The Official Journal of the International Society on Infant Studies
Infancy

Infancy, ***16***(6), 655–668, 2011

ISSN: 1525-0008 print / 1532-7078 online
DOI: 10.1111/j.1532-7078.2011.00067.x

WILEY-BLACKWELL

Auditory Stream Segregation Improves Infants' Selective Attention to Target Tones Amid Distracters

Nicholas A. Smith
Perceptual Development Laboratory
Boys Town National Research Hospital

Laurel J. Trainor
Department of Psychology, Neuroscience and Behaviour
McMaster University

This study examined the role of auditory stream segregation in the selective attention to target tones in infancy. Using a task adapted from Bregman and Rudnicky's 1975 study and implemented in a conditioned head-turn procedure, infant and adult listeners had to discriminate the temporal order of 2,200 and 2,400 Hz target tones presented alone, preceded and followed by 1,460 Hz flanker tones, and presented within a series of 1,460 Hz captor tones meant to release the target tones from the effects of the flankers by capturing the flankers into a separate stream. Infants showed the same pattern of discrimination across conditions as adults: discrimination of target tones in the target-alone condition, a decrease in performance when flanker tones were introduced, and a return to target-alone level in the captor condition. These results suggest that infants' perceptual organization of tones is similar to that of adults, and that their ability to selectively attend to target sounds and ignore distracters depends on the structural properties and perceptual organization of the nontarget sounds.

Most acoustical environments consist of multiple sound sources, whose sound waves combine into a complex mixture in the ear. Attending to a

Correspondence should be sent to Nicholas A. Smith, Boys Town National Research Hospital, 555 North 30th St., Omaha, NE 68131. E-mail: nicholas.smith@boystown.org

sound source of interest can become difficult for a number of reasons. ‘‘Energetic masking’’ occurs when other sounds interfere with how the target sound is encoded in the auditory periphery (i.e., the inner ear or cochlea). However, sounds that produce no energetic masking effect on the target, for example, those with very different spectral properties, can nevertheless interfere with the perception of the target—a phenomenon referred to as ‘‘informational masking’’ (Durlach et al., 2003; Neff & Green, 1987; Pollack, 1975; Watson & Kelly, 1981). Although the target sounds are encoded in the periphery, listeners have difficulty selectively attending to these sounds because of the distracting effects of the other sounds.

Stimulus uncertainty is an important factor in informational masking. In most psychophysical tasks, adults have lower thresholds and better performance when the target signal is presented at an expected frequency and time (Egan, Greenberg, & Schulman, 1961; Greenberg & Larkin, 1968). However, this ability to selectively attend to target sounds appears to be immature in infants. Bargones and Werner (1994) compared the ability of 7- to 9-month-old infant and adult listeners to detect tones at expected and unexpected frequencies. They found that whereas adult listeners performed better with expected frequencies, infants did not, and even outperformed adults at the detection of tones outside the expected frequency. Other work has shown that for tone detection in broadband noise, 6-month-old infants’ performance suffers more than can be accounted for by energetic masking, suggesting an increased susceptibility to nonsensory, distraction effects (Werner & Bargones, 1991). Distracting sounds have also been found to impede 6- to 8-month-old infants’ perception of speech sounds (Polka, Rvachew, & Molnar, 2008). This study examines the role of perceptual organization in infants’ ability to selectively attend to changes in target tones in the presence of other distracting tones.

Bregman (1990) has used the term ‘‘auditory scene analysis’’ to refer to the process of perceptually organizing sound. The goal of auditory scene analysis is to isolate and identify individual sound sources from the incoming mixture of sounds. This goal is achieved using different kinds of processes. According to Bregman, primitive processes are stimulus-driven and are thought to function in a preattentive and presumably unlearned way. They are akin to Gestalt grouping principles in that they segregate the auditory scene (or group elements within the scene) into ‘‘streams’’ that correspond to objects or sound sources in the world. In contrast, in schema-based processes, listeners use their knowledge of stimulus structure (i.e., linguistic, musical) to pull familiar sound sources out of the mixture.

Although stream segregation can be achieved on the basis of a variety of stimulus features (Bregman, 1990), some of the earliest and most robust demonstrations have used frequency differences. One basic finding is that as the frequency separation between rapidly alternating tones increases, listeners find it more difficult to hear the tones as a coherent pattern, and the high- and low-frequency tones separate into isolated streams (Bregman & Campbell, 1971; Miller & Heise, 1950; van Noorden, 1975).

Stream segregation is difficult to assess in infants. Early demonstrations of streaming in adults relied on phenomenological and subjective methods (van Noorden, 1975) that could not be translated to experimental paradigms suitable for infants. One approach to evaluating infants' stream segregation abilities has been to examine infants' dishabituation to changes in tone patterns that vary in their discriminability depending on whether they are perceptually organized as one or two streams. Demany (1982) constructed tone sequences that were more easily discriminated from their retrogradations (i.e., the sequences played backward) when they were perceived as a single integrated stream. Using a visual habituation procedure, he found that 7- to 15-week-old infants' dishabituation patterns to changes in the order of the sequence corresponded to adults' discrimination responses, suggesting that infant stream segregation is similar to that of adults. Fassbender (1993) replicated and extended Demany's study, using a high amplitude sucking procedure. Infants, 2–5 months of age, were habituated to either an ascending or descending series of target tones, which were interleaved with random tones from the same frequency range. The interleaved tones make it very difficult for adults to discriminate whether the sequence is ascending or descending. The same is true for infants in that when the order of the target tones was reversed, infants did not dishabituate to the retrogradation. However, they did dishabituate to the retrogradation when the random interleaving tones were shifted up in frequency, decreased in amplitude, or changed in timbre from sine tones to complex tones. Fassbender thus argued that changing the stimulus properties of the random tones caused infants to perceive the target and random tones in different streams, similarly to adults.

The earliest behavioral evidence for auditory perceptual organization in infants comes from McAdams and Bertoncini (1997) who tested 3- to 4-day-old infants. In a sucking paradigm, infants heard repeating four-note ascending and descending melodic sequences comprised of notes of two different timbres presented from speakers at two different locations. When three of the four notes had the same timbre/location infants could discriminate ascending and descending contours, presumably because infants integrated the tones on the basis of timbral/spatial cues. However, when the same melodic contours were presented using notes that alternated in timbre/location, infants failed to discriminate the rising and falling contours, suggesting that infants segregated the melodies into separate streams.

The presence of stream segregation ability at birth is corroborated by electrophysiological methods. Winkler et al. (2003) tested newborns in an event-related potential study in which infants showed a mismatch negativity (MMN) response to infrequent intensity increases in a series of 1,813 Hz tones. When intervening tones of variable intensity and frequency were added, the MMN response disappeared, but only when the intervening tones were presented in the same frequency range as the target tones. When the intervening tones were presented two to three octaves lower than the target, the MMN persisted, suggesting that the segregation of the target and intervening tones into different streams allowed infants to detect the intensity changes in the target series.

Together these studies suggest that even very young infants have the basic capacity to organize the auditory scene in a similar manner as adults, though other work has shown that the development of these abilities continues through childhood (Sussman, Wong, Horvath, Winkler, & Wang, 2007). However, because the existing research has either used habituation to test sensitivity to changes in the sound pattern as a whole, or used electrophysiological methods to measure neural responses in a passive task, the role that stream segregation plays in infant selective attention remains largely unexplored. One primary question is whether selective attention, which appears to be immature in the case of tone detection in noise (Bargones & Werner, 1994), exhibits similar immaturity in other tasks, in which stream segregation of the target may facilitate performance. For example, the ability to detect changes in clearly audible, suprathreshold, target tones presented amid distracter tones may benefit from the segregation of the target tones into a separate stream.

The role of attention in the formation and maintenance of auditory streams is complex (Carlyon, 2004). Bregman (1990) has described stream segregation as a preattentive process, arguing that ‘‘stream formation has the job of grouping sounds that have arisen from the same source and attention must select distinct sources for further analysis’’ (p. 208). However, although studies showing an MMN response to streaming stimuli suggest that while attention is not required, the MMN response is modulated by attention (Alain & Izenberg, 2003; Snyder, Alain, & Picton, 2006; Sussman, Horváth, Winkler, & Orr, 2007; Sussman, Ritter, & Vaughan, 1999). Furthermore, in behavioral studies, the formation of auditory streams is reduced when listeners are asked to perform a competing attention demanding task (Carlyon, Cusack, Foxton, & Robertson, 2001).

One line of evidence supporting some degree of independence of streaming and attention comes from experiments that show that the perceptual organization of nonattended streams facilitates the processing of an attended target stream. One example of this is comodulation masking release (CMR). In CMR the effect of a narrow-band noise masker centered on a target tone is reduced when its temporal envelope fluctuates with the temporal envelope of noise bands remote in frequency (Hall, Haggard, & Fernandes, 1984). These covarying elements are perceptually organized into a separate stream, thereby reducing their effects on the target tones. This phenomenon has been shown in children (Grose, Hall, & Gibbs, 1993; Hall, Grose, & Dev, 1997), though it has not been tested in infants.

Bregman and Rudnicky's (1975) experiment provides another example of how the organization of a nonattended stream facilitates the processing of an attended stream. They presented listeners with pairs of short target tones that differed in frequency. Listeners were asked to judge whether the temporal order of target tones differed between the standard and comparison interval. In the comparison interval, the target tones were immediately preceded and followed by lower-frequency flanker tones, which obscured the temporal order of the intervening target tones. However, the disruptive effect of the flanker tones was mitigated when the comparison sequence was presented within a series of "captor" tones that matched the flankers in terms of frequency and temporal regularity. The flankers were captured into a separate stream, thereby releasing the target tones from interference.

This study consists of a translation of Bregman and Rudnicky's (1975) study to an infant paradigm in which we test infants' temporal order discrimination under three conditions using a conditioned head-turn procedure. If stream segregation is operational in infants, then a similar pattern of results to those found in adults would be expected: successful discrimination of the target tone order when presented alone, a drop in performance when the target tones are surrounded by flanker tones, and a recovery in performance when captor tones are introduced.

One advantage of this paradigm is that it is possible to evaluate streaming in terms of how it affects performance on a psychophysical task. Unlike previous studies (Demany, 1982; Fassbender, 1993; McAdams & Bertoncini, 1997) streaming can be demonstrated by improved discrimination, rather than decreased discrimination, thereby permitting the examination of how perceptual organization interacts with attention and other cognitive processes. Furthermore, direct comparisons between infants and adults can be made using the same stimuli and procedure, which was not possible with earlier studies of auditory scene analysis in infants using habituation.

METHOD

Participants

Eighteen adults and 40 infants participated. The adult listeners had a mean age of 39.5 years (*SD* = 8.9 years), and reported normal hearing. Each adult completed all three conditions. Infant subjects were between 6.5 and

8 months of age. The decision to test infants of this age was made in part for methodological reasons—they are ideally suited to the conditioned head-turn procedure (Werker, Polka, & Pegg, 1997)—and in part in order to facilitate comparisons to some previous studies (Bargones & Werner, 1994; Polka et al., 2008; Werner & Bargones, 1991). Infants were assigned to one of two groups, described below: captor group (n = 20, M age = 7.50 months, SD = 0.52 months) and flanker group (n = 20, M age = 7.30 months, SD = 0.57 months). The two infant groups did not differ significantly in age, $t(38) = 1.11$, *ns*. They were recruited through Boys Town Pediatrics Clinics in the Omaha area, where they had visited pediatricians for well-baby check-ups. All infants had passed newborn hearing screening tests.

An additional 60 infants were tested but are not reported here because they either did not pass training (n = 27) or failed to complete two full blocks of trials because of fussiness, crying, or inattention (n = 33). High attrition rates, up to 50%, are common in conditioned head-turn procedures, particularly when the stimuli are difficult to discriminate (Werker et al., 1997). Our procedure required that each infant transition seamlessly through training and multiple test blocks to ensure that within-subject comparisons could be made, and also to make it clear that the target stimulus and the task remained the same across blocks. However, these high demands likely contributed to the high attrition rate in this study.

Stimuli and procedure

The goal of this experiment was to compare infants' stream segregation with that of adults by translating the stimulus paradigm designed by Bregman and Rudnicky (1975) to a conditioned head-turn procedure capable of evaluating discrimination in infants. All tones were 65 msec long, including 10 msec linear onset and offset ramps and were presented at 48 dB SPL. These tones were arranged into three different configurations described below and illustrated in Figure 1. In all cases, the listener's task was to discriminate, or detect changes in, the temporal order of two target tone pairs presented within a continuously repeating series of background stimuli.

In the first block, all subjects completed the target-alone condition in which the target tones consisted of two immediately successive tones of 2,200 and 2,400 Hz. These frequencies are discriminable, even to young infants (Olsho, Koch, Halpin, & Carter, 1987; Sinnott & Aslin, 1985). The background stimuli consisted of these tones in low to high order. These pairs of target tones were repeated with a stimulus onset asynchrony of 975 msec. On change trials the order of the tones was reversed (i.e., high to low order). On control trials, the tones were identical to the background stimuli.

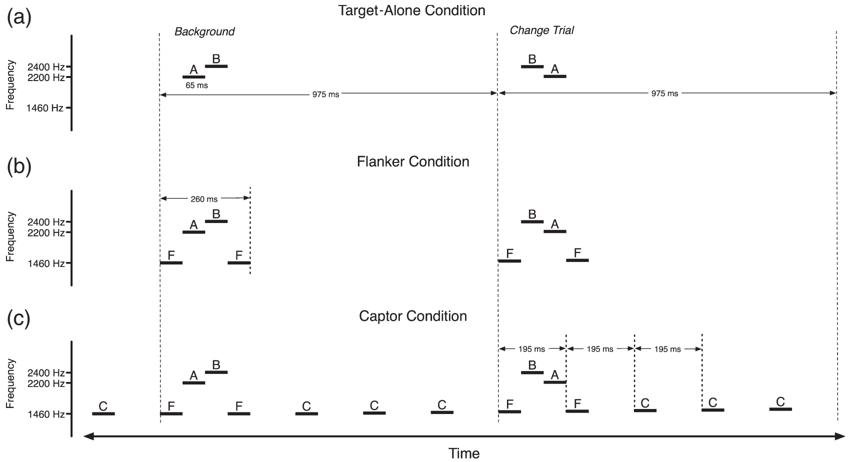

Figure 1 A schematic illustration of the configuration of stimuli in the three conditions tested. In the target-alone condition (a), background stimuli consisting of pairs of tones (A and B) are repeated. A change trial is shown in which the order of target tones is reversed. In the flanker condition (b) the target tones are immediately preceded by tones (F). In the captor condition (c) additional tones (C) are introduced, which together with the flanker tones form an isochronous rhythm.

In the second and third blocks adults completed both the flanker and captor conditions, with the order of conditions counterbalanced across subjects. Infant subjects performed a second, but not a third block, during which half the infants (the flanker group) completed the flanker condition, and the other half (the captor group) completed the captor condition.[1] In the flanker condition the pair of target tones was presented in the same manner as in the target-alone condition, but each pair was flanked by 65 msec 1,460 Hz tones immediately preceding and following the target tones. The captor condition was identical to the flanker condition, but with additional 1,460 Hz tones between successive target pairs. Their timing was such that these tones created an isochronous rhythm with the flanking tones, with a stimulus onset asynchrony of 195 msec. Because the 1,460 Hz tones immediately before and after the target tones are the same in the flanker and captor condition, any forward or backward masking effects on discrimination of the target tones are presumed to be the same.

[1]Our original intention was to use the same experimental design for infant as with adults, with each infant performing in each of the three experimental conditions. However, too few infants were able to make it through the third block of trials, because of fussiness or boredom, that we decided to use a two-block design and compare performance in the flanker and captor conditions between-subjects.

The experiment was conducted in a single-walled 10 × 10′ sound attenuated booth (Industrial Acoustic Company, Bronx, NY). Participants were seated facing the experimenter; infants on their parents' lap. The stimuli were generated and the experiment was controlled by a Mac Pro computer (Apple, Inc., Cupertino, CA) running MAX/MSP software (Cycling '74, San Francisco, CA), located outside the booth. The sound was produced by a Roland Edirol FA101 external firewire audio interface (Roland Corporation, Los Angeles, CA), which routed the signal through a Crown D-75A amplifier (Crown Audio, Inc., Elkhart, IN), and delivered the sound to the listeners by a GSI audiometric loudspeaker (Grason-Stadler, Eden Prairie, MN). The speaker was located on the listeners' right-hand side, approximately 175 cm away. Five-sec long, silent reinforcing animations were presented on an LCD display located just below the speaker.

A conditioned head-turn procedure was used, and is described below. For the sake of consistency and comparability, adults were tested in the same manner as infants, though they were asked to raise their hand to indicate their detection of a change.

Prior to beginning test trials, infants and adults performed a short training session with the target-alone stimuli. To facilitate training, the target stimuli on change trials were highlighted by making them 6 dB louder (54 dB SPL) than the background stimuli. The training procedure consisted of a mix of demonstration trials, in which the visual reinforcer automatically appeared 1 sec after the reversal of the target tones regardless of the infant's behavior, and training trials in which the reinforcers were only presented when the infant turned within 2 sec of the change. To pass training and proceed to the experimental trials, infant and adult listeners were required to respond on three of the previous four trials. No control trials were used during training.

Once the training criterion was met, the blocks of experimental trials began without interruption. Each block consisted of 20 trials: 10 change trials in which the order of the target tones was reversed, and 10 control trials in which target tones were identical to the background series. The control trials were used to minimize possible effects of differences in response bias across listeners (i.e., some listeners responding more often than others). The experimenter called for a trial on a hidden keypad when the infant was centered, or looking forward. The type of trial (i.e., change or control) was randomly determined by the computer program, and the experimenter and the infant's parent were unaware of the type of trial because they listened to masking music over headphones. When the infant turned toward the display, the experimenter recorded this on the keypad. If the head turn occurred within 2 sec of a change trial, the computer presented the visual reinforcer. Head turns on control trials (i.e., false alarms) were not reinforced.

Because adults performed all three conditions (with the order of the flanker and captor conditions counterbalanced across subjects), differences between performance on flanker and captor trials could be examined as a within-subjects factor. However, because infants were not able to complete both conditions, comparison of performance on flanker and captor trials was examined as a between-subjects factor for infants.

RESULTS

For each subject, for each block/condition, the data consisted of the number of responses on the 10 change trials and 10 control trials. These numbers are analogous to hit and false-alarm rates. Percent correct scores[2] were calculated by adding the number of hits and correct rejections (i.e., 10 minus the number of false alarms), and dividing by the total number of trials (i.e., 20). Because each infant only completed two blocks of trials, whereas adults completed all three, statistical tests were performed on each age group separately. A preliminary analysis using one-sample *t* tests revealed that performance was significantly above chance (50%) for both infant and adult listeners in all experimental conditions (all ps $< .001$).

As shown in Figure 2, adult listeners performed at near ceiling levels in the target-alone condition, but their performance dropped when flanker tones were introduced. As expected on the basis of Bregman and Rudnicky's (1975) study, performance improved when captor tones were introduced. These effects were confirmed by a two-way analysis of variance (ANOVA) with a within-subjects factor of stimulus condition (target alone, flanker, and captor) and a between-subjects factor of block order. Performance varied significantly across condition, $F(2, 32) = 12.96$, $p < .001$, $\eta^2 = .447$. Pairwise comparisons revealed that performance in the flanker condition was significantly lower than both the target-alone ($p < .005$) and the captor ($p < .005$) conditions, which were not significantly different from each other. No significant effect of block order, $F(1, 16) = .160$, *ns*, or condition $\times$ block order interaction was found, $F(2, 32) = .149$, *ns*.

Percent correct scores for the infants, also shown in Figure 2, were submitted to a two-way ANOVA, with the between-subjects factor of group (captor or flanker) and the within-subjects factor of block. As both groups were treated identically in the first block, in which they heard target tones alone, but differed in the second block, in which either the captor or the flanker conditions were presented, any difference between the conditions should

[2]These results were also analyzed in terms of d' using modified hit and false-alarm rates and a comparable pattern of results was found.

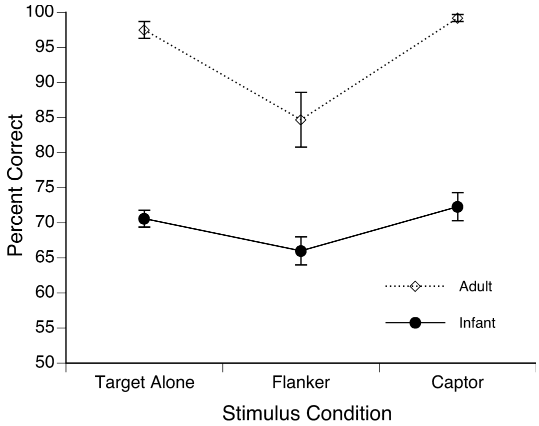

Figure 2 Overall performance in the three stimulus conditions for adult and infant listeners. Chance performance is equal to 50%. Standard error bars are shown.

be manifest as a block × group interaction. No significant main effects were found for block, $F(1, 38) = 1.31$, *ns*, $\eta^2 = .033$, or for group, $F(1, 38) = 1.01$, *ns*, $\eta^2 = .026$. As expected, a significant block × group effect was found, $F(1, 38) = 9.28$, $p < .01$, $\eta^2 = .196$. An examination of the within-subjects simple effects in this interaction revealed a significant drop in performance in the flanker group from the first (target alone) to the second (flanker) block, $F(1, 38) = 8.77$, $p < .005$. In contrast, the captor group's performance did not differ significantly between the first (target alone) and second (captor) block, $F(1, 38) = 1.81$, *ns*. The corresponding examination of between-subjects simple effects found no significant difference between the groups in the first (target alone) block, $F(1, 38) = .53$, *ns*, but significantly higher performance for the captor group than the flanker group in the second block, $F(1, 38) = 5.03$, $p < .01$.

DISCUSSION

A few previous studies have suggested that at least some aspects of auditory scene analysis are operative in infants (Demany, 1982; Fassbender, 1993; McAdams & Bertoncini, 1997). The present results extend these findings by showing that infants cannot only organize successive sounds into two streams, but that they can ignore information in one stream while attending to information in another. Specifically, infants show the same flanker-related masking effect and captor-related streaming effect as adults. Both infants and adults successfully perceived differences in the temporal order of the target tones when presented alone, experienced difficulty when the flanker tones were added, but returned to, and even slightly exceeded, target-alone levels when captor tones were introduced that organized the flanker tones into a different stream than the target tones, consistent with Bregman and Rudnicky (1975). This result suggests that infants show similar benefits of perceptual organization of the nonattended stream as adults.

In contrast to these findings, infants often perform like broadband listeners, being unable to benefit from prior expectations about the frequency of target tones (Bargones & Werner, 1994), and showing large distraction effects for sounds at remote frequencies (Werner & Bargones, 1991). In the context of this work, the decrement in infants' performance with the addition of flanker tones is consistent with these findings, although this decrement was not particularly larger for infants than for adults. However, the lack of decrement in infants' performance with captor tones is surprising in the context of this earlier work because the greater number of tones competing for attention would be predicted to result in an even greater performance decrement. Thus, it appears that not only can infants readily form separate auditory streams, but that a segregated stream does not necessarily have a distracting effect on infants' processing of information in a target stream.

There are several differences between this study and previous studies that might contribute to infants' ability to ignore irrelevant information. One is that previous studies involved the detection of sounds presented near threshold levels, whereas in the present case the task was the discrimination of the temporal order of clearly audible sounds. Near threshold, infants might be more affected by distracting sounds than when detecting features of sounds above threshold. Other factors that might affect infants' ability to ignore sounds are variability in the irrelevant information, whether distracter sounds are discrete or continuous, whether the sounds are tonal or noise, and whether distracter and target tones are played simultaneously or successively. In this study, nonsimultaneous tonal distracters were used and their frequency and timing in the captor condition was perfectly regular and predictable. For detection near threshold, Leibold and Werner (2006) found that thresholds for both infant and adult listeners are lower when continuously repeating tonal maskers are presented at fixed versus random frequencies. This is consistent with other work illustrating the role of masker variability in informational masking (Allen & Wightman, 1995; Kidd, Mason, Deliwala, Woods, & Colburn, 1994; Neff & Green, 1987). However, infants are much more affected than adults by the simple presence of continuous noise that overlaps the target tones, a condition of very low masker variability. Specifically, Werner and Bargones (1991) found that infants show excessive masking near threshold, when the masking noise is presented continuously throughout a test session.

We suggest that the idea that perceptual segregation might be necessary in order for infants to easily ignore masking stimuli is worth exploring. For infants, it is possible that stream segregation is difficult for sounds near threshold, that segregation of sequential sounds might be easier than simultaneous sounds, and that continuous noise lacks sufficient patterned structure to easily promote its perceptual segregation from the target tones. However, under conditions that promote stream segregation, infants appear to be able to attend to the information in one stream and ignore the information in another stream. In the present case, the use of predicable, nonsimultaneous, above-threshold tonal stimuli likely provided good conditions for stream segregation.

Exploring the development of auditory scene analysis can also shed light on the nature of this process. The Gestaltists viewed it to be unlearned or innate and largely bottom up. Finding that these principles are operational in young infants provides some suggestive evidence for this position. On the other hand, Bregman (1990) has argued that there are both innate and learned aspects to auditory scene analysis, and innate influences must collaborate with, or "bootstrap," a learning process. Over time, having acquired some knowledge of the stimulus structure through learning processes, listeners can use top-down, schema-based processes to detect, discriminate and identify particular targets that provide clues to auditory object segregation in the incoming auditory information. For example, infants can use various acoustical or statistical cues in the continuous speech stimulus to segment and isolate individual words for the purposes of word learning, and then eventually use that lexical knowledge to facilitate later speech segmentation in a top-down manner (Saffran, 2001). Infants are able to attend to a single female voice among competing male voices, although these abilities remain immature (Newman & Jusczyk, 1996). Furthermore, their increased sensitivity to their own name (Newman, 2005) and their own mother's voice (Barker & Newman, 2004) suggests that even infants can use their knowledge to analyze the auditory scene.

Infants need to learn about the objects in their environment, and one aspect of this is the sounds those objects make. Thus, auditory scene analysis is crucial for differentiating people, learning language, perceiving music, and interpreting environmental sounds. The results presented here illustrate that auditory scene analysis also plays an important role in cognitive development, in that it affects infants' ability to selectively attend to relevant target sounds. A clearer picture of the interaction between the operation of basic principles of perceptual organization and attentional processes will lead to a better understanding of infants' sensitivity to, and capacity to learn about, their environment across domains.

ACKNOWLEDGMENTS

We thank Colleen Shane, Sandy Estee, Katie Gorga, and Abbey Correll for assistance with subject recruitment and data collection. This research was supported in part by the NIH (P30 DC4662).

REFERENCES

Alain, C., & Izenberg, A. (2003). Effects of attentional load on auditory scene analysis. *Journal of Cognitive Neuroscience*, *15*, 1063–1073.

Allen, P., & Wightman, F. (1995). Effects of signal and masker uncertainty on children's detection. *Journal of Speech and Hearing Research*, *38*, 503–511.

Bargones, J., & Werner, L. (1994). Adults listen selectively; infants do not. *Psychological Science*, *5*, 170–174.

Barker, B., & Newman, R. S. (2004). Listen to your mother! The role of talker familiarity in infant streaming. *Cognition*, *94*, 45–53.

Bregman, A. S. (1990). *Auditory scene analysis*. Cambridge, MA: MIT Press.

Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, *89*, 244–249.

Bregman, A. S., & Rudnicky, A. I. (1975). Auditory segregation: Stream or streams. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 263–267.

Carlyon, R. P. (2004). How the brain separates sounds. *Trends in Cognitive Sciences*, *8*, 465–471.

Carlyon, R. P., Cusack, R., Foxton, J. M., & Robertson, I. H. (2001). Effects of attention and unilateral neglect on auditory stream segregation. *Journal of Experimental Psychology. Human Perception and Performance*, *27*, 115–127.

Demany, L. (1982). Auditory stream segregation in infancy. *Infant Behavior & Development*, *5*, 261–276.

Durlach, N. I., Mason, C. R., Kidd, G., Arbogast, T. L., Colburn, H. S., & Shinn-Cunningham, B. G. (2003). Note on informational masking. *Journal of the Acoustical Society of America*, *113*, 2984–2987.

Egan, J., Greenberg, G., & Schulman, A. (1961). Interval of time uncertainty in auditory detection. *Journal of the Acoustical Society of America*, *33*, 771–778.

Fassbender, C. (1993). *Auditory grouping and segregation processes in infancy*. Norderstedt: Kaste Verlag.

Greenberg, G. Z., & Larkin, W. D. (1968). Frequency-response characteristic of auditory observers detecting signals of a single frequency in noise: The probe-signal method. *Journal of the Acoustical Society of America*, *44*, 1513–1523.

Grose, J. H., Hall, J. W., & Gibbs, C. (1993). Temporal analysis in children. *Journal of Speech and Hearing Research*, *36*, 351–356.

Hall, J. W., Grose, J., & Dev, M. (1997). Auditory development in complex tasks of comodulation masking release. *Journal of Speech, Language and Hearing Research*, *40*, 946–954.

Hall, J. W., Haggard, M. P., & Fernandes, M. A. (1984). Detection in noise by spectro-temporal pattern analysis. *Journal of the Acoustical Society of America*, *76*, 50–56.

Kidd, G., Mason, C. R., Deliwala, P. S., Woods, W. S., & Colburn, H. S. (1994). Reducing informational masking by sound segregation. *Journal of the Acoustical Society of America*, *95*, 3475–3480.

Leibold, L., & Werner, L. A. (2006). Effect of masker-frequency variability on the detection performance of infants and adults. *Journal of the Acoustical Society of America*, *119*, 3960–3970.
McAdams, S., & Bertoncini, J. (1997). Organization and discrimination of repeating sound sequences by newborn infants. *Journal of the Acoustical Society of America*, *102*, 2945–2953.
Miller, G., & Heise, G. (1950). The trill threshold. *Journal of the Acoustical Society of America*, *22*, 637–638.
Neff, D. L., & Green, D. M. (1987). Masking produced by spectral uncertainty with multi-component maskers. *Perception & Psychophysics*, *41*, 409–415.
Newman, R. S. (2005). The cocktail party effect in infants revisited: Listening to one's name in noise. *Developmental Psychology*, *41*, 352–362.
Newman, R. S., & Jusczyk, P. (1996). The cocktail party effect in infants. *Perception & Psychophysics*, *58*, 1145–1156.
Olsho, L. W., Koch, E., Halpin, C., & Carter, E. (1987). An observer-based psychoacoustic procedure for use with young infants. *Developmental Psychology*, *23*, 627–640.
Polka, L., Rvachew, S., & Molnar, M. (2008). Speech perception by 6- to 8-month-olds in the presence of distracting sounds. *Infancy*, *13*, 421–439.
Pollack, I. (1975). Auditory informational masking. *Journal of the Acoustical Society of America*, *57*, S5.
Saffran, J. (2001). Words in a sea of sounds: The output of infant statistical learning. *Cognition*, *81*, 149–169.
Sinnott, J., & Aslin, R. (1985). Frequency and intensity discrimination in human infants and adults. *Journal of the Acoustical Society of America*, *78*, 1986–1992.
Snyder, J. S., Alain, C., & Picton, T. W. (2006). Effects of attention on neuroelectric correlates of auditory stream segregation. *Journal of Cognitive Neuroscience*, *18*, 1–13.
Sussman, E., Horváth, J., Winkler, I., & Orr, M. (2007). The role of attention in the formation of auditory streams. *Perception & Psychophysics*, *69*, 136–152.
Sussman, E., Ritter, W., & Vaughan, H. G. (1999). An investigation of the auditory streaming effect using event-related brain potentials. *Psychophysiology*, *36*(1), 22–34.
Sussman, E., Wong, R., Horvath, J., Winkler, I., & Wang, W. (2007). The development of the perceptual organization of sound by frequency separation in 5-11-year-old children. *Hearing Research*, *225*, 117–127.
van Noorden, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences*. (Unpublished doctoral dissertation). Eindhoven University of Technology, Eindhoven, the Netherlands.
Watson, C. S., & Kelly, W. J. (1981). The role of stimulus uncertainty in the discrimination of auditory patterns. In D. J. Getty & J. H. Howard (Eds.), *Auditory and visual pattern recognition* (pp. 37–59). Hillsdale, NJ: Erlbaum.
Werker, J., Polka, L., & Pegg, J. (1997). The conditioned head turn procedure as a method for testing infant speech perception. *Early Development and Parenting*, *6*, 171–178.
Werner, L. A., & Bargones, J. (1991). Sources of auditory masking in infants: Distraction effects. *Perception & Psychophysics*, *50*, 405–412.
Winkler, I., Kushnerenko, E., Horváth, J., Ceponiene, R., Fellman, V., Huotilainen, M., ... Sussman, E. (2003). Newborn infants can organize the auditory world. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(20), 11812–11815.