40% and Pattern 3 occurring 20% of the time (Fig. 1A). Stimulus onset asynchrony (SOA) from high tone to high tone was 220 ms; SOA from high tone to low tone was 110 ms; and 440 ms separated the four-tone patterns (Fig. 1A). Thus, the inter-stimulus interval (ISI) for the high tone stream (the interval between successive triplets of the high tones) was isochronous, but the ISI for the low tone stream was jittered.

2.1.3. Logic of the stimulus design—The MMN component is elicited by deviants detected as violating a previously detected regularity in a stimulus sequence that is neurophysiologically represented in auditory memory. The present experiment was designed so that one type of deviant would be encountered if listeners integrated all the tones into a single perceptual stream (pattern deviants), and a different type of deviant would be encountered if they segregated the high and low tones into separate streams (loudness deviants). In this way, elicitation of the MMN component would indicate whether neurophysiological representations of one or more organizations were held in memory simultaneously. If all the tones were experienced in a single stream the three patterns would be present (Fig. 1C). In this case, Pattern 3 would be the infrequent (deviant) pattern, and

could elicit MMN if the integrated organization were held in memory. On the other hand, if the high and low tones formed separate streams, the H tones would split from the L tones, forming triplets, and the L tones would comprise a stream in which the oddball was an intensity change (12% randomly presented 81 dB vs. 88% 69 dB L tones, Fig. 1B) with a jittered ISI. In the segregated percept, the 81 dB tones are deviant with respect to the 69 dB L tones in the same stream. The H tones vary in intensity above and below the L tone intensity values, with the four intensity levels of the high tones (65, 73, 77, and 85 dB) occurring equally often in a randomized order (Fig. 1B), so that if all tones were heard as a single stream there would be a variety of intensity levels, none of them standing out as deviant. Therefore, only when the L tones split by frequency into their own stream can the loudness detection task be performed (i.e., comparison between the 69 and 81 dB tones). The infrequent louder intensity low tone was the MMN-eliciting deviant when sounds were segregated. Pattern 3 was the MMN eliciting deviant when the sounds were integrated.

Because louder intensity sounds can produce larger obligatory components (e.g., N1) that can overlap with the MMN response, a control condition was presented to obtain a comparison tone for the louder intensity deviant that had the same physical characteristics and could serve as the comparison tone to the louder deviant intensity. This was done by reversing the intensity value of the frequent-infrequent low tones (i.e., 81 dB for the frequent L tones and 69 dB for the infrequent) so that the 81 dB tones were not deviant under this control condition. All other parameters were the same as under the main condition, including task requirements. Thus, under the control condition participants pressed the response key to the softer intensity deviant when they performed the loudness detection task. 672 patterns were presented under the control condition.

2.1.4. Procedures—Participants sat in a comfortable chair in front of a computer monitor during the experiment. There were two tasks: pattern identification and loudness detection. The pattern identification task was performed by pressing the response key for the designated pattern (Fig. 1C). There was no indication made to participants about deviant patterns. The participants' task was only to identify and press the response key to the designated pattern (P1, P2, or P3). For the loudness detection task, participants were instructed to segregate out the low tones, focus on their intensity value, ignore the higher tones, and press the response key when the louder L tones were detected in the low tone stream (Fig. 1B). Trials were randomly intermingled, visually cueing the participant as to which task to perform on the upcoming trial (Fig. 1D). Practice was provided for both tasks prior to the experiment. There were 30 trials for each different task. Each trial consisted of 42 patterns, each trial lasting approximately 36 s. The visual cue was presented on screen for 625 ms, and the sounds for the trial began in 1.375 s from the offset of the visual cue (Fig. 1D). Total session time was approximately 2.5 h, which included electrode cap placement and breaks.

2.1.5. Electroencephalogram (EEG) recording—EEG was recorded using an electrode cap with the following electrode sites: Fpz, Fp1, Fp2, Fz, F3, F4, F7, F8, FC1, FC2, FC5, FC6, Cz, C3, C4, CP1, CP2, CP5, CP6, T7, T8, Pz, P3, P4, P7, P8, Oz, O1, O2 (modified 10–20 system, American Electroencephalographic Society, 1991) and the left and

right mastoids (LM and RM). Fp1 and an electrode placed below the left eye in a bipolar configuration was used to monitor the vertical electro-oculogram (VEOG). An electrode placed on the tip of the nose was used as the reference. PO9 was used as the grounding electrode. EEG was digitized with a 500 Hz sampling rate (0.05–100 Hz band pass), and then band pass filtered off-line between 0.5 and 30 Hz. Epochs were 700 ms, starting 100 ms before stimulus onset. Artifact rejection criterion was set at $\pm 75 \mu\text{V}$ on all electrodes after baseline correction was applied to the entire epoch. Stimuli were averaged separately in each trial by status (intensity deviant, pattern deviant, standard) regardless of task. Comparison ERPs for the 81 dB SPL intensity deviants when the loudness detection task was performed were obtained by averaging together the evoked response to the 81 dB SPL 1046 Hz tones from the control blocks. Comparison ERPs for the Pattern 3 target (deviant) when the pattern identification task was performed were obtained by averaging together the ERPs to the Pattern 3 trials when they were non-targets, in the trials when Pattern 1 and Pattern 2 were the targets.

2.1.6. Data analysis

2.1.6.1. Behavioral data analysis: Reaction time (RT), hit rate (HR), false alarm rate (FAR), and d' measures were calculated for all targets.

Responses were considered correct in the loudness detection task if button presses occurred between 40 and 900 ms from onset of the deviant intensity 1046 Hz (L) tones, and RT measured from the onset of the deviant. Misses were counted as the absence of a button press to deviants. Correct rejections were the absence of button presses to the non-deviant L tones, and false alarms were button presses to any non-deviant L tone. False alarm rate was calculated using the total number of non-deviant L tones, not the total number of non-deviant stimuli. d prime was calculated using HR and FAR based on these data.

For the pattern identification task, correct responses were recorded if button presses were made to the designated target pattern between 40 and 900 ms from onset of the first tone of the pattern for P1, the second tone of the pattern for P2, and the third tone of the pattern for P3. RT was calculated, separately, from the onset of each respective tone for the three patterns. False alarms were recorded if button presses occurred to either of the other two patterns than the target (e.g., if P1 was the target, responses to P2 and P3 were considered false alarms). Correct rejections were the absence of responses to the non-target patterns and misses were the absence of response to the designated target pattern. d prime was calculated using these data for HR and FAR.

One-way repeated measures Analysis of Variance (ANOVA) were calculated on d' and RT to assess task effects. Greenhouse–Geisser corrected p values were reported. Tukey HSD was used for post-hoc calculations.

2.1.6.2. ERP data analysis: To obtain amplitude measurements for statistical analysis, the peak latency of the ERP components (MMN and P3b) were identified from the grand-mean difference waveforms (ERP evoked by the deviant-minus- ERP evoked by the comparison standard). The electrode of known maximal signal-to-noise ratio was used to choose the peak latency (Fz for MMN and Pz for P3b) in each task, and deviant type

separately. Mean amplitude was then measured in each individual using a 50 ms interval centered on the peak. Four-way repeated-measures analysis of variance (ANOVA) with factors of task type (loudness detection, pattern identification), attention (attended, unattended), stimulus type (deviant, standard), and electrode position (Fz, Cz and Pz) were used to determine presence of components and scalp topography, and effects of task type and attention on elicitation of the components. Greenhouse–Geisser corrections for sphericity were applied and the p values reported. Post-hoc analyses were calculated using Tukey HSD.

2.2. Results

2.2.1. Behavioral results—Table 1 shows the HR, FAR, and d' for the ambiguous (5 ST) condition of Experiment 1, displayed for each target separately. There was a significant main effect of trial type on RT ($F_{3,33} = 36.61$, $\varepsilon = 0.65$, $p = 0.0001$). Post-hoc calculations revealed that RT was significantly shorter to the Pattern 3 targets than all the other targets, with no significant differences among the other targets. There was a significant main effect of trial type on d' ($F_{3,33} = 4.73$, $\varepsilon = 0.56$, $p = 0.027$). Post-hoc calculations showed that sensitivity to the target as measured by d' was lower to the Pattern 2 target compared to the Pattern 3 target. No other comparisons were significantly different from each other.

2.2.2. ERP results—Fig. 2 shows the ERPs evoked by the deviant and control standard stimuli under the Ambiguous condition (left) when loudness was attended (top) and when the patterns were attended (bottom). Table 2 shows the mean amplitudes of the ERP waveforms elicited by standard and deviants for attended and unattended targets, used to measure the MMN. Table 3 shows the mean amplitudes of the ERP waveforms elicited by standard and deviants for attended and unattended targets, used to measure the P3b component.

Fig. 3 shows the difference waveforms and scalp voltage distribution, delineating the MMN (top row) and P3b (bottom row) components.

2.2.2.1. MMN component: The key result of this study, which addresses the question of whether multiple representations are simultaneously maintained in memory, was that MMN was elicited by unattended deviants when either task was performed (i.e., intensity deviants when the pattern task was performed and pattern deviants when the loudness detection task was performed) (Fig. 3). This is consistent with more than one neurophysiological representation being maintained in memory during task performance. Further, there were no significant differences in the MMN amplitude elicited by attended, task-specific (target) deviants or by unattended (non-target) deviants (Fig. 3, top row), whereas the task-specific P3b component was only elicited by attended deviants (Fig. 3, bottom row). Statistics supporting these results are described below. For the task factor, the ‘attended task’ refers to the set of sounds that the participant is using to perform the designated task, and the response to those selected sounds (as described in the Section 2.1).

2.2.2.1.1. Main effects: There was a main effect of task type on amplitude (loudness vs. pattern) ($F_{1, 11} = 12.43$, $p = 0.0048$), with overall mean amplitude being more negative for

the loudness than the pattern task. There was a main effect of stimulus type ($F_{1, 11} = 13.691$, $p = 0.0035$). The amplitude of the deviant was more negative than the standard. This is consistent with MMN elicitation overall (all deviants contributing to the greater negativity). There was a significant main effect of electrode ($F_{2, 22} = 10.41$, $\varepsilon = 0.55$, $p = 0.0007$). Post-hoc calculations revealed that the amplitude at Fz ($-0.52 \mu\text{V}$) was overall more negative than that at Cz ($0.08 \mu\text{V}$) or Pz ($0.38 \mu\text{V}$), with no difference between Cz and Pz. This suggests a frontal scalp distribution for the MMN. There was no main effect of attention on MMN ($F_{1, 11} = 2.56$, $p = 0.138$), meaning that overall amplitude did not change whether the target was attended or unattended.

2.2.2.1.2. Interactions- two-way: There was a significant interaction of task type and attention ($F_{1, 11} = 10.78$, $p = 0.0073$). Post-hoc calculations revealed that this was due to there being no significant task type amplitude difference between loudness ($0.26 \mu\text{V}$) and pattern ($0.14 \mu\text{V}$) when they were non-targets (unattended), but the amplitude for attended loudness task ($-1.06 \mu\text{V}$) was more negative than the attended pattern task ($0.52 \mu\text{V}$). A significant attention by electrode interaction ($F_{2, 22} = 6.27$, $\varepsilon = 0.75$, $p = 0.014$) was due to the amplitude being more negative at Fz than at Cz or Pz (with no difference between Cz and Pz) for attended tasks, and Fz more negative than Cz and Pz, and Cz more negative than Pz for unattended tasks. There was no effect of attention at Fz (no difference between attended and unattended task deviants), only at Cz and Pz, where the amplitude for attended task was more negative than for unattended task at those electrodes. For the significant stimulus by electrode interaction ($F_{2, 22} = 8.7$, $\varepsilon = 0.61$, $p = 0.008$), post-hoc calculations revealed that the deviant at Fz was more negative than at Cz and Pz and no difference between Cz and Pz. There was no difference between standard responses by electrode. Further, the deviant at Fz was more negative than the standard at Fz and the deviant at Cz more negative than the standard at Cz, but no difference between deviant and standard at Pz. This is consistent with MMN topography.

2.2.2.1.3. Interactions- three- and four-way: There was one significant three-way interaction of task type by attention by electrode ($F_{2, 22} = 12.12$, $\varepsilon = 0.54$, $p = 0.00395$). Post-hoc calculations showed that the interaction occurred because the attended amplitudes were more negative than unattended amplitudes for the same task type at Fz and Cz for the loudness task but only at Fz for the pattern task. The four-way interaction (task type by attention by stimulus type by electrode) was also significant ($F_{2, 22} = 12.54$, $\varepsilon = 0.80$, $p = 0.000776$). The post-hoc results confirmed the presence and topography of the MMN for all deviants (i.e., Fz amplitude for deviants more negative than standard for both tasks and both types of attention). The interaction was demonstrated because there was no significant difference between stimulus types at the Pz electrode for either task type or attention state.

There were no other significant interactions (task type $\times$ stimulus type $p = 0.31$; task type $\times$ electrode $p = 0.12$; attention $\times$ stimulus type $p = 0.76$; task type $\times$ attention $\times$ stimulus type $p = 0.53$; task type $\times$ stimulus type $\times$ electrode $p = 0.13$; and attention $\times$ stimulus type $\times$ electrode $p = 0.66$).

To summarize, MMN was elicited by all deviants regardless of what task participants performed. When the loudness task was performed, H and L tones were segregated to

perform the task, yet Pattern 3 deviants elicited significant MMNs (the integrated organization was irrelevant to the task); and when the pattern identification task was performed, H and L tones were integrated to perform the task, yet intensity deviants within the L tones elicited significant MMNs (the segregated organization was irrelevant to the task). This pattern of results suggests that task performance did not abolish the alternative representation available in the ambiguous input. When one task was performed MMN was elicited by both pattern and intensity deviants regardless of the task that participants were performing, indicating that both the integrated and the segregated organizations were neurophysiologically represented in memory during task performance.

2.2.2.2. P3b component: The key finding for the P3b component was that it was elicited only for task-relevant (attended) targets and not for task-irrelevant (unattended) targets regardless of the task type being performed. Further, the scalp voltage distribution was consistent with the P3b component, having maximal amplitude at the Pz electrode. Statistics supporting these results are described below.

2.2.2.2.1. Main effects: There was a main effect of attention ($F_{1,11} = 9.398$, $p = 0.01$), with the amplitude for the attended task ($1.06 \mu\text{V}$) overall more positive than the amplitude for the unattended task ($0.10 \mu\text{V}$). There was a main effect of stimulus type ($F_{1,11} = 18.29$, $p = 0.001$), with the overall amplitude of the deviant ($1.47 \mu\text{V}$) more positive than the overall amplitude of the standard ($-0.31 \mu\text{V}$). There was a main effect of electrode ($F_{2,22} = 50.16$, $\varepsilon = 0.75$, $p.0.0001$). Post-hoc calculations showed that the overall amplitude was most positive at Pz ($\text{Pz} > \text{Cz} > \text{Fz}$). There was no main effect of task type ($F_{1,11} < 1$, $p = 0.51$).

2.2.2.2.2. Interactions, two-way: There was an interaction between attention and stimulus type ($F_{1,11} = 8.84$, $p.0.013$). Post-hoc calculations showed the amplitude of the deviants ($2.7 \mu\text{V}$) to be more positive than the amplitude of the standards ($-0.57 \mu\text{V}$) for the attended tasks, but with no amplitude difference between deviant ($0.25 \mu\text{V}$) and standard ($-0.05 \mu\text{V}$) for the unattended tasks. Also, there was no significant difference between standard stimuli, only between deviants (i.e., attended deviant larger than unattended deviant). This demonstrates the presence of P3b only for the attended deviants. The interaction between attention and electrode ($F_{2,22} = 14.79$, $\varepsilon = 0.56$, $p.0.0001$) was due to the Pz electrode having a greater positive amplitude in the attended vs. unattended at the Pz and Cz electrodes but not at Fz, where there was no amplitude difference. There was an interaction between stimulus type and electrode ($F_{2,22} = 67.86$, $\varepsilon = 0.78$, $p.0.0001$). Post-hoc calculations showed that this interaction was due to there being no significant amplitude difference for the standards ($\text{Fz} = \text{Cz} = \text{Pz}$), and with the amplitude of the deviant most positive at the Pz electrode and larger than the standard at Pz, next largest amplitude of the deviant at Cz and larger than the standard at Cz, with no difference in amplitude between standard and deviant at Fz. There were no interactions between task type and attention ($p = 0.22$); task type and stimulus type ($p = 0.64$); or task and electrode ($p = 0.37$).

2.2.2.2.3. Interactions, three- and four-way: There was a significant three-way interaction between attention, stimulus type, and electrode ($F_{2,22} = 15.24$, $\varepsilon = 0.56$, $p = 0.0016$). Post-hoc tests revealed that this interaction was due to there being no significant

difference between the standards and deviants at any electrodes in the unattended tasks, whereas the deviant was significantly more positive than the standard at Pz (5.12 μ V vs. -0.36 μ V respectively) and at Cz, but not at Fz for the Attended tasks. This is consistent with the presence of the P3b for the attended tasks and its absence in the unattended tasks. There was a significant four-way interaction between task type, attention, stimulus type, and electrode ($F_{2, 22} = 5.91$, $\varepsilon = 0.61$, $p = 0.025$). This result confirmed the presence of a significantly more positive amplitude for the deviants compared to the standards in the attended tasks, with no such difference in the unattended tasks at Cz and Pz, and no such differences for attended or unattended at Fz. These results reflect the topography of the P3b, with amplitude largest at the parietal electrode site elicited for the attended targets and absent for the unattended targets.

There were no other significant interactions (task type, attention and stimulus type ($p = 0.41$); task type, attention and electrode ($p = 0.73$); or task type, stimulus type and electrode ($p = 0.21$)).

To summarize, a significant P3b component was only elicited by task-relevant deviants: intensity deviants elicited P3b when performing the loudness detection task, and Pattern 3 deviants elicited P3b when performing the pattern identification task. No P3b components were elicited by the alternative (unattended) organization during task performance.

3. Experiment 2: unambiguous condition (1 ST & 15 ST)

3.1. Purpose

Whereas the results of the Ambiguous condition (Experiment 1) were clear: multiple representations were maintained in memory regardless of which task was being performed by the participant, the results were also somewhat surprising in light of the theory of mutual exclusivity. Therefore, we conducted a second experiment to confirm that when the stimulus-driven input was stronger toward the integrated (1 ST) or segregated (15 ST) organizations, only one MMN would be elicited by the respective deviant for that organization (i.e., pattern 3 for the integrated organization and intensity for the segregated organization). That is, there should be no MMN to the unrepresented (or less robust) organization. The same procedures were used, but with two different frequency distances between H and L tones to create more clearly segregated and integrated percepts (i.e., the input was not ambiguous as to the organization).

3.2. Methods and materials

3.2.1. Participants—Eleven healthy young adults participated in the study (8 male) aged 21–41 years old ($M = 29$ years, $SD = 5$). Participants gave written informed consent after the procedures were explained to them. All participants passed a hearing screen (20 dB HL at 500, 1000, 2000, and 4000 Hz bilaterally). None of the participants in Experiment 2 participated in Experiment 1. We used a between-subjects design to avoid potential learning or training effects from having performed this task previously.

All recording parameters, procedures, and data analysis techniques were the same as in Experiment 1, except for the frequency distance between the H and L tones. Based on

previous studies by ourselves and others (Bregman, 1990; Carlyon et al., 2001; Sussman et al., 2007a, 2007b; Sussman et al., 2014; Van Noorden 1975), we used a 15 ST difference between H and L tones (H tone was 2489 Hz) to create more strongly (unambiguously) segregated sounds under one condition, and a 1 ST distance (H tone was 1109 Hz) to create unambiguously integrated sounds under the other condition (see also Denham et al. (2013) for factors influencing perception of streaming). The task trials were randomly interwoven throughout stimulus blocks. A visual stimulus occurring before each auditory trial designated which task to perform (pattern identification or loudness detection). Participants performed only the loudness detection task under the 15 ST condition and only the pattern identification task under the 1 ST condition.

3.2.2. Data analysis

3.2.2.1: Behavioral data were analyzed the same way as in experiment 1. One-way repeated measures Analysis of Variance (ANOVA) were calculated on d' and RT to assess task effects. Greenhouse–Geisser corrected p values were reported. Tukey HSD was used for post-hoc calculations.

3.2.2.2: ERP data were analyzed in separate calculations of four-way analysis of variance (ANOVAs) with factors of trial type (15 ST vs. 1 ST), attention (attended vs. unattended), stimulus type (deviant vs. standard), and electrode (Fz, Cz, Pz). The stimulus type factor determines the presence of the ERP component and the electrode factor indicates scalp topography. As a reminder, in the 15 ST trials, MMN elicited by the intensity deviants would index segregation, and in the 1 ST trials, MMN elicited by the pattern 3 deviants index integration. No MMNs were expected to the unattended pattern deviants in the 15 ST trials, or the unattended intensity deviants in the 1 ST trials. The MMN has a fronto-central scalp voltage distribution and the P3b is maximal at the Pz electrode.

3.3. Results

3.3.1. Behavioral results—Table 1 shows the HR, FAR, and d' for the unambiguous (15 ST and 1 ST) conditions of Experiment 2, displayed for each target separately. There was a significant main effect showing differences among the RT scores ($F_{3,30} = 25.03$, $\varepsilon = 0.58$, $p < 0.0001$). Post-hoc calculations revealed that RT was significantly shorter to the Pattern 3 targets than all the other targets, with no significant differences among the other targets. There was a significant main effect showing differences among the d' scores ($F_{3,30} = 22.27$, $\varepsilon = 0.86$, $p < 0.0001$). Post-hoc calculations showed that sensitivity to the loudness target as measured by d' was significantly higher than all the Pattern targets. Sensitivity to P1 targets was not significantly different than to P2 or P3. Sensitivity to the P2 target was significantly lower than to the P3 target.

3.3.2. ERP results—Fig. 2 shows the ERPs evoked by the deviant and control standard stimuli under the Unambiguous conditions (right column) when loudness was attended (top) and when the patterns were attended (bottom). Table 4 shows the mean amplitudes of the ERP waveforms elicited by standard and deviants for attended and unattended targets, used to measure the MMN. Table 5 shows the mean amplitudes of the ERP waveforms elicited by

standard and deviants for attended and unattended targets, used to measure the P3b component.

Fig. 4 shows the difference waveforms and scalp voltage distribution, delineating the MMN (top row) and P3b (bottom row) components.

3.3.2.1. MMN component: The significant four-way interaction ($F_{2,20} = 12.43$, $p = 0.00031$) supports the key findings for MMN elicitation and scalp voltage distribution. Post-hoc calculations revealed that significant MMNs were elicited, the deviant was significantly more negative than the standard at Fz ($-1.89 \mu\text{V}$, $0.004 \mu\text{V}$, respectively) for the attended loudness targets in the 15 ST trials and for the attended pattern targets in the 1 ST trials ($-0.98 \mu\text{V}$, $1.12 \mu\text{V}$, respectively). In contrast, there was no significant difference between deviant and standard for the unattended pattern deviants at Fz in the 15 ST trials ($-0.09 \mu\text{V}$, 0.42 , respectively), or the unattended intensity deviants in the 1 ST trials ($-1.77 \mu\text{V}$, $-1.38 \mu\text{V}$ respectively). At the Cz electrode, the same differences between standards and deviants were observed, but at the Pz electrode, no differences between deviants and standards for any of the deviants in any trials. This pattern of results is consistent with a fronto-central scalp distribution of MMN. Thus, MMNs were elicited by the attended intensity deviants in the 15 ST trials (segregated) and for the attended pattern deviants in the 1 ST trials (integrated), and no MMNs to the unattended deviants.

3.3.2.1.1. Main effects: There was a main effect of stimulus type ($F_{1,10} = 17.11$, $p = 0.002$), due to the deviant amplitude being overall more negative than the standard amplitude. There was a main effect of electrode ($F_{2,20} = 8.72$, $p = 0.0079$). Post-hoc tests revealed that overall the amplitude at the Fz electrode was more negative than that at Cz and Pz electrodes (and with no difference between Cz and Pz). There was no main effect of trial type ($F < 1$, $p = 0.66$) or attention ($F < 1$, $p = 0.83$).

3.3.2.1.2. Interactions, two-way: There was a significant interaction between trial type and attention ($F_{1,10} = 13.08$, $p = 0.0047$). Post-hoc tests showed that the amplitude evoked by the attended trial type was more negative than the amplitude of the unattended trial type. There was a significant interaction between attention and stimulus type ($F_{1,10} = 26.01$, $p = 0.0005$). Post-hoc calculations showed that the amplitude of the deviant was significantly more negative than the amplitude of the standard for the attended but not for the unattended targets. The significant interaction between trial type and electrode ($F_{2,20} = 23.21$, $\epsilon = 0.79$, $p < 0.0001$) was due to a more negative overall amplitude for the responses in the 1 ST trials at Fz but a more positive overall amplitude for the responses in the 1 ST trials at Pz. The stimulus type by electrode interaction ($F_{2,20} = 18.82$, $\epsilon = 0.89$, $p < 0.0001$) was due to the overall amplitude for the deviant stimuli at Fz being significantly more negative than that at Cz, and Cz more negative than Pz ($Fz < Cz < Pz$), and with all deviants being more negative than the standards overall, but with the amplitude of the standards at Fz being more negative than the standards at Cz and Pz. There was no interaction between trial type and stimulus type ($F < 1$, $p = 0.48$). There was a significant three-way interaction between trial type, attention and electrode ($F_{2,20} = 20.01$, $\epsilon = 0.59$, $p < 0.0006$). No other interactions were significant (no interactions between trial type, attention, and stimulus: $F = 1.93$, $p = 0.19$;

trial type, stimulus type, and electrode: $F = 2.05$, $p = 0.18$; attention, stimulus type, and electrode $F < 1$, $p = 0.67$).

3.3.2.2. P3b component: The key finding for the P3b component was the three-way interaction between attention, stimulus type, and electrode ($F_{2,20} = 30.81$, $p < 0.0001$). Post-hoc calculations revealed that the amplitude of the ERP to the deviant stimulus ($4.62 \mu\text{V}$) was significantly more positive than the amplitude of the ERP to the standard stimulus ($-0.26 \mu\text{V}$) at Pz for the attended tasks (as well as a significant difference between deviant and standard at Cz and Fz) but with no significant difference between deviant and standard amplitudes at any electrode for the unattended tasks. This demonstrates elicitation of the P3b component only for task-relevant deviants.

3.3.2.2.1. Main effects: There was a significant main effect of trial type ($F_{1,10} = 42.76$, $p < 0.001$), with loudness task amplitude more positive than pattern task amplitude overall; a main effect of attention ($F_{1,10} = 8.54$, $p = 0.015$) with the attended task amplitude more positive than the unattended task amplitude overall; a main effect of stimulus type ($F_{1,10} = 25.93$, $p < 0.005$), with the ERP amplitude of the deviant more positive than the ERP amplitude of the standard overall; and a main effect of electrode ($F_{2,20} = 39.39$, $\epsilon = 0.96$, $p < 0.0001$), with post-hoc calculations showing that the amplitude was most positive at the Pz electrode ($\text{Pz} > \text{Cz} > \text{Fz}$).

3.3.2.2.2. Interactions, two-way: Significant interactions were found between trial type and attention ($F_{1,10} = 19.52$, $p = 0.013$); and between trial type and stimulus ($F_{1,10} = 15.49$, $p = 0.028$), with post-hoc calculation showing that the greater positivity comes from the deviants and not the standards (no significant difference between standards). There was an interaction between attention and stimulus type ($F_{1,10} = 17.15$, $p = 0.002$). Post-hoc calculations showed that the difference between deviant and standard comes from the attended task and not the unattended (no significant difference between unattended deviant and standard). There was also an interaction between attention and electrode ($F_{2,20} = 22.25$, $\epsilon = 0.67$, $p = 0.0002$); and between stimulus type and electrode ($F_{2,20} = 20.33$, $\epsilon = 0.78$, $p < 0.0001$). The interaction between trial type and electrode was not significant ($F < 1$, $p = 0.74$).

3.3.2.2.3. Interactions, three- and four-way: Another significant three-way interaction occurred between trial type, attention and electrode ($F_{2,20} = 6.28$, $\epsilon = 0.73$, $p = 0.016$). There were no other significant three- or four-way interactions (trial type, attention and stimulus type, $F_{1,10} = 2.34$, $p = 0.16$); trial type, stimulus type and electrode, $F < 1$, $p = 0.66$); or among all factors, $F_{2,20} = 2.39$, $\epsilon = 0.62$, $p = 0.144$).

4. General discussion

This study tested the hypothesis that attention would resolve ambiguous auditory input by neurophysiologically representing in memory the sound organization used to perform a task. Specifically, we sought to determine whether focusing on one of two possible organizations means that the neurophysiological representation of the other is suppressed, or whether both organizations are represented even when only one organization appears in perception. The

key finding was that MMN was elicited by organizational deviants specific to both organizations, integration (pattern deviants) and segregation (loudness deviants), regardless of the task being performed (Fig. 3). This suggests that task performance does not exclude alternative organizations for ambiguous input.

It might be expected that performing a task would enhance or bring to the foreground the organization used to perform the task, while suppressing or diminishing the unattended, alternative organization (Mesgarani and Chang, 2012). However, that was not the case. The unattended organization not needed to perform the task was still accessible, evoking a change-detection response by deviants specific to the unattended organization. This is consistent with the results of Pressnitzer and Hupé (2006), who found that the alternative organization intruded into perception when actively grouping or segregating sounds. The current results demonstrate that even when one organization dominates in perception, alternative organizations are represented in the neural trace. The perception of one state does not abolish representation of the other state even during the dominant percept. This result is consistent with perceptual switching studies suggesting that multiple potential organizations are maintained simultaneously and compete for representation (Winkler et al., 2012; Denham et al., 2014).

Notably, evidence for both organizations occurred only when the stimulus input was ambiguous (5 ST, Experiment 1) and not when the stimuli had physical characteristics that induced a stronger representation of stream segregation (15 ST distance between H and L tones) or stream integration (1 ST distance between H and L tones). For the unambiguous cases (Experiment 2), deviants elicited MMN only with respect to the attended task that matched the stimulus organization used to perform the task. There was no indication that the opposite organization was neurophysiologically represented (deviants of the unattended organization did not elicit significant MMNs).

There was evidence that the alternative organization was detected and neurally represented in the ambiguous case because deviants (non-targets) elicited MMNs. However, the deviants of the alternative organization were not processed as alternative targets. This was shown by the absence of a response-related P3b component to the deviants of the alternative sound organization. Even though they elicited MMNs, the target-related P3b component was only elicited by the deviants that were also targets (Fig. 3, bottom panel, thick line) and were not elicited by unattended, non-targets of the alternative organization (Fig. 3, bottom panel, thin line).

Taken together, these results demonstrate that when the stimulus input is ambiguous more than one organization is neurophysiologically represented despite the task being performed. Task performance did not abolish the alternative representation available in the ambiguous input. This pattern of results supports the hypothesis that multiple representations are maintained in memory, and demonstrate a level of flexibility of the auditory system that would be advantageous, for example, for performing multiple tasks. Having access to multiple neural representations of sounds in the environment would facilitate the behavioral ability to perceive and respond to different events quickly when there are multiple competing

sound sources. These results further indicate that multiple alternative representations are available to perception.

In Experiment 1 (Ambiguous 5 ST condition), the same set of sounds was presented and only the instruction to the participants to switch tasks altered how the sounds were perceived. The physical input was the same across the whole experiment and attention modulated what was perceived within the set of sounds: integration or segregation. Thus, a possible explanation for the current set of results is that both organizations were neurophysiologically represented because the task required switching back and forth between the two organizations across the trials within each stimulus block. That is, trial type was randomized so that the participant had to be ready to perform the loudness task (requiring segregation) or the pattern identification task (requiring integration) as each trial was visually cued. The task-switching may thus have provided the basis for both organizations to remain active throughout the task. The prediction would be that if only one task were being performed with the sounds, under a condition in which only a single task was performed in a block of trials, then it is possible that the one organization involved in the task would dominate in perception, and MMNs would not be elicited by the alternate, non-task organization. However, the results of the spontaneous switching experiments (Pressnitzer and Hupé, 2006; Winkler et al., 2012; Denham et al., 2014) suggest that multiple organizations would be present despite the focus of attention because the switching arises from lower-level competition and not from task demands. Thus, these spontaneous switching studies would predict that multiple alternative organizations should always be present.

The current findings are also consistent with the idea that stimulus-driven processes of grouping do not fully create one organization or the other in perception, but pass “suggestions” about possible groupings to higher processes, with the strength proportional to the weight of evidence favoring each grouping. Top-down processes make selections based on both the weighting induced by the stimulus-driven strength of the organization and on task requirements, biases, or expectations. The neural traces of the possible organizations may persist for some time, even when not being used, making it easier to reinstate a recently used task. In these experiments, trials on the “attend pattern” and “attend loudness” tasks were randomly intermingled, so they occurred in close temporal proximity and possibly served to keep the neural traces ‘active’.

In sound environments when there are multiple possible organizations occurring simultaneously, overlapping acoustics from many sources, maintaining several representations may facilitate the ability to switch attention around the room or area with facility and ease. Listening to an orchestra involves perceiving the global ‘picture’ of the overall harmony (how all the instruments sound together), as well as hearing the individual melodies of various instruments. To be able to switch from listening to the global harmony to the individual melodies, we propose that neural representations of both organizations are maintained to allow access to multiple representations simultaneously. This would allow for rapid and flexible switching. Our results thus support the hypothesis that maintenance of multiple neural representations facilitates behavioral goals in multi-source environments by allowing access to several sound organizations.

Rather than a competitive model of the system, in which the multiple organizations are in competition for a winner-take-all resolution, results of the current study suggest that multiple representations are neurophysiologically represented simultaneously regardless of task performance, even when one organization appears in perception. This type of rapid and flexible adaptive processing is needed in everyday situations to maintain stable auditory representations of the environment and allow rapid scanning of interesting events in busy places.

Acknowledgments

This research was supported by the National Institutes of Health (R01DC004263, E.S.).

The authors thank Jean DeMarco and Julia Wang for their help with data collection. These data were presented in part at the 15th World Congress of Psychophysiology of the International Organization of Psychophysiology (I.O.P.), Budapest, Hungary September 1–4, 2010.

References

- Bregman, AS. Auditory Scene Analysis: The Perceptual Organization of Sound. MIT Press; Cambridge MA: 1990.
- Bregman AS. Auditory streaming is cumulative. *J Exp Psychol: Hum Percept Perform.* 1978; 4(3): 380–387. [PubMed: 681887]
- Brochard R, Drake C, Botte MC, McAdams S. Perceptual organization of complex auditory sequences: effect of number of simultaneous subsequences and frequency separation. *J Exp Psychol: Hum Percept Perform.* 1999; 25:1742–1759. [PubMed: 10641316]
- 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]
- Cusack R, Deeks J, Aikman G, Carlyon RP. Effects of location, frequency region, and time course of selective attention of auditory scene analysis. *J Exp Psychol: Hum Percept Perform.* 2004; 30(4): 643–656. [PubMed: 15301615]
- Denham S, Böhm TM, Bendixen A, Szalardy O, Kocsis Z, Mill R, Winkler I. Stable individual characteristics in the perception of multiple embedded patterns in multistable auditory stimuli. *Front Neurosci.* 2014; 8:25. [PubMed: 24616656]
- Denham SL, Gyimesi K, Stefanics G, Winkler I. Perceptual bi-stability in auditory streaming: how much do stimulus features matter? *Learn Percept.* 2013; 5:73–100.
- Horváth J, Czigler I, Sussman E, Winkler I. Simultaneously active pre-attentive representations of local and global rules for sound sequences in the human brain. *Brain Res Cogn Brain Res.* 2001; 12(1):131–144. [PubMed: 11489616]
- Mesgarani N, Chang EF. Selective cortical representation of attended speaker in multi-talker speech perception. *Nature.* 2012; 485(7397):233–236. [PubMed: 22522927]
- Micheyl C, Tian B, Carlyon RP, Raucher JP. Perceptual organization of tone sequences in the auditory cortex of awake macaques. *Neuron.* 2005; 48:139–148. [PubMed: 16202714]
- Müller D, Widmann A, Schröger E. Auditory streaming affects the processing of successive deviant and standard sounds. *Psychophysiology.* 2005; 42(6):668–678. [PubMed: 16364062]
- Näätänen R, Sussman E, Salisbury D, Shafer V. Mismatch negativity (MMN) as an index of cognitive deficit. *Brain Topogr.* 2014; 27(4):451–466. [PubMed: 24838819]
- Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I. “Primitive intelligence” in the auditory cortex. *Trends Neurosci.* 2001; 24(5):283–288. [PubMed: 11311381]
- Nelken I. Processing of complex stimuli and natural scenes in the auditory cortex. *Curr Opin Neurobiol.* 2004; 14(4):474–480. [PubMed: 15321068]
- Pressnitzer D, Hupé JM. Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. *Curr Biol.* 2006; 16(13):1351–1357. [PubMed: 16824924]

- Rahne T, Böckmann M, von Specht H, Sussman E. Visual cues can modulate integration and segregation of objects in auditory scene analysis. *Brain Res.* 2007; 1144:127–135. [PubMed: 17306232]
- Rahne T, Sussman E. Neural representations of auditory input accommodate to the context in a dynamically changing acoustic environment. *Eur J Neurosci.* 2009; 29(1):205–211. [PubMed: 19087164]
- Shamma SA, Elhilali M, Micheyl C. Temporal coherence and attention in auditory scene analysis. *Trends Neurosci.* 2011; 34(3):114–123. [PubMed: 21196054]
- Sterzer P, Kleinschmidt A, Rees G. The neural bases of multistable perception. *Trends in Neurosci.* 2009; 13(7):310–318.
- Sussman E. Integration and segregation in auditory scene analysis. *J Acoust Soc Am.* 2005; 117(3):1285–1298. [PubMed: 15807017]
- Sussman E. A new view on the MMN and attention debate: auditory context effects. *J Psychophysiol.* 2007; 21(3–4):164–175.
- Sussman E. What gets modeled in complex auditory environments: finding regularity in a changing auditory scene? (Proceedings of the 15th World Congress of Psychophysiology). *Int J Psychophysiol.* 2010; 77(3):215.
- Sussman E. Attention matters: pitch vs. pattern processing in adolescence. *Front Dev Psychol.* 2013; 4(333):1–8.
- 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.* 2005; 5(1):93–110. [PubMed: 15913011]
- Sussman E, Horváth J, Winkler I, Orr M. The role of attention in the formation of auditory streams. *Percept Psychophys.* 2007a; 69(1):136–152. [PubMed: 17515223]
- Sussman E, Ritter W, Vaughan HG Jr. Attention affects the organization of auditory input associated with the mismatch negativity system. *Brain Res.* 1998; 789(1):130–138. [PubMed: 9602095]
- 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, Steinschneider M. Neurophysiological evidence for context-dependent encoding of sensory input in human auditory cortex. *Brain Res.* 2006; 1075(1):165–174. [PubMed: 16460703]
- Sussman E, Steinschneider M, Lee W, Lawson K. Auditory scene analysis in school-aged children with developmental language disorders. *Int J Psycho-physiol.* 2014 in press.
- 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. *Hear Res.* 2007b; 225:117–127. [PubMed: 17300890]
- Sussman-Fort J, Sussman E. The effect of stimulus context on the buildup to stream segregation. *Front Neurosci.* 2014; 8:93. [PubMed: 24808822]
- Szalárdy O, Böhm TM, Bendixen A, Winkler I. Event-related potential correlates of sound organization: early sensory and late cognitive effects. *Biol Psychol.* 2013; 93(1):97–104. [PubMed: 23384511]
- Van Noorden, LPAS. Temporal coherence in the perception of tone sequences. Eindhoven University of Technology; Eindhoven, The Netherlands: 1975. Unpublished doctoral dissertation
- Winkler I, Denham S, Mill R, Böhm TM. Multistability in auditory stream segregation: a predictive coding view. *Philos Trans Royal Soc Lond Ser B Biol Sci.* 2012; 367:1591.
- Winkler I, Denham SL, Nelken I. Modeling the auditory scene: predictive regularity representations and perceptual objects. *Trends Cogn Sci.* 2009; 13(12):532–540. [PubMed: 19828357]

Fig. 1.

Schematic of the stimulus paradigm for Experiment 1. A. Stimulus sequence. Illustration of the three stimulus patterns (Pattern 1, Pattern 2, and Pattern 3) that comprise the stimulus sequence. The y-axis depicts tone frequency in Hz (5 semitones between H and L tones), and the x-axis depicts the timing in ms. The three patterns are presented randomly throughout each trial. B. Loudness Detection Task. Participants segregated out the L tones, focusing on the intensity (69 dB) of the tones and made a button response whenever a random deviant intensity tone (81 dB) was detected. C. Pattern Identification Task. Participants integrated H and L tones to identify the three different patterns. Button presses were made to the designated target pattern. D. Randomized Trials. Visual cues were used to randomly designate prior to each trial which task was to be performed (see text for further details).

Fig. 2.

Event-related potentials elicited by the target (top row) and non-target (bottom row) deviants (thick solid line) and standards (thin solid line) are displayed for Experiment 1 (left column) and Experiment 2 (right column), showing responses for the Loudness detection task (top panels) and the Pattern identification tasks (bottom panels).

Attend pattern
[target]

308 ms

Attend intensity
[non-target]

308 ms

Fig. 3.

Experiment 1 Ambiguous condition (5 ST). Difference (deviant-minus-standard) waveforms are displayed for the MMN (top panel) and P3b (bottom panel) components at the electrode of greatest signal-to-noise ratio. MMN and P3b elicited by the Intensity deviants for the attended target (thick, solid line) and the unattended non-target (thin, solid line) are shown in the left column, and for the Pattern deviants in the right column. Gray bars indicate the peak of the components used to determine measuring interval for obtaining mean amplitudes for statistical analyses. Displayed above the waveforms are maps of the scalp voltage distribution obtained at the peak of the MMN and P3b components.

Fig. 4. Experiment 2 Unambiguous condition (15 ST and 1 ST). Difference (deviant-minus-standard) waveforms are displayed for the MMN (top panel) and P3b (bottom panel) components at the electrode of greatest signal-to-noise ratio. MMN and P3b elicited by the Intensity deviants are shown in the left column for the attended target and the unattended non-targets for the 15 ST (thick line) and 1 ST (thin line) conditions. Pattern deviant responses are shown in the right column for the attended target and the unattended non-targets for the 1 ST (thick line) and 15 ST (thin line) conditions. Gray bars indicate the peak of the components used to determine measuring interval for obtaining mean amplitudes for statistical analyses. Displayed above the waveforms are maps of the scalp voltage distribution obtained at the peak of the MMN and P3b components.

Table 1

Behavioral results.

Experiment 1				
Trial Type/Task	HR	FAR	d'	RT
5 ST trials				
Loudness	0.68 (0.25)	0.11 (0.17)	2.36 (1.74)	0.414 (0.082)
Pattern 1	0.74 (0.17)	0.08 (0.11)	2.32 (0.99)	0.447 (0.070)
Pattern 2	0.62 (0.17)	0.17 (0.12)	1.40 (0.89)	0.425 (0.065)
Pattern 3	0.81 (0.11)	0.03 (0.04)	3.07 (0.89)	0.269 (0.052)
Experiment 2				
1 ST trials				
Pattern 1	0.81 (0.14)	0.08 (0.10)	2.70 (1.07)	0.484 (0.065)
Pattern 2	0.71 (0.17)	0.09 (0.08)	2.13 (0.94)	0.439 (0.034)
Pattern 3	0.80 (0.15)	0.02 (0.02)	3.18 (0.75)	0.297 (0.055)
15 ST trials				
Loudness	0.88 (0.05)	0.001 (0.001)	4.16 (0.25)	0.429 (0.046)

HR=hit rate; FAR=false alarm rate; RT=reaction time

Table 2

Experiment 1. Ambiguous 5 ST condition.

MMN						
Electrode	Attend loudness			Unattend pattern		
	Dev	STD	Difference	DEV	STD	Difference
Fz	-2.40 (1.32)	-0.94 (1.05)	-1.46	-1.73 (1.59)	-0.29 (1.34)	-1.44
Cz	-1.93 (2.46)	-0.24 (0.94)	-1.69	-0.16 (1.69)	0.78 (0.91)	-0.93
Pz	-1.00 (3.39)	0.14 (0.86)	-1.13	0.93 (1.97)	1.29 (0.76)	-0.36
Unattend loudness						
Attend pattern						
Fz	-0.67 (1.46)	1.45 (1.21)	-2.12	-0.65 (1.00)	0.97 (0.74)	-1.62
Cz	0.31 (1.31)	1.07 (1.34)	-0.76	-0.13 (1.55)	0.94 (0.99)	-1.08
Pz	0.87 (1.45)	0.06 (1.24)	0.81	0.19 (1.64)	0.26 (0.93)	-0.07

Table 3

Experiment 1. Ambiguous 5 ST condition.

P3b						
Electrode	Attend loudness			Unattend pattern		
	Dev	STD	Difference	DEV	STD	Difference
Pz	5.38 (4.37)	0.021 (1.50)	5.36	1.24 (1.57)	0.96 (1.08)	0.28
Cz	2.93 (3.35)	-0.30 (1.29)	3.22	0.64 (1.82)	0.69 (1.14)	-0.04
Fz	0.23 (2.52)	-0.65 (0.93)	0.88	-0.44 (1.56)	-0.04 (0.78)	-0.40
Unattend loudness						
Pz	4.87 (3.43)	-0.73 (1.85)	5.60	1.16 (1.74)	-0.41 (0.83)	1.57
Cz	3.29 (3.00)	-1.28 (1.88)	4.57	0.20 (1.92)	-0.89 (0.983)	1.08
Fz	-0.50 (1.90)	-0.51 (1.73)	0.01	-1.30 (1.25)	-0.60 (0.99)	-0.70

Table 4

Experiment 2. Unambiguous conditions.

MMN						
15 ST trials						
Electrode	Attend loudness			Unattend pattern		
	Dev	STD	Difference	DEV	STD	Difference
Fz	-1.89 (1.57)	0.00 (0.98)	-1.89	-0.09 (0.89)	0.42 (0.62)	-0.50
Cz	-2.00 (1.86)	0.49 (0.98)	-2.49	0.14 (0.91)	0.26 (0.39)	-0.11
Pz	-1.72 (1.53)	0.58 (0.87)	-2.29	0.20 (1.12)	-0.28 (0.29)	0.48
1 ST trials						
Electrode	Attend pattern			Unattend loudness		
	Dev	STD	Difference	DEV	STD	Difference
Fz	-0.98 (1.50)	1.12 (1.15)	-2.10	-1.77 (1.23)	-1.38 (1.14)	-0.39
Cz	-0.30 (1.57)	1.19 (1.31)	-1.49	-0.84 (0.84)	-0.49 (0.94)	-0.35
Pz	0.05 (1.24)	0.53 (1.29)	-0.49	0.15 (0.64)	0.27 (0.58)	-0.12

Table 5

Experiment 2. Unambiguous conditions.

P3b						
15 ST trials						
Electrode	Attend loudness			Unattend pattern		
	Dev	STD	Difference	DEV	STD	Difference
Pz	7.04 (3.19)	0.39 (0.99)	6.68	0.28 (1.14)	-0.68 (0.63)	0.96
Cz	5.57 (3.57)	0.23 (0.97)	5.34	0.61 (0.85)	-0.97 (1.00)	1.58
Fz	2.85 (3.13)	0.12 (0.90)	2.73	0.38 (0.69)	-0.73 (1.15)	1.12
1 ST trials						
Electrode	Attend pattern			Unattend loudness		
	Dev	STD	Difference	DEV	STD	Difference
Pz	2.19 (2.63)	-0.89 (1.11)	3.08	0.55 (1.34)	0.67 (0.39)	-0.12
Cz	1.29 (2.48)	-1.28 (1.30)	2.57	0.267 (1.35)	0.61 (0.72)	-0.34
Fz	-1.01 (1.76)	-1.04 (1.31)	0.03	-0.30 (1.33)	0.042 (0.52)	-0.34