Published in final edited form as:

Neuropsychologia. 2009 February ; 47(3): 771–785. doi:10.1016/j.neuropsychologia.2008.12.007.

Attention effects on auditory scene analysis in children

Elyse Sussman¹ and Mitchell Steinschneider²

¹ Department of Neuroscience, Albert Einstein College of Medicine, Bronx, NY

² Department of Neurology, Albert Einstein College of Medicine, Bronx, NY

Abstract

Auditory scene analysis begins in infancy, making it possible for the baby to distinguish its mother's voice from other noises in the environment. Despite the importance of this process for human behavior, the question of how perceptual sound organization develops during childhood is not well understood. The current study investigated the role of attention for perceiving sound streams in a group of school-aged children and young adults. We behaviorally determined the frequency separation at which a set of sounds was detected as one integrated or two separated streams and compared these measures with passively and actively obtained electrophysiological indices (MMN and P3b) of the same sounds. In adults, there was a high degree of concordance between passive and active electrophysiological indices of stream segregation that matched with perception. In contrast, there was a large disparity in children. Active electrophysiological indices of streaming were concordant with behavioral measures of perception, whereas passive indices were not. In addition, children required larger frequency separations to perceive two streams compared to adults. Our results suggest that differences in stream segregation between children and adults reflect an under-development of basic auditory processing mechanisms, and indicate a developmental role of attention for shaping physiological responses that optimize processes engaged during passive audition.

Keywords

Attention; auditory stream segregation; event-related potentials (ERPs); mismatch negativity (MMN); P3b

Introduction

Attention is intimately tied to learning. Practice of a new skill leads to automaticity of that skill (Schneider & Shiffrin, 1977). Developing automaticity in learned skills is beneficial to enhance performance and manage the intricacy of common complex tasks. The link between attention and learning has been well documented. It has long been known that attention modulates auditory cortical physiology (Hubel, Henson, Rupert, & Galambos, 1959). In addition, it has been shown that learning reshapes tuning characteristics by modulating receptive fields (RFs) (Weinberger, 2003), altering spike firing rate (Pandya et al., 2005), and changing temporal response properties of cells (Zhou & Merzenich, 2007) in task-specific ways (Fritz, Elhilali, & Shamma, 2005; Fritz, Shamma, Elhilali, & Klein, 2003; Weinberger, 2003; Zhang, Bao, &

Corresponding author: Elyse Sussman, Ph.D., Department of Neuroscience, Albert Einstein College of Medicine, 1410 Pelham Parkway S, Bronx, NY 10461, USA, Ph: 718-430-3313, Fax: 718-430-8821, Email: esussman@acom.yu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Merzenich, 2001). Learning-induced reshaping of RFs in auditory cortex has also been shown to correlate with the behavioral importance of a stimulus (Rutkowski & Weinberger, 2005). In humans, perceptual learning also brings about physiological changes (Stevens, Fanning, Coch, Sanders, & Neville, 2008; Tallal et al., 1996; Tremblay & Kraus, 2002). Children trained to distinguish two within-category speech sounds increased their behavioral discrimination on post-training trials, which was accompanied by increased amplitude of auditory cortical evoked potentials (Tremblay & Kraus, 2002). Thus, there is considerable evidence in animals and humans highlighting the significance of attention for altering cortical physiology to enhance perception of the sounds in our environment.

The sound environment is dynamically changing, with overlapping acoustic information originating from simultaneously active sources. Despite the complexity of natural environments, we readily perceive multiple distinct sound events. Fundamental neural mechanisms facilitate this process from infancy. No one needs to teach a baby how to segregate the sound of a rattle from their mother's voice. However, the acuity with which we process the sound environment in adulthood is greater than in early childhood. Studies in infants indicate that maturation of the auditory system may shape the stream segregation process. Infants require much larger separations between acoustic dimensions (Demany, 1982; McAdams & Bertoncini, 1997) and slower paced stimuli (Demany, 1982; McAdams & Bertoncini, 1997) to perceive separated sound sources than do adults. These findings related to auditory stream segregation are consistent with literature showing that infants have higher auditory thresholds than adults for detecting differences in pitch (Nozza, & Wilson, 1984; Ruben, 1992; Sinnott, Pisoni, & Aslin, 1983; Trehub, Schneider, & Endmand, 1980; Werner, Marean, Halpin, Spetner, & Gillenwater, 1992), spatial location (Clifton, Morrongiello, Kulig, & Dowd, 1981; Muir & Field, 1979) and temporal features (Morrongiello, & Trehub 1987; Werner et al., 1992; Wightman, Allen, Dolan, Kistler, & Jamieson, 1989). These maturational differences in the perception of acoustic parameters needed for accurate auditory scene perception suggest that changes occurring during the development of the auditory system contribute to the enhanced acuity of the segregation process observed in adults. However, the acuity of scene analysis seen in adulthood takes a longer developmental course than maturation of discrimination processes for the specific acoustic features required for stream segregation (Sussman, Wong, Horváth, Winkler, & Wang, 2007).

How does the finer acuity experienced in adulthood develop? Although studies suggest that the basic neural mechanisms underlying stream segregation are present from birth, factors that contribute to the accurate development of auditory scene analysis in childhood are not well understood. In particular, it is not clear the degree to which attention may be needed during development to refine the cortical physiology associated with sound organization. Our recent research investigating the role of stimulus-driven (automatic) and goal-directed (attention-based) processes in scene analysis demonstrated that automatic mechanisms group low-level properties extracted by the auditory system to build separate, physiologically-defined descriptions of the distinctive sound streams in adults (Sussman, 2005). Importantly, attention interacts with these low-level processes, regulating and maintaining neural representations of the distinct sound objects, according to task-specific goals (Sussman, Ritter, & Vaughan, 1998; Sussman, Winkler, Houtilainen, Ritter, & Näätänen, 2002). In the present study, behavioral and electrophysiological measures were used to compare active and passive auditory stream segregation processes in children and adults. Effects of attention on physiological markers denoting stream segregation were determined using sound streams that had varying frequency differences between two sets of sounds. Our hypothesis was that attention modulates cortical physiology in children by influencing the acuity of automatic auditory processes in auditory scene analysis.

To test this hypothesis, we determined whether actively segregating sounds would modulate neurophysiological indices of sound organization. We compared perception of sound streams with the evoked brain responses obtained under passive listening conditions and under active attentional control. The purpose was to determine whether electrophysiological indices of stream segregation in both passive and active listening conditions would parallel the perception of streaming in 9–12 year-old children.

The mismatch negativity (MMN) and P3b event-related brain potentials (ERPs) were the primary measures used to evaluate unattended (MMN) and attended (P3b) physiological sound processing. The MMN reflects a passive deviance detection process that is based on the auditory system's ability to detect and maintain regularities in the acoustic signal. The most common way to elicit MMN is to present a randomly occurring infrequent (oddball) stimulus that deviates from a frequently repeating (standard) stimulus. Simple deviations in stimulus features, such as frequency or intensity, elicit the MMN (Näätänen, 1992). The main neural generators of the MMN are located bilaterally in the supratemporal plane, indicating their auditory cortical origin (Halgren et al., 1995; Sams & Hari, 1991; Scherg, Vajsar, & Picton, 1989). MMN reflects the outcome of a change detection process that is based upon the memory of sound regularities in ongoing auditory input (Näätänen, 1992; Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001).

Because MMN is elicited by the violation of detected auditory regularities, it can be used to determine which regularities (e.g., pattern of sounds) are represented in memory at the time the “deviant” occurs. The deviant stimuli are used as probes to determine whether a set of sounds is segregated or integrated (Sussman et al., 1998, Sussman, Ritter, & Vaughan, 1999; Sussman, 2005). MMN is thus an effective tool for the current study because it can index sound organization processes in both passive and active listening conditions (Sussman, 2007). MMN under passive conditions shows good correspondence with the perception of the same sounds in the performance of active tasks, including, but not limited to, the ability to index large vs. small changes in sound frequency (Tiitinen, May, Reinikainen, & Näätänen, 1994), to index the distinction between one and two auditory streams (Sussman, Cepenione, Shestakova, Näätänen, & Winkler, 2001; Sussman et al., 1999), and to index learning effects (Atienza, Cantero, & Dominquez-Marin, 2002; Näätänen, Schröger, Karakas, Tervaniemi, & Paavilainen, 1993).

The P3b component (Sutton, Braren, Zubin, & John, 1965), a positive-going waveform that is sometimes called the “target P3” to distinguish it from the “novelty” P3a, also indicates deviance detection. In contrast to MMN, the P3b reflects attentive processing, as it is elicited only when attention is focused on the sounds and a task is being performed. P3b is non-modality specific, with widespread neural generators that results in a largely posterior scalp distribution in adults (Friedman, Cycowicz, & Gaeta, 2001; Halgren, Marinkovic, & Chauvel, 1998). Another of the task-dependent processing components is the N2b component, which in adults is generally elicited prior to the P3b component by deviant sounds even without an overt response, provided that attention is directed toward the sounds (Näätänen, Simpson, & Loveless, 1982; Novak, Ritter, Vaughan, & Witznitzer, 1990). The N2b is a negative-going waveform that often overlaps with the MMN. However, it can be distinguished from the MMN by its scalp topography, being maximum in centro-parietal electrode sites with no polarity reversal at the mastoids, whereas the MMN has a more frontal maximum and generally inverts in polarity at electrode sites located over the mastoids (Novak et al., 1990). Little is known about the N2b in children, whereas the positive-going P3b waveform is a robust response and is easily distinguished from the negative-going deviance detection processing components. Therefore, for the purposes of the current study, the MMN and P3b, will provide physiological evidence of passive and active indices to compare with behavioral indices of stream segregation.

If automatic and attentive stream segregation processes have different thresholds, then there will be a pronounced disparity in the MMN results obtained in the IGNORE and ATTEND conditions. The degree of concordance will be determined by how closely the elicitation of MMN matches with the perception of the sounds as two streams. Concordance of neurophysiological responses obtained in active conditions with perception will be demonstrated if MMN, and P3b, are elicited in the same conditions in which two streams are reportedly heard. Finally, if attention plays a role in shaping automatic auditory processes during development, then an additional discrepancy between children and adults should be observed with respect to the agreement between passive elicitation of MMN and perception of streaming.

MATERIALS and METHODS

Participants

Forty-six children and ten adults participated in the study. Participants were recruited by flyers posted in the immediate medical/research community and in local area schools. Adults gave informed consent, children gave written assent and their accompanying parent gave written consent after the protocol was explained to them. The protocol was approved by the Committee for Clinical Investigations at the Albert Einstein College of Medicine and all procedures were carried out according to the Declaration of Helsinki. All participants were paid for their participation. Table 1 shows the assignment of the participants by experiment, their age and gender distribution¹. Recruits were pre-screened by phone interview (the parent for children) to exclude past or present diagnoses of learning, speech/language, hearing, behavioral, or neurological disorders, or report of special educational services or school grade retention. Recruits who met the pre-screen criteria underwent a two-hour psychometric testing session with a licensed psychologist to meet criteria for the behavioral and ERP recording sessions². All participants passed a hearing screen (hearing threshold of 20 dB HL or better for 250, 1000 Hz; hearing threshold of 20 dB HL or better for 250, 2000, and 4000 Hz).

Stimuli

Pure tones, 50 ms in duration (7.5 ms rise/fall times), were created using NeuroStimTM software and presented binaurally via E-a-rtone[®] 3A insert earphones. Two types of tones were presented in each sound sequence. One tone was presented at a fixed frequency in all the sequences and the other tone varied in its spectral distance from the fixed tone in different conditions of frequency separation. The two tones were presented every 100 ms (onset-to-onset pace) (Figure 1B), where the 'X' represents a tone with the fixed frequency and the 'O' represents a tone that was either near or far in frequency from the fixed 'X' tone. An auditory oddball condition was also conducted, in which the 'X' tone was presented without the intervening 'O' tones (Figure 1A) at a rate of once every 300 ms (the pace the 'X' tone occurs in the experimental conditions). The frequency separations for the conditions were chosen according to the logarithmic musical scale, in semitone (ST) steps. This method was used so that the intervallic relationship between any two tones could be maintained across subject groups, even if the absolute frequency values used for the fixed 'X' tone differed between adult and child groups, as described below.

¹Some children participated in more than one experiment.

²Standard scores on all screening tests for cognitive and language function, reading/decoding skills, and phonological processing abilities had to be at least within the average range (standard scores of at least 85) on the following instruments: Wechsler Abbreviated Scale of Intelligence (WASI) for cognitive function; Woodcock-Johnson-III Tests of Achievement (WJ-III) Letter/Word Identification and Word Attack for reading/decoding; WJ-III Understanding Directions and Children's Essentials of Language Fundamentals-3 (CELF-3); The Phonological Awareness Test (PAT) Rhyming; and The Comprehensive Test of Phonological Processing (CTOPP) core screening tests for phonological awareness, phonological memory, and rapid naming. In addition, participants or their parents had to report fewer than 6 symptoms of inattention and hyperactivity/impulsivity on a DSM-IV checklist.

For children, the standard 'X' tone intensity was 65 dB SPL, occurring with 88% probability. The deviant was an 'X' tone with an intensity increment (80 dB SPL) that occurred 12% of the time. The intervening 'O' tones had any of four different intensity values (60, 70, 75, and 85 dB), distributed equiprobably, that spanned above and below the standard and deviant intensity values of the 'X' tones. A smaller, 12 dB, difference was used with adults, and the intervening intensity values were in four-dB steps. These intensity values for children and adults were chosen to ensure that the intensity contrasts (between standard and deviant 'X' tones) would be well above threshold for all participants. The intensity variability of the 'O' tones ensured that the 'X' tones had neither the highest nor the lowest intensity value of the six different intensity values presented, and minimized the possibility that the standard or deviant 'X' tone could be easily detected regardless of whether or not the 'X' and 'O' tones segregated from one another on the basis of frequency separation. Tones were calibrated using a sound-level meter with an artificial ear (Brüel & Kjær 2210).

Four experiments comprised the study: **Experiment 1:** Children IGNORE. **Experiment 2:** Children ATTEND. **Experiment 3:** Adults IGNORE. **Experiment 4:** Adults ATTEND. Adults participated in both IGNORE and ATTEND conditions on the same day.

Experiment 1: Children IGNORE—There were 11 experimental conditions, which included ten conditions of frequency separation measured by the distance in semitones (ST) above the 'X' tone (3, 5, 7, 9, 11, 15, 19, 23, 27, and 31 ST) and one Oddball condition. The IGNORE conditions were conducted in three separate experiments, with each participant receiving a total of either four or five conditions (three or four ST conditions along with one oddball condition, see Table 1).

The stimulus frequency of the 'X' tone was 880 Hz for the 3–19 ST conditions. The frequency of the 'O' tone was 3 ST: 1046.5 Hz; 5 ST: 1174.7 Hz; 7 ST: 1318.5 Hz; 9 ST: 1480 Hz; 11 ST: 1661.2 Hz; 15 ST: 2093 Hz; 19 ST: 2637 Hz. For the 23–31 ST conditions, the fixed frequency of the 'X' tone was lowered to 440 and the frequency of the 'O' tone was then 23 ST: 1661.2 Hz; 27 ST: 2093 Hz; 31 ST: 2637 Hz. A total of 3000 'X' tones were presented in each experimental condition and 600 'X' tones in the control conditions, to obtain a standard comparison for delineating MMN (see below).

Experiment 2: Children ATTEND—Deviance Detection Task: There were five experimental conditions, which comprised four ST conditions (3, 7, 11, and 19 ST), and one Oddball. The stimulus frequency of the fixed 'X' tone was 440 Hz and the frequency of the 'O' tone was 3 ST: 523.25 Hz; 7 ST: 659.26 Hz; 11 ST: 830.61 Hz; and 19 ST: 1318.5 Hz. Practice sequences used a 1 ST difference (466.16 Hz) for "one-stream" and 23 ST (1661.2 Hz) for "two-streams". A total of 2400 'X' tones were presented in each experimental condition and 600 'X' tones in the control conditions (see below).

One vs. Two Streams: A separate behavioral experiment was conducted in which the same ST conditions were presented to participants but in short, 25 s sequences. Five sequences were played for each ST condition in a randomized order.

Experiments 3–4: Adults IGNORE and ATTEND—Deviance Detection Task: There were five experimental conditions, which comprised four ST conditions (1, 5, 7, and 11 ST), and one Oddball. The stimulus frequency of the fixed 'X' tone was 1046.5 Hz and the frequency of the 'O' tone was 1 ST: 1108 Hz; 5 ST: 1396 Hz; 7 ST: 1568 Hz; and 11 ST: 1975 Hz. Practice sequences were 1 ST (1108 Hz) for "one-stream" and 15 ST (2489 Hz) for "two-streams". A 12 dB difference was used for the standard (71 dB) and deviant (83 dB) 'X' tone intensity. The intensity value of the 'O' tones was equiprobably distributed among four intensities (67, 75,

79, 87 dB). A total of 1500 'X' tones were presented in each experimental condition and 500 'X' tones in the control conditions (see below).

One vs. Two Streams: A separate behavioral experiment was conducted in which the same ST conditions were presented to participants but in 20 s sequences. Five sequences were played for each ST condition in a randomized order.

Procedures

Participants sat in a reclining chair in an electrically-shielded and sound-attenuated booth. The presentation order of conditions was counterbalanced across participants. In the IGNORE conditions, participants watched a silent captioned video of their choice. The experimenter monitored the EEG for regular eye saccades to ensure that participants were watching the movies and reading the captions.

The ATTEND conditions began with a behavioral experiment to obtain a measure of the perception of the sequences as one or two streams (*One vs. Two Streams*). Participants were given a brief explanatory session, followed by practice and testing sessions. In the explanatory session, sounds were presented to allow participants to distinguish the difference between one versus two streams, using a sequence that was clearly heard as one stream (1 ST higher than the fixed tone) and another that was clearly heard as two streams (15 ST higher than the fixed tone for adults and 23 ST for children). Participants listened to the sounds and were then informed whether it was a one- or two-stream sequence. Visual depictions were displayed simultaneously (Figure 2A and 2B). Next, these two sequences were randomly presented (6 trials) and subjects were asked whether they heard one-stream or two-streams. The experimenter gave feedback after each trial. After this training session, a short practice was given prior to the testing session. Both adults and children were told that they would hear sounds in short sequences as they just did but that the sounds would not be exactly the same. They were to listen to the whole sequence and then report whether the sounds they heard were *more like one* or *more like two* streams of sound. The experimenter recorded their responses before starting the next sequence. No feedback was given during the testing session. There were five trials per condition. The order was counterbalanced across participants.

In the ATTEND conditions in which ERPs were recorded (*Deviance Detection Task*), the session began with task instruction and practice. Visual depiction was provided with instruction. The task was to press the response keypad whenever a louder (or softer in the control conditions) 'X' tone was detected (Figure 2C). The Oddball sequence was presented first for practice with the task. Then a 15 ST condition for adults and a 23 ST condition for children (neither part of the experimental conditions) were used for practice. Participants were instructed to listen to the 'X' tone stream and press the key whenever a louder intensity 'X' tone (or softer in the control condition) occurred within it. For the children, the image of 'knocking on a door' was given and they were told to listen for the knocking on the door sounds and press the response key whenever a louder/softer one was heard. Participants were told that for some sequences it would be easier to find the knocking on the door than in others. The separate runs of the ST conditions were presented in two different randomized orders, with half the participants receiving one randomization and half receiving the other, for each experiment. The purpose for randomizing the run order (opposed to the ST condition order) was to avoid possible carryover effects of the integrated/segregated organizations. For the Adults, in which both IGNORE and ATTEND conditions were conducted in the same session, the IGNORE conditions were presented first to all participants to avoid potential influences from attending to the tones. The run order was randomized, separately within each experiment, and two different run orders were counterbalanced across participants.

Electrode placement and EEG recording—The EEG recordings were obtained using a 32-channel electrode cap that incorporated a subset of the International 10–20 system (Jasper, 1958). Additional electrodes were placed over the left and right mastoids (LM and RM, respectively). The tip of the nose was used as the reference electrode during recordings. F7 and F8 electrode sites were used in a bipolar configuration to monitor the horizontal electro-oculogram (EOG). FP1 and an electrode placed below the left eye were used in a bipolar configuration to monitor the vertical EOG. All impedances were maintained below 5 k Ω for adults and 10 k Ω for children. The EEG and EOG were digitized (Neuroscan Synamps amplifier, Compumedics Corp., El Paso, Texas) at a sampling rate of 1000 Hz (0.05–200 Hz bandpass). EEG was filtered off-line with a bandpass of 1–15 Hz. Artifact rejection was set to exclude activity exceeding 75 μ V for adults and 100 μ V for children after EEG epochs were baseline corrected. Epochs were 600 ms in duration, starting 100 ms pre-stimulus onset and ending 500 ms post-stimulus onset. Epochs were extended to 700 ms (100 ms pre-stimulus and 600 ms post-stimulus) for the Adult ATTEND conditions due to the later peak latency of the P3b (see Results for details).

Data analysis

Behavioral Responses: For each participant, reaction time (RT), hit rate (HR), false alarm rate (FAR), and d' (Green & Swets, 1966) were calculated in each condition separately. A response was considered correct when it was recorded 100–1200 ms after stimulus onset. We calculated d' as a measure of the subject's ability to detect the louder intensity tones within the 'X' tone stream, which presupposed that the 'X' tones were heard in a separate stream from the 'O' tones. 'Hits' were correct button presses to the infrequent intensity 'X' tones; 'misses' were the absence of a button press when the deviant intensity 'X' tones occurred. "Correct rejections" were correct "no-go" responses to any of the 'X' tones, and "false alarms" were any button presses that were not to the deviant intensity 'X' tones. There was a very low FAR (<1%) for all age groups in all conditions. Button presses were only required on a small percentage of the tones. As there were many instances in which subjects had perfect performance, which can result in statistically infinite d' , we adjusted the scores slightly to avoid infinite d' by adding 0.5 to all the cells (hits, misses, false alarms, and correct rejections) and dividing the resulting scores by the number of trials ($n+1$) related to the proportion (Macmillan & Creelman, 2005).

Repeated measures ANOVA were calculated to compare the perception of one vs. two streams across ST conditions (*One vs. two streams* behavioral experiment) and to compare sensitivity for intensity deviance detection, using the d' scores, across ST conditions (*Deviance detection task* obtained during ERP recording).

Electrophysiological Responses: In the IGNORE experiments, ERPs were averaged separately after artifact rejection for the infrequent 'X' tone in each ST condition and for the frequent 'X' tone elicited in each of the corresponding control blocks. In the ATTEND experiments, the ERPs elicited by the infrequent 'X' tones were additionally averaged separately by correct and incorrect responses in each condition (Table 2). These are the responses to the louder intensity deviants from the experimental blocks. Additionally, the ERPs elicited by the softer 'X' tones in the control conditions, and the standards from the main blocks were averaged (correct and incorrect responses together) and are additionally presented in figures.

Control for delineating MMN: A corresponding control condition was conducted for each experimental condition in both IGNORE and ATTEND experiments. To limit stimulus-specific ERP effects, the intensity of the standard and deviant 'X' tones was reversed (e.g., standard intensity was 80 dB SPL and deviant intensity was 65 dB SPL). All other parameters

were identical to the corresponding experimental conditions. The purpose was to obtain a comparison tone that was physically identical to the “deviant” (Jacobsen et al., 2003). Thus, the louder standard tones from the control blocks were used as control tones (i.e. standards) for the deviants (i.e. the louder tones) in the experimental conditions. That is, the intensity value of the control standards matched the intensity value of the deviants. In the ATTEND conditions, participants practiced and detected softer intensity sounds in their corresponding conditions.

The MMN, N2b and P3b amplitudes were measured using a 40 ms window centered on their respective peak latencies obtained in the deviant-minus-control difference curves (at Fz for MMN, Cz for N2b, and Pz for P3b) of the grand-mean averages, for each condition and group separately. However, because the N2b overlaps with the MMN when the tones are attended, the mastoid peak was used in the ATTEND experiment conditions in adults to distinguish the peak latency of the MMN from the peak latency of the N2b. Where there was no observable peak, the mean interval was used from the nearest condition in which the component was present. The peak latencies used to measure each component are displayed in Table 3. Latency measurements were obtained using a peak detection program (Neuroscan 4.3).

To further determine presence of response components, one-sample *t* tests were used to verify whether the mean amplitude was significantly different from zero. Repeated measures ANOVAs were used to compare amplitudes of the MMN and P3b components across ST conditions, separately in adults and children in a two-way model with factors of response accuracy (correct vs. incorrect) and ST condition (including the Oddball). Hundt-Feldt corrections were used. Post hoc tests were calculated using Tukey HSD.

RESULTS

Behavior results

The behavioral results of Adults and Children are summarized in Fig. 3 and Table 4.

ADULTS

One vs. two streams

adults heard two streams more than one stream in the 5, 7, and 11 ST conditions and they heard one stream significantly more in the 1 ST condition than the other conditions ($F_{3,27}=51$, $p<0.00001$, $\epsilon=.36$).

Deviance detection

Sensitivity (d') for detecting the louder sounds in the attended stream was revealed in a main effect of condition ($F_{4,36}=63$, $p<0.001$, $\epsilon=.37$) showing that deviants were detected with greater accuracy in the Oddball, 11, and 7 ST conditions than the 1 and 5 ST conditions. There was no statistical difference in detecting the deviants among the Oddball, 11, and 7 ST conditions. The mean difference between 5 ST and 1 ST, in which deviants weren't accurately detected, was large but was not significantly different ($p=0.27$). This was probably due to the high degree of variance in the 5 ST condition (Table 4), in which the tones were often perceived as one stream. There was a negative relationship of reaction time (RT) with Δf ($F_{3,27}=65.99$, $p<0.0001$, $\epsilon=.72$). RT significantly increased as frequency separation between the tones decreased.

Additionally, RT was significantly faster for the Oddball deviants compared to that of all ST conditions.

CHILDREN

One vs. Two streams

Two streams were heard in the 19 and 11 ST conditions and significantly more often than in the 7 and 3 ST conditions ($F_{3,33}=65, p<0.00001, \epsilon=.50$). Two streams were heard more in the 7 than 3 ST conditions. No significant difference was observed between the 19 and 11 ST conditions.

Deviance Detection

A main effect of ST difference on accuracy for children ($F_{4,44}=126, p<0.00001, \epsilon=.66$) showed no difference in sensitivity for detecting loudness deviants in the attended stream in the Oddball and 19 ST conditions. Oddball was better than 11, 7, and 3 ST. There was no difference between 19 and 11, or between 11 and 7. Detection was better in the 7 than 3 ST condition. RT to deviants was significantly faster in the Oddball condition compared to all ST conditions ($F_{3,33}=14.85, p<0.0001, \epsilon=.58$). Among the ST conditions, only 19 ST was significantly faster than 7 ST.

GROUP COMPARISON

Whereas children required larger frequency separations to perceive two versus one sound stream, the shapes of performance curves were very similar to those seen in adults (Figure 3). Both groups had clearly integrated (1 ST for adults, 3 ST for children) and clearly segregated (11 ST for adults, 19 ST for children) organizations, with an ambiguous region of 5 ST for adults and 7 ST for children. These behavioral results indicate that children require slightly larger frequency separations to hear segregated streams but otherwise have similar behavioral functions for stream segregation. Further, RT increased as ST differences decreased in both children and adults.

Comparing common conditions across groups (Oddball, 11 ST, 7 ST) showed that sensitivity (d') for detecting the louder sounds in the attended stream was greater in the Oddball condition (main effect of ST difference, $F_{2,40}=63, p<0.00001, \epsilon=.98$). There was no significant difference in detecting the louder sounds between the 11 and 7 ST conditions. There was a trend towards children being less sensitive in detecting the louder sounds than adults ($F_{1,20}=3.95, p=0.061$). There was no interaction.

Electrophysiology results

Ignore Conditions

ADULTS

Obligatory responses

Fig. 4 displays the ERPs elicited by deviant and control stimuli in the Ignore conditions (right column). Only the P1 component is prominent in the ST conditions, occurring at the stimulation rate (Sussman, Steinschneider, Gumenyuk, Grushko, & Lawson, 2008).

MMN

MMNs were elicited by deviants (louder intensity 'X' tones) in the Oddball, 11, and 7 ST conditions (Fig. 5). ANOVA comparing the MMN amplitude revealed a main effect of condition ($F_{4, 36}=19.27, p<0.001$), in which the Oddball amplitude was larger than all the ST conditions. There was no difference between 11 and 7 ST (in which the amplitude was significantly less than zero voltage) or between the amplitude of the waveforms measured in the MMN range for the 5 and 1 ST conditions (in which the amplitude was not significantly

less than zero voltage). The voltage maps showing fronto-central negative scalp distribution in the Oddball, 11, and 7 ST conditions (Fig. 5, top row) are consistent with MMN generators in bilateral auditory cortices (Alho, 1995). These results obtained in the IGNORE conditions compare favorably with behavioral data, and indicate that the MMN is a reliable measure of stream segregation at all but ambiguous ST conditions.

CHILDREN

Obligatory responses

Fig. 6 displays the ERPs elicited by deviant and control stimuli elicited in the Ignore conditions (right column). The P1 and N2 components, occurring at the stimulation rate, are prominent in the Oddball condition (Sussman et al., 2008). The P1/N2 complex is also prominent in the 31 ST condition. At 19 ST and smaller, only the P1 is prominent, and is most evident to the 'X' tones occurring every 300 ms.

MMN

MMNs were elicited by deviants (louder intensity 'X' tones) only in the Oddball and 31 ST conditions when the sounds were ignored (Fig. 7). A statistically non-significant MMN was observed in the 27 ST condition. There was no significant difference between the MMN amplitude in the Oddball and 31 ST conditions ($t_{11} < 1$, $p > 0.45$). These results indicate that when the input was ignored, stream segregation, as measured electrophysiologically, occurred only at the largest tested frequency difference between sounds in children. Importantly, this result compares unfavorably with behavioral measures, and is in stark contrast to the findings in adults. As in adults, the voltage maps show a fronto-central negative scalp distribution of the MMN (Fig. 7, top row).

Attend Conditions

The number of correct and incorrect ERP responses remaining after artifact rejection is summarized by condition for Adults and Children in Table 2.

ADULTS

ERPs

Fig. 4 (left column) displays the ERPs elicited by the deviant and control stimuli in all attend conditions. The P1 component is evident in all ST conditions. The obligatory components at Fz evoked by the deviant stimuli overlap with a processing-contingent negativity that likely reflects both MMN and N2b (except for the 1 ST condition in which target detection components were not elicited). The P3b (dashed line) is seen in the deviant ERPs at Pz in all but the 1 ST condition.

MMN component

Deviants that were correctly detected elicited MMN in the Oddball, 11, 7, and 5 ST conditions but not in the 1 ST condition (Fig. 8). The mastoid trace is displayed in Fig. 8 to show the polarity inversion, and indicate the peak latency of the MMN, which is overlapped with the slightly later N2b at the frontal electrode sites. MMNs were also elicited in the Oddball, 11, and 7 conditions for the incorrect responses but not in the 5 ST or 1 ST conditions (Fig. 8, bottom row). Thus, the MMNs elicited by missed deviants in the ATTEND conditions were generally similar to those conditions in which deviants elicited MMN in the IGNORE conditions (i.e., Oddball, 11 ST, and 7 ST). Only the correctly identified targets elicited MMN in the 5 ST condition. Comparison of the MMN amplitude revealed a main effect of response accuracy ($F_{1,9} = 5.96$, $p < 0.05$), wherein the MMN amplitude was larger when deviants were

correctly identified than when missed, but there was no main effect of condition and no interaction.

N2b component

N2b was elicited along with MMN and P3b in the same ST conditions in which deviants were correctly identified (Oddball, 11, 7, and 5 ST; Fig. 8), but was either not elicited or was markedly smaller for incorrect responses. Mean amplitude in the N2b range was larger for correct than incorrect responses ($F_{1,9} = 24.819$, $p < 0.001$) but the N2b amplitude did not statistically differ as a function of ST difference ($F_{4,36} = 2.47$, $p = 0.10$, with no interaction).

P3b component

Deviants that were correctly identified elicited P3b in the Oddball, 11, 7, and 5 ST conditions (Fig. 8), the same conditions as those in which significant MMNs were elicited by correct responses. P3b was significantly elicited in the 11 and 7 ST conditions by incorrect trials but was not elicited in the 5 or 1 ST conditions. The repeated measures ANOVA comparing amplitude in the P3b range for response accuracy (correct vs. incorrect) and ST condition revealed a main effect of condition ($F_{4,36} = 6.86$, $p < 0.05$). Amplitude was larger in the Oddball, 11 ST, 7 ST for correct and incorrect responses and 5 ST for correct responses than the 1 ST for correct responses and 5 ST and 1 ST for incorrect responses. In addition, there was a main effect of response accuracy ($F_{1,9} = 18.14$, $p < 0.01$). P3b amplitude elicited by correct responses was larger than incorrect responses. Fig. 8 also displays the centro-parietal voltage distribution typical of the P3b component.

CHILDREN

ERPs

Fig. 6 (left column) displays the ERPs elicited by the deviant and control stimuli when the children attended to the sounds. P1 is prominently recorded in all ST conditions (Sussman et al., 2008). The obligatory components in the deviant ERP at Fz overlaps with the negativity involved in processing target detection, except for the 3 ST condition where the children generally heard one stream. The P3b (dashed line) is seen in the deviant ERPs most prominently at Pz.

Processing-contingent negativity (PCN)

It was not possible to distinguish components associated with target detection (i.e., MMN and N2b) in children in the same way as it is possible in adults. In adults, the N2b temporally follows MMN but precedes P3b. It also has a more centro-parietal scalp distribution than MMN. In children, however, the same temporal and topographic characteristics were not observed. The parietal P3b did not follow the negativity, but overlapped with the negativity seen at Fz. Thus, the distribution for target-detection components in children contrasted that found for adults, in which the negative followed by positive waveforms are separated in time. Without additional information about possible neural generators contributing to this response, it is not possible in children to differentiate with confidence that the negativity was MMN and N2b components. The negativity in children likely represents multiple processing-contingent components when sounds are attended. To be cautious, we will call the negativity elicited in the ATTEND experiments, a Processing-Contingent Negativity (PCN). We measured two peaks observed in the negative deflection at Fz elicited by deviant 'X' tones (see Table 3 – PCN1 and PCN2).

The average amplitude of the first peak negativity elicited by the deviant when it was correctly identified was compared with the average amplitude when the deviant was missed. These data are summarized in Fig. 9, along with voltage maps showing the scalp distribution. PCN1 was

elicited in the Oddball, 19, 11, and 7 ST conditions when correctly identified and in the Oddball and 19 ST conditions when incorrectly identified. The PCN1 amplitude of the Oddball was larger than in any other condition (main effect of condition, $F_{4, 44}=10.61$, $p<0.001$). PCN1 amplitude was not larger when elicited by correctly identified deviants (i.e., no main effect of response accuracy and no interaction). This contrasts the adult MMN findings, in which there was an effect of response accuracy showing a larger MMN amplitude when correctly identified.

The mean amplitude of the second peak in the negativity (PCN2) was significantly less than zero for correct responses in the 19 ST, 11 ST, and 7 ST conditions, and was not significantly present in the 3 ST or for any of the incorrect responses (Fig. 9). The amplitude of PCN2 did not differ among conditions ($F_{2, 22}<1$, $p=.54$).

P3b component

The P3b dominates the attended waveform in the children. P3b was elicited in all conditions when the deviant was correctly identified. P3b was also significantly elicited by deviant tones in the Oddball condition when the target was missed (Figs. 6, 9). The P3b amplitude was larger when the deviant stimulus was correctly identified (main effect of response accuracy $F_{1, 11}=25.0$, $p<0.001$), with no main effect of condition or interaction. Fig. 9 displays the voltage distribution for P3b, which shows a posterior focus similar to P3b elicited in adults.

Group comparison of P3b latency

Comparing the common active conditions between children and adults (Oddball, 11 ST, and 7 ST), the P3b peak latency was significantly earlier in children than adults, (main effect of group: $F_{1, 20}=337.6$, $p<0.0001$); the P3b latency increased with decreased frequency separation (main effect of condition: $F_{2, 40}=139.0$, $\epsilon=.96$, $p<0.0001$); and there was an interaction between group and condition ($F_{2, 40}=44.6$, $\epsilon=.96$, $p<0.0001$) because P3b latency increased with decreasing frequency separation for adults (Oddball $<11<7$), but in children P3b latency for the Oddball was shorter than both ST conditions, which did not significantly differ from each other (11 and 7 ST).

ANALYSIS OF THE CONTROL CONDITION ERPS

Although the comparison response used to delineate the MMN component had the same physical parameters as the deviant sound, the possibility exists that the response to the louder intensity sound for eliciting MMN may reflect an N1 enhancement rather than a true MMN. The strongest evidence against this explanation is the analysis of the control condition data, in which an intensity *decrement* was the deviant intensity 'X' tone. While these control conditions were originally intended to obtain a comparison waveform for delineating the MMN in the main conditions, they provide an excellent opportunity to examine the waveforms in a situation in which an intensity decrement rather than an intensity increment was the MMN-eliciting deviant (probe) tone. The ERP responses to the "standard" 'X' tones from the main experimental conditions (65 dB for children; 71 dB for adults) were subtracted from the ERP responses to the intensity decrement 'X' tone (65 dB, children; 71 dB, adults) from the control conditions. Even though fewer deviant tones were obtained per participant in each condition, this analysis revealed the same pattern of results in both the IGNORE and ATTEND conditions obtained with the louder intensity tones of the main experiment (Figures 10–11). Thus, MMN was elicited concordantly in passive and active conditions in adults, but not in children. In children, MMN was elicited only in the Oddball and 31 ST conditions (Fig. 10). It is rather striking that even with the lowered signal-to-noise ratio occurring because of a smaller number of deviants, the averaged responses appear nearly as robust as from the main conditions. However, the peak latencies appear to be longer in the control than those of the main conditions for the MMN (e.g., Oddball condition for Adults 172 vs. 123 ms, and for Children, 197 vs. 159

ms) and P3b (e.g., Oddball condition for Adults 337 vs. 290 ms, and for Children, 336 vs. 234 ms) components. This may be due to somewhat greater difficulty detecting an intensity decrement than an intensity increment.

DISCUSSION

The principal finding of this study was that there was a high degree of concordance between passive and active electrophysiological indices of stream segregation that matched with perception in the mature adult auditory system. In contrast, perception of stream segregation in children matched with active but not passive electrophysiological measures. Although the basic neural mechanisms of stream segregation may be operative in childhood, indicated by passive measures, the enhanced physiological response with task performance suggests that attention plays a major role in shaping the acuity of complex auditory scene analyses. We speculate that physiological changes brought about by attending to behaviorally relevant input may refine and optimize automatic auditory processes used in daily life.

Perception of stream segregation

The current study shows that 9–12 year-old children required a slightly larger frequency separation than adults to perceive two streams, which is consistent with our previous behavioral study from the same age group (Sussman et al., 2007). This result indicates that auditory mechanisms involved in sound organization are still immature in this age group even though by age 9, frequency discrimination abilities are adult-like (Thompson, Cranford, & Hoyer, 1999). This suggests that frequency discrimination acuity is not the only contributing factor to the ability to analyze the auditory scene.

Although children required a larger frequency separation to hear segregated streams compared to adults, they otherwise had similar behavioral functions for stream segregation. There was a specific frequency difference that led to the perception of clearly segregated, clearly integrated, and ambiguous region in both groups. Furthermore, when the streams were segregated, there were no significant differences in the ability of children compared to adults to detect the louder intensity tones within the 'X' stream. This is consistent with literature showing that intensity discrimination abilities are adult-like by three years of age (Berg & Boswell, 2000). Thus, the current and previous behavioral results indicate that maturation of complex sound patterns follows a longer developmental course than maturation of simple sound feature discrimination.

Comparing active and passive physiological measures with perception

In adults, active and passive measures indexing stream segregation were largely concordant and matched with perception. In the ST conditions that two streams were perceived, MMNs were elicited along with P3b when actively detecting deviants in the attended stream. In the ambiguous range for perceiving one or two streams (5 ST; consistent with the results of Carlyon, Cusack, Foxton, & Robertson, 2001; Sussman et al., 2007; van Noorden, 1975), MMN was not elicited in the passive state. Two-streams were passively indicated by MMN only at unambiguous frequency separations (7 ST and 11 ST). This contrasts the results during active segregation of the sounds, in which MMN was also elicited in the 5 ST condition, suggesting that attention sharpens pre-existing sound representations based on passive auditory system mechanisms.

Children's electrophysiological responses were qualitatively similar to adults in the active conditions. Electrophysiological indices of stream segregation were elicited in the same ST conditions that two streams were clearly perceived when children actively detected deviants (11 ST and above), as well as in the perceptually ambiguous region for children (7 ST). Thus,

as in adults, in general, there was good correspondence between perception of streams and active electrophysiological indices of stream segregation.

In contrast to the active conditions, electrophysiological measures in children were qualitatively different than adults in the passive conditions. According to behavior alone, children required a slightly larger frequency separation to hear two streams. This might lead to the expectation, as was shown with adults during passive listening, that a slightly larger frequency separation would also be needed for two streams to be indexed by MMN. Thus, MMN should have been elicited at about 11 ST, a clearly segregated organization according to behavior. However, in contrast to adults, in the passive IGNORE conditions MMN did not appear until 31 ST. Thus, there was a dramatic difference between the frequency separation required during passive and active processes for physiological indices to indicate two streams.

One interpretation for the difference between the MMN results in the passive and active conditions is that larger contrasts are required in passive than active listening conditions for children but not adults. This interpretation suggests that the key factor influencing the response is not stream segregation but attention, which leads to easier detection of sound feature contrasts. This interpretation is mitigated in the current study, given that MMN was elicited in some passive conditions but not others even though the loudness contrast itself was the same in all.

The close concordance in adults between behavioral and passive but mature electrophysiological indices of stream segregation shows that ERPs can provide reliable indices of sound pattern organization. The pronounced disparity between behavioral and passive electrophysiological indices of stream segregation in children indicates that passive auditory processing mechanisms of sound pattern organization are immature in children of this age. Physiological changes at meaningful frequency separations were seen in the children only when they attended to the task. The enhanced capacity of children to detect sound patterns both behaviorally and electrophysiologically, relative to passively obtained physiological measures, strongly suggests that attention plays a crucial role during normal development in shaping the automatic processes and strengthening the neural networks used for auditory scene analysis. Attention is thus proposed to be a necessary prerequisite during development for learning to process sounds in one's environment.

We cannot exclude the possibility that extended passive exposure to sounds during maturation is at least partially responsible for the better concordance in adulthood. Although in the adult brain, cortical plasticity is not induced solely by passive sound stimulation (Recanzone, Schreiner, & Merzenich, 1993), cortical plasticity to passive sound exposure has been observed in the developing animal (Zhang, Bao, & Merzenich, 2002). Nonetheless, passive experience is not always fully adequate for plastic changes in adulthood (Polley, Steinberg, & Merzenich, 2006). Thus, the specific ways in which passive auditory stimulation alters cortical fields during the different developmental stages remains open. Moreover, the connection between alterations in neural response properties and resulting perceptual abilities is largely unknown.

Correct versus incorrect target responses

The ERPs were sorted according to whether or not the intensity deviant 'X' tones were correctly identified. In adults, attending to a stream of sounds and detecting deviants elicited the P3b component along with a negativity that included overlap between MMN and N2b. Elicitation of this processing-contingent negativity (PCN) complex and P3b provided further indication that the sounds were physiologically segregated and deviants detected in the attended stream. These ERP responses (N2b/P3b) can also be elicited by attended non-targets (Novak et al., 1990; Sussman et al., 2002). In other words, noticing the deviants without making an overt response also can elicit these components. For the current study, this means that missed targets

might also elicit PCN and P3b when the sounds are segregated. This can be observed in the ERPs elicited by correct and incorrect responses. PCN and P3b were elicited by missed targets in the same ST conditions as MMNs were elicited in the IGNORE experiment (7 and 11 ST conditions). Even though the targets were behaviorally missed, the deviants were noticed at some level when the representation of streaming was already established. Thus, the final response decision occurred outside the physiological indication of active detection indexed by PCN and P3b. This is also suggested by the relationship between behavioral response latency and peak latency of the ERP components. RT latency partially overlapped with the peak latency of the P3b component for adults (e.g., both RT and P3b peaked at 336 ms in the 11 ST condition). Thus, the target detection process occurred prior to P3b. In contrast, the PCN peak latency occurred approximately 125 ms earlier than RT in all conditions, suggesting that it might reflect a component of the target detection process.

The attention-related components seen in the ERP of adults were not paralleled by similar potentials in children. First, while there was a demonstrative presence of a processing-contingent negativity (PCN), MMN and N2b could not be identified as separable components. Second, the peak latency of the P3b component was considerably earlier in children than in adults. Even at the slowest 300 ms SOA rate of the Oddball condition, the P3b had an earlier peak latency than the same condition in adults. Moreover, in children, the PCN and P3b overlapped in time, occurring approximately 250 ms from stimulus onset. RT was considerably longer than both PCN and P3b, occurring approximately 150 ms later than both of these physiological responses. Thus, in contrast to adults, both physiological indicators of target detection precede the behavioral response. The earlier P3b peak latency for children compared to adults contrasts with findings that suggest P3b latency gets shorter with maturation (e.g., Pearce, Crowell, Tokioka, & Pacheco, 1989; Zenker & Barajas, 1999). The current study differs from previous studies by using a rapid stimulus rate within a complex paradigm. The faster rate produced overlap of responses in children (Sussman et al., 2008), and may have affected the latency of the response. Further studies are clearly needed to characterize processing-contingent components associated with attention in children, and to ascertain their maturation to adult patterns in complex sound processing paradigms.

P3b was readily visible in the ERPs of children when two streams were perceived. P3b was also significantly elicited in the smallest 3 ST condition for correctly identified targets, even though PCN was not observed in that condition. This effect was observed with detection of only 30 deviants, thus showing the robustness of this response in children.

Notably, there were both similarities and dissimilarities between children and adults in relating neural activity and perception to correct and incorrect responses. Among the similarities, both PCN and P3b were elicited by incorrect responses. In children, when the targets were missed in the 19 ST condition, deviant 'X' tones still elicited a significant PCN. This suggests that once the sounds were segregated by attention, a PCN could be elicited by deviants even without a behavioral response. This result further indicates that the presence of these physiological measures may be necessary but are not sufficient to predict behavior. Furthermore, the average RT in both children and adults was later than the peak of the PCN. Therefore, together, these results suggest that PCN is either unrelated to behavior, or it precedes other neural processes that can then abort the decision to make an overt response.

The most remarkable dissimilarity observed for correct vs. incorrect responses between adults and children was that in adults the incorrect responses paralleled the passive measures; MMNs were elicited in conditions of 7 ST and above. In children, on the other hand, PCN was elicited at 19 ST for incorrect responses, but was not passively elicited until 31 ST. This suggests that during the IGNORE experiment, the sounds were not segregated by passive auditory processes at frequency separations less than 31 ST in children. That is, the automatic segregation that

occurred in adults during passive listening was not indicated in children. The results thus show that attention strongly facilitated automatic sound segregation and deviance detection, and that this facilitation occurred even when children did not identify targets by making an overt behavioral response, as long as they were actively listening to the sounds.

Relationship to disordered auditory processing

We suggest that attention, as a prerequisite of learning, is crucial for accurate sound organization. This takes on added significance when considering individuals who have difficulties with sound processing, such as children with language impairments or individuals with hearing loss. Although new theories stress the importance of attentional dysfunction in developmental language disorders (Hari & Renvall, 2001; Petkov, O'Connor, Benmoshe, Baynes, & Sutter, 2005; Stevens, Sanders, & Neville, 2006), our results suggest that attention is not just a general and non-specific factor influencing sound processing. Instead, results suggest that attentional mechanisms play a specific role in reorganizing passive, low-level acoustic processes during normal development. For children with disordered auditory processes, attentional impairments may hamper the ability to develop and refine these automatic sound organization processes and ultimately impact on perception.

Acknowledgements

This research was funded by the National Institutes of Health (grant #DC006003). We thank Dr. Katherine Lawson for performing and analyzing the psychometric tests. We thank Jean DeMarco for recruiting and scheduling the participants, and Wei Wei Lee for collecting the data.

References

- Alho K. Cerebral generators of mismatch negativity MMN and its magnetic counterpart (MMNm) elicited by sound changes. *Ear and Hearing* 1995;16(1):38–51. [PubMed: 7774768]
- Atienza M, Cantero JL, Dominguez-Marin E. The time course of neural changes underlying auditory perceptual learning. *Learning & Memory* 2002;93:138–150. [PubMed: 12075002]
- Berg KM, Boswell AE. Noise increment detection in children 1 to 3 years of age. *Perception & Psychophysics* 2000;62(4):868–873. [PubMed: 10883590]
- Carlyon RP, Cusack R, Foxton JM, Robertson IH. Effects of attention and unilateral neglect on auditory stream segregation. *Journal of Experimental Psychology Human Perception and Performance* 2001;27:115–127. [PubMed: 11248927]
- Clifton RK, Morrongiello JW, Kulig JW, Dowd JW. Newborns' orientation toward sound: possible implications for cortical development. *Child Development* 1981;52:833–838. [PubMed: 7285655]
- Demany L. Auditory stream segregation in infancy. *Infant Behavior & Development* 1982;5:261–276.
- Friedman D, Cycowicz YM, Gaeta H. The novelty P3: an event-related brain potential ERP sign of the brain's evaluation of novelty. *Neuroscience and Biobehavioral Reviews* 2001;25:355–373. [PubMed: 11445140]
- Fritz JB, Elhilali M, Shamma SA. Differential dynamic plasticity of A1 receptive fields during multiple spectral tasks. *Journal of Neuroscience* 2005;2533:7623–7635. [PubMed: 16107649]
- Fritz J, Shamma S, Elhilali M, Klein D. Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nature Neuroscience* 2003;611:1216–1222.
- Green, DM.; Swets, JA. *Signal detection theory and psychophysics*. New York: Wiley; 1966.
- Halgren B, Baudena P, Clarke JM, Heit G, Liegeois C, Chauvel P, Musolino A. Intracerebral potentials to rare target and distractor auditory and visual stimuli. I Superior temporal plane and parietal lobe. *Electroencephalography and Clinical Neurophysiology* 1995;94:191–220. [PubMed: 7536154]
- Halgren E, Marinkovic K, Chauvel P. Generators of the late cognitive potentials in auditory and visual oddball tasks. *Electroencephalography and Clinical Neurophysiology* 1998;1062:156–64. [PubMed: 9741777]

- Hari R, Renvall H. Impaired processing of rapid stimulus sequences in dyslexia. *Trends in Cognitive Science* 2001;52:525–532.
- Hubel DH, Henson CO, Rupert A, Galambos R. Attention units in the auditory cortex. *Science* 1959;129(3358):1279–1280. [PubMed: 13658956]
- Jacobsen T, Horenkamp T, Schröger E. Preattentive memory-based comparison of sound intensity. *Audiology & Neuro-otology* 2003;86:338–346. [PubMed: 14566104]
- Jasper HH. The 10–20 electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology* 1958;10:370–375.
- Javitt DC, Steinschneider M, Schroeder CE, Arezzo JC. Role of cortical N-methyl-D-aspartate receptors in auditory sensory memory and mismatch negativity generation: Implications for schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America* 1996;93:11962–11967. [PubMed: 8876245]
- Macmillan, NA.; Creelman, CD. *Detection theory: A user's guide*. Vol. 2. New York: Cambridge University Press; 2005.
- McAdams S, Bertoncini J. Organization and discrimination of repeating sound sequences by newborn infants. *The Journal of the Acoustical Society of America* 1997;102:2945–2953. [PubMed: 9373981]
- Morrongiello BA, Trehub SE. Age-related changes in auditory temporal perception. *Journal of Experimental Child Psychology* 1987;44:413–426. [PubMed: 3694124]
- Muir D, Field J. Newborn infants orient to sounds. *Child Development* 1979;50:431–436. [PubMed: 487883]
- Näätänen, R. *Attention and Brain Function*. New York: Erlbaum; 1992.
- Näätänen R, Schröger E, Karakas S, Tervaniemi M, Paavilainen P. Development of a memory trace for a complex sound in the human brain. *Neuroreport* 1993;4:503–506. [PubMed: 8513127]
- Näätänen R, Simpson M, Loveless NE. Stimulus deviance and evoked potentials. *Biological Psychology* 1982;14:53–98. [PubMed: 7104425]
- Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I. Pre-attentive cognitive processing “primitive intelligence” in the auditory cortex as revealed by the mismatch negativity MMN. *Trends in Neuroscience* 2001;24(5):283–288.
- Novak GP, Ritter W, Vaughan HG Jr, Wiznitzer ML. Differentiation of negative event-related potentials in an auditory discrimination task. *Electroencephalography and Clinical Neurophysiology* 1990;75:255–75. [PubMed: 1691075]
- Nozza RJ, Wilson WR. Masked and unmasked pure-tone thresholds of infants and adults: development of auditory frequency selectivity and sensitivity. *Journal of Speech and Hearing Research* 1984;27:613–622. [PubMed: 6521469]
- Pandya PK, Moucha R, Engineer ND, Rathbun DL, Vazquez J, Kilgard MP. Asynchronous inputs alter excitability spike timing and topography in primary auditory cortex. *Hearing Research* 2005;203(1–2):10–20. [PubMed: 15855025]
- Pearce JW, Crowell DH, Tokioka A, Pacheco GP. Childhood developmental changes in the auditory P300. *Journal of Child Neurology* 1989;4:100–106. [PubMed: 2715604]
- Petkov CI, O'Connor KN, Benmoshe G, Baynes K, Sutter ML. Auditory perceptual grouping and attention in dyslexia. *Cognitive Brain Research* 2005;24:343–354. [PubMed: 15993772]
- Polley DB, Steinberg EE, Merzenich MM. Perceptual learning directs auditory cortical map reorganization through top-down influences. *Journal of Neuroscience* 2006;26:4970–4982. [PubMed: 16672673]
- Recanzone GH, Schreiner CE, Merzenich MM. Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *Journal of Neuroscience* 1993;13:187–103. [PubMed: 8423485]
- Ruben RJ. The ontogeny of human hearing. *Acta Otolaryngologica* 1992;112:192–196.
- Rutkowski RG, Weinberger NM. Encoding of learned importance of sound by magnitude of representational area in primary auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America* 2005;102:13664–13669. [PubMed: 16174754]

- Sams M, Hari R. Magnetoencephalography in the study of human auditory information processing. *Annals of the New York Academy of Sciences* 1991;620:102–117. [PubMed: 2035937]
- Scherg M, Vajsa J, Picton TW. A source analysis of the late human auditory evoked potentials. *Journal of Cognitive Neuroscience* 1998;1:336–355.
- Schneider W, Shiffrin RM. Controlled and Automatic human information processing: I Detection search and attention. *Psychological Review* 1977;84:1–66.
- Sinnott JM, Pisoni DB, Aslin RN. A comparison of pure tone auditory thresholds in human infants and adults. *Infant Behavior & Development* 1983;6:3–17.
- Stevens C, Fanning J, Coch D, Sanders L, Neville H. Neural mechanisms of selective auditory attention are enhanced by computerized training electrophysiological evidence from language-impaired and typically developing children. *Brain Research* 2008;1205:55–69. [PubMed: 18353284]
- Stevens C, Sanders L, Neville H. Neurophysiological evidence for selective auditory attention deficits in children with specific language impairment. *Brain Research* 2006;1111(1):143–152. [PubMed: 16904658]
- Sussman E. Integration and segregation in auditory scene analysis. *Journal of the Acoustical Society of America* 2005;117:1285–1298.
- Sussman E. A New View on the MMN and Attention Debate: The Role of Context in Processing Auditory Events. *Journal of Psychophysiology* 2007;21(3–4):164–175.
- Sussman E, Cepenione R, Shestakova A, Näätänen R, Winkler I. Auditory stream segregation processes operate similarly in school-aged children as adults. *Hearing Research* 2001;153:108–114. [PubMed: 11223301]
- Sussman E, Ritter W, Vaughan HG Jr. Attention affects the organization of auditory input associated with the mismatch negativity system. *Brain Research* 1998;789:130–138. [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, Gumenyuk V, Grushko J, Lawson K. The maturation of human evoked brain potentials to sounds presented at different stimulus rates. *Hearing Research* 2008;236:61–79. [PubMed: 18207681]
- Sussman E, Winkler I, Houtilainen M, Ritter W, Näätänen R. Top-down effects on the initially stimulus-driven auditory organization. *Cognitive Brain Research* 2002;13:393–405. [PubMed: 11919003]
- 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. *Hearing Research* 2007;225:117–127. [PubMed: 17300890]
- Sutton S, Braren M, Zubin J, John ER. Evoked potential correlates of stimulus uncertainty. *Science* 1965;150:1187–1188. [PubMed: 5852977]
- Tallal P, Miller SL, Bedi G, Byma G, Wang X, Nagarajan SS, Schreiner C, Jenkins WM, Merzenich MM. Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science* 1996;271(5245):81–84. [PubMed: 8539604]
- Thompson NC, Cranford JL, Hoyer E. Brief-tone frequency discrimination by children. *Journal of Speech and Hearing Research* 1999;42:10681–1068.
- Tiitinen H, May P, Reinikainen K, Näätänen R. Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature* 1994;372(6501):91–92.
- Trehub SE, Schneider BA, Endmand M. Developmental changes in infants' sensitivity to octave-band noises. *Journal of Experimental Child Psychology* 1980;29:282–293. [PubMed: 7365426]
- Tremblay KL, Kraus N. Auditory training induces asymmetrical changes in cortical neural activity. *Journal of Speech, Language, and Hearing Research* 2002;45:564–572.
- van Noorden, LPAS. Temporal coherence in the perception of tone sequences. Unpublished doctoral dissertation: Eindhoven University of Technology; 1975.
- Weinberger NM. The nucleus basalis and memory codes: auditory cortical plasticity and the induction of specific associative behavioral memory. *Neurobiology of Learning and Memory* 2003;80:268–284. [PubMed: 14521869]
- Werner LA, Marean GC, Halpin CF, Spetner NB, Gillenwater JM. Infant auditory temporal acuity: gap detection. *Child Development* 1992;63:260–272. [PubMed: 1611932]

- Wightman F, Allen P, Dolan T, Kistler T, Jamieson D. Temporal resolution in preschool children. *Child Development* 1989;60:611–624. [PubMed: 2737011]
- Zenker F, Barajas JJ. Auditory P300 development from an active, passive and single-tone paradigms. *International Journal of Psychophysiology* 1999;33:99–111. [PubMed: 10489075]
- Zhang LI, Bao S, Merzenich MM. Persistent and specific influences of early acoustic environments on primary auditory cortex. *Nature Neuroscience* 2001;4(11):1123–1130.
- Zhang LI, Bao S, Merzenich MM. Disruption of primary auditory cortex by synchronous auditory inputs during a critical period. *Proceedings of the National Academy of Sciences of the United States of America* 2002;99(4):2309–14. [PubMed: 11842227]
- Zhou X, Merzenich MM. Intensive training in adults refines A1 representations degraded in an early postnatal critical period. *Proceedings of the National Academy of Sciences of the United States of America* 2007;104(40):15935–15940. [PubMed: 17895375]

Figure 1. Schematic diagram of the stimulus paradigm

Rectangles labeled 'X' and 'O' represent tones. The abscissa shows the passage of time and the ordinate indicates the frequency scale. The shading of the rectangles represents the intensity value of the tones. **A. Oddball sequence.** 'X' tones are presented at the pace they occur in the semitone conditions, but without the intervening 'O' tones. The deviant (louder intensity probe tone) is displayed in the 4th position. This condition served as a control, determining the ability to detect louder intensity tones without intervening tones, indexed by behavioral and ERP measures. **B. Semitone (ST) Conditions.** The 'X' tones were a fixed frequency. The conditions are created by varying the spectral distance of the intervening 'O' tones from the 'X' tones. When the 'O' tones are near in frequency, the 'X' and 'O' tones are perceived as integrated into one stream (XOOXOO), whereas when the 'O' tones are distant in frequency from the 'X' tones, they are perceived as being in their own stream: segregated from the 'X' tones, which occur at the same onset-to-onset pace as in the Oddball condition (X—X—X—X—). The 'X' tone intensity was fixed, with the probe tone 12 dB higher in intensity for adults and 15 dB higher for children. The intervening 'O' tones varied in four different values above and below the intensity value of both of the 'X' tones, with an equiprobable distribution among them. Thus, when the tones segregate by frequency can the louder 'X' tone (probe) be distinguished from the other 'X' tones.

Figure 2. Behavioral Tasks

A-B. One vs. Two Streams. A visual depiction was used along with sound sequences to illustrate one vs. two streams for practice in the behavioral experiment that participants reported how they perceived each sound sequence. **C. Deviance Detection Task.** Visual depiction used to illustrate the deviance detection task, where gray values illustrate the soft vs. louder intensity of the 'X' tones with an arrow pointing to the target tone.

Behavioral Results

—○— Adults
—□— Children

Figure 3. Behavioral Results

The behavioral results are displayed for the perception of One vs. Two Streams (left panel) and for the Deviance Detection Task (right panel) for Adults (blue dashed line) and Children (black solid line). The abscissa displays the ST conditions in each respective age group. The ordinate displays the proportion that two streams were heard (left) and the hit rate for deviance detection (right). Children needed a slightly larger frequency separation between 'X' and 'O' tones to perceive two streams or detect 'X' tone deviants, but generally performed similarly as adults.

Figure 4. Event-related brain potentials for Adults

The grand-averaged ERPs elicited at Fz by the deviant (thick solid line) and control (thin solid line) tones are displayed for the ATTEND (left) and IGNORE (right) experiments. The average of the correct responses are displayed for ERPs elicited by the deviant in the ATTEND conditions. The P1 peak is notable in each condition, peaking at approximately 100 ms in the ST conditions when the sounds were ignored (right). For the ATTEND conditions, the ERP elicited by the deviant is also displayed at the Pz electrode (dashed line) to show the P3b (denoted with an arrow).

Figure 5. IGNORE Conditions for Adults

The five conditions of the IGNORE experiment are displayed for the adults, showing the difference waveforms at Fz (thick solid line) and the left mastoid (LM; thick dashed line). Difference waveforms were obtained by subtracting the ERPs elicited by the standard (louder 'X' tones from the control conditions of the IGNORE experiments) from the ERPs elicited by the deviant (louder 'X' tones in the corresponding experimental conditions). Statistical significance is reported with two stars (<0.01) or one star (<0.05), representing the p values. Arrows point to the significantly elicited MMNs and to the inversion seen at the mastoid. Maps displaying the scalp distribution of the voltage taken at the peak latency of the MMN component are displayed above each waveform. Negative voltage is represented with a blue shade and positive voltage with a red shade; the deeper the color, the greater the voltage value. The bilateral, frontal distribution of the MMN is seen in the Oddball, 11, and 7 ST conditions. The MMN elicited in the Oddball condition is largest in amplitude and earlier in latency than those elicited in the ST conditions.

Figure 6. Event-related brain potentials for Children

The grand-averaged ERPs elicited at Fz by the deviant (thick solid line) and control (thin solid line) tones are displayed for the ATTEND (left) and IGNORE (right) experiments. The average of the correct responses are displayed for ERPs elicited by the deviant in the ATTEND conditions. The 31 ST condition, in which a significant difference between the deviant and control was found, is displayed for the IGNORE experiment in addition to the conditions that were common to both. The P1 peak is notable in each condition, most prominently at the rate of the 'X' tones (300 ms SOA). For the ATTEND conditions, the ERP elicited by the deviant is also displayed at the Pz electrode (dashed line) to show the P3b (denoted with an arrow).

CHILDREN IGNORE

Figure 7. IGNORE Conditions for Children

The eleven conditions of the IGNORE experiment is displayed for the children, showing the difference waveforms at Fz (thick solid line) and the left mastoid (LM; thick dashed line). Difference waveforms were obtained by subtracting the ERPs elicited by the standard (louder 'X' tones from the control conditions of the IGNORE experiments) from the ERPs elicited by the deviant (louder 'X' tones in the corresponding experimental conditions). Statistical significance is reported with two stars (<0.01), one star (<0.05), or a cross (<0.10), representing the *p* values. Arrows point to the significantly elicited MMNs and to the inversion seen at the mastoid. Maps displaying the scalp distribution of the voltage taken at the peak latency of the MMN component (or at the largest negative peak, where statistical measurements were made) are displayed above each waveform. Negative voltage is represented with a blue shade and positive voltage with a red shade; the deeper the color, the greater the voltage value. The bilateral, frontal distribution of the MMN is seen in the Oddball, 11, and 7 ST conditions. The MMN elicited in the Oddball condition is largest in amplitude and earlier in latency than those elicited in the ST conditions.

Figure 8. ATTEND Conditions for Adults

The five conditions of the ATTEND experiment are displayed for the adults, showing the difference waveforms (obtained by subtracting the ERPs elicited by the louder, standard 'X' tones of the control conditions from the ERPs elicited by the louder, deviant 'X' tones of the corresponding experimental conditions) at Fz (thick solid line), Pz (thin solid line) and the left mastoid (LM, thick dashed line). ERPs are displayed according to whether they were correctly identified (Correct vs. Incorrect). Note that the epoch duration is longer than in the other figures (−100 to 600 ms) to accommodate the longer P3b peak latency. Statistical significance is reported with two stars (<0.01) or one star (<0.05), representing the *p* values. The arrows point to significantly elicited components, from the electrodes where they are best observed. Maps displaying the scalp distribution of the voltage taken at the peak latency of the MMN and P3b components are displayed above each waveform. Negative voltage is represented with a blue shade and positive voltage with a red shade; the deeper the color, the greater the voltage value. The bilateral, frontal distribution of the MMN is seen in the Oddball, 11, 7, and 5 ST conditions. The broad, posterior distribution of the P3b is seen in the same conditions in which MMN was elicited.

CHILDREN ATTEND

Figure 9. ATTEND Conditions for Children

The five conditions of the ATTEND experiment are displayed for the children, showing the difference waveforms (obtained by subtracting the ERPs elicited by the louder, standard ‘X’ tones of the control conditions from the ERPs elicited by the louder, deviant ‘X’ tones of the corresponding experimental conditions) at Fz (thick solid line), Pz (thin solid line) and the left mastoid (LM, thick dashed line). ERPs are displayed according to whether they were correctly identified (Correct vs. Incorrect). Statistical significance is reported with two stars (<0.01) or one star (<0.05), representing the *p* values. The arrows point to significantly elicited components, from the electrodes where they are best observed. PCN was measured at an early and later peak (PCN1 and PCN2). Maps displaying the scalp distribution of the voltage taken at the peak latency of the PCN1 and P3b components are displayed above each waveform. Negative voltage is represented with a blue shade and positive voltage with a red shade; the deeper the color, the greater the voltage value. The right, frontal distribution of the PCN is seen in the Oddball, 19, 11, and 7 ST conditions. The tight posterior distribution of P3b is seen with overlap of the PCN.

IGNORE CONDITIONS

SOFTER INTENSITY DEVIANTS

Figure 10. Softer intensity deviants in IGNORE conditions

Difference waveforms, obtained by subtracting the ERPs elicited by the softer, standard 'X' tones in the experimental conditions of the IGNORE experiments from the ERPs elicited by the intensity decrements (softer deviant 'X' tones) in the corresponding control conditions, are displayed. A subset of the data is shown to demonstrate the pattern of MMN results (indicated with arrows) similarly found for the louder intensity 'X' deviants of the main experiment. Children's difference waveforms for the Oddball (thick solid line), 31 ST (thin solid line), and 27 ST (thin dashed line) conditions are displayed in the top row. The bottom row displays the adults for the Oddball (thick solid line), 11 ST (thin solid line), 7 ST (short dashed line), and 1 ST (long dashed line) conditions.

Figure 11. Softer intensity deviants in ATTEND conditions

Difference waveforms, obtained by subtracting the ERPs elicited by the softer, standard 'X' tones in the experimental conditions of the ATTEND experiments from the ERPs elicited by the intensity decrements (softer deviant 'X' tones) in the corresponding control conditions, are displayed. A subset of the data is shown to demonstrate the pattern of MMN and P3b results (indicated with arrows) similarly found for the louder intensity 'X' deviants of the main experiments. Children's difference waveforms for the Oddball (left column), 19 ST (middle column), and 3 ST (right column) conditions are displayed in the top row. Adult difference waveforms are displayed in the bottom row showing the Oddball (left), 11 ST (middle), and 1 ST (right). The mastoid trace (thick dashed line) is displayed to delineate the peak of the MMN, which is overlapped with N2b in the Fz trace (thick solid line). P3b is best delineated with the Pz trace (thin solid line).

Table 1

Assignment of Participants to Experiments

CHILDREN - Ignore

ST Conditions	Age Range	Mean (SD)	Count	M/F
3, 5, 7, 9	9–12 years	10 y 7 mo (1.2)	10	1/9
11, 15, 19	9–12 years	10 y 9 mo (1.1)	12	5/7
23, 27, 31	9–12 years	10 y 3 mo (0.8)	12	4/8

CHILDREN - Attend

ST Conditions	Age Range	Mean (SD)	Count	M/F
3, 7, 11, 19	10–12 years	11 y (.7)	12	4/8

ADULTS – Attend + Ignore

ST Conditions	Age Range	Mean (SD)	Count	M/F
1, 5, 7, 11	22–46	31 y 7 mo (7.2)	10	6/4

Table 2
Correct vs. Incorrect ERP responses (after artifact rejection).

		DEVIANCE DETECTION							
			Oddball	19	11	07	05	03	01
ADULTS	Correct		136	--	117	112	90	--	42
	Incorrect		6	--	28	34	53	--	102
CHILDREN	Correct		154	139	129	113	--	30	--
	Incorrect		13	31	44	61	--	130	--

Table 3

Peak Latencies of ERP Components

Adults				
Condition	Ignore MMN (Fz)	MMN (Lm)	Attend N2b (Cz)	P3b (Pz)
Oddball	123 (14)	106 (14)	154 (15)	290 (10)
11 ST	177 (15)	140 (16)	202 (14)	336 (18)
7 ST	171 (17)	175 (16)	211 (18)	410 (18)
5 ST	•	207 (19)	221 (15)	426 (16)
1 ST	•	•	•	•
Children				
Condition	Ignore MMN (Fz)	PCN1 (Fz)	Attend PCN2 (Fz)	P3b (Pz)
Oddball	159 (15)	153 (14)	•	234 (18)
31 ST	236 (16)	--	--	--
27 ST	266 (17)	--	--	--
19 ST	•	172 (12)	256 (14)	222 (15)
11 ST	•	166 (16)	251 (16)	285 (16)
7 ST	•	190 (16)	258 (16)	286 (15)
3 ST	•	•	•	391 (19)

Note: Peak mean latencies in milliseconds (and standard deviation) of the MMN component in the IGNORE experiment and of the MMN, N2b (Adults), two peaks of the PCN (Children), and P3b components for *correct responses* in the ATTEND experiment, in each semitone (ST) condition for each age group.

• denotes that the component was not significantly elicited.

-- denotes that the condition was not conducted in that experiment. The electrode used to measure the presence of the components is listed in parentheses under the component name.

Table 4

Deviance Detection Task

Condition	ADULTS					CHILDREN				
	1	5	7	11	Odd	3	7	11	19	Odd
HR (SD)	26% (23)	65% (18)	79% (16)	81% (10)	96% (2)	18% (16)	63% (17)	74% (16)	82% (10)	92% (3)
d' (SD)	1.12 (.85)	2.06 (2.13)	3.59 (.49)	3.67 (.49)	4.58 (.37)	1.53 (.51)	3.17 (.57)	3.47 (.62)	3.78 (.39)	4.10 (.34)
RT (SD)	572 (90)	409 (65)	368 (80)	336 (70)	281 (57)	526 (102)	437 (79)	409 (82)	397 (98)	361 (98)

Note: Hit Rate - **HR**; d prime - **d'**; Reaction Time - **RT**; Standard Deviation - **SD**