Attention is critical for spatial auditory object formation

Benjamin H. Zobel¹ · Richard L. Freyman² · Lisa D. Sanders¹

Published online: 5 May 2015
© The Psychonomic Society, Inc. 2015

Abstract The precedence effect provides a novel way to examine the role of attention in auditory object formation. When presented with two identical sounds from different locations separated by a short stimulus onset asynchrony (SOA), listeners report a single auditory object at the location of the lead sound. When the SOA is above the echo threshold, listeners report hearing two auditory objects with different locations. Event-related potential (ERP) studies have shown that the number of perceived auditory objects is reflected in an object-related negativity (ORN) 100–250 ms after onset and in a posterior late positivity (LP) 300–500 ms after onset. In the present study, we tested whether these ERP effects are modulated by attention by presenting lead/lag click pairs at and around listeners' echo thresholds, while in separate blocks the listeners (1) attended to the sounds and reported whether the lag sound was a separate source, and (2) performed a two-back visual task. When attention was directed away from the sounds, neither the ORN nor the LP observed in the attend condition was evident. Instead, unattended click pairs above the echo threshold elicited an anterior positivity 250–450 ms after onset. However, an effect resembling an ORN was found in comparing the ERPs elicited by unattended click pairs with SOAs below the attended echo threshold, indicating that the echo threshold may have been lowered when attention was directed away from the sounds. These results suggest that attention modulates early perceptual processes that are critical for auditory object formation.

Keywords Selective attention · Audition · Evoked potentials

Attention has been shown to be important for visual object formation (Treisman & Gelade, 1980; Treisman & Schmidt, 1982). Much less is known about the role of attention in auditory object formation, in part because defining auditory objects and measuring their perception presents unique challenges. However, the *precedence effect* is a phenomenon in which millisecond differences in the timing of sounds result in categorical differences in the number of perceived auditory objects (Wallach, Newman, & Rosenzweig, 1949). Specifically, when identical sounds are presented from different locations with a stimulus onset asynchrony (SOA) above the echo threshold, they are perceived as two auditory objects; when the same sounds are presented with a slightly shorter SOA, below the echo threshold, they are perceived as a single auditory object localized at the position of the lead sound. The object-related negativity (ORN) provides an ERP index of early auditory object perception that can be measured in the complete absence of attention to sounds. Previous ORN studies (Alain, Arnott, & Picton, 2001; Alain & Izenberg, 2003; Alain, Schuler, & McDonald, 2002; Dyson, Alain, & He, 2005; Hautus & Johnson, 2005) have suggested that attention plays little to no role in forming auditory objects based on harmonic grouping or dichotic pitch. However, the precedence effect involves auditory object formation that includes higher-level processes. We hypothesized that attention does play an important role in auditory object perception in the precedence effect, such that attention would be shown, for the first time, to modulate the ORN. Such evidence would suggest that under complex conditions, attention is similarly critical for both visual and auditory object formation.

✉ Benjamin H. Zobel
bzobel@psych.umass.edu

¹ Department of Psychological and Brain Sciences, University of Massachusetts, Amherst, MA 01003, USA

² Department of Communication Disorders, University of Massachusetts, Amherst, MA 01003, USA

The precedence effect

The precedence effect can be demonstrated by positioning a listener in front of two spatially separated loudspeakers and presenting a simple simulation of a direct sound, followed by its echo. When identical sounds are presented from both loudspeakers as a lead/lag pair with an SOA on the order of

milliseconds, the sounds will be perceptually fused into a single auditory object with a location dominated by the lead sound (for a review, see Litovsky, Colburn, Yost, & Guzman, 1999). As the lead/lag SOA is increased beyond the range of the precedence effect, the listener will begin to identify the lag sound as a separate auditory object at the location of the lag loudspeaker. The SOA at which this occurs is called the listener's *echo threshold* (Blauert, 1997). Extensive psychophysical examination of the precedence effect has shown that echo thresholds vary widely across stimulus types, from approximately 5 to 10 ms for clicks (Freyman, Clifton, & Litovsky, 1991), to upward of 50 ms for music (Wallach et al., 1949) and speech (Haas, 1951). The stimulus features shown to influence the echo threshold include amplitude, envelope, duration, pitch, and lead/lag correlation (Blauert, 1997; Blodgett, Wilbanks, & Jeffress, 1956; Goverts, Houtgast, & van Beek, 2000; Miller, Litovsky, & Kluender, 2009; Saberi & Antonio, 2003; Schubert & Wernick, 1969; Seeber & Hafter, 2011; Shinn-Cunningham, Zurek, Durlach, & Clifton, 1995). Additionally, echo thresholds tend to be highly variable among listeners, and can also vary within an individual according to the location of the lead sound (Sanders, Joh, Keen, & Freyman, 2008).

Since basic stimulus features can influence the echo threshold, and many of the calculations involved in localizing direct sound energy occur subcortically, the precedence effect is likely to begin very early in auditory processing, involving interactions among peripheral filtering operations, hair cell responses, and binaural cross-correlations (Hartung & Trahiotis, 2001). Neuropsychological evidence in humans and single-cell recording studies in several animal models have suggested that processing all along the early auditory pathway, from the auditory nerve to primary auditory cortex, contributes to the precedence effect (for a review, see Litovsky et al., 1999). Key findings also point to higher-level mechanisms involved in the precedence effect. Infants, despite being born with the ability to localize single-source sounds, do not begin to orient to lead/lag stimuli until approximately four months of age (Clifton, Morrongiello, Kulig, & Dowd, 1981; Muir, Clifton, & Clarkson, 1989), and an echo threshold and source localization for lead/lag stimuli take several years to fully develop (Litovsky & Godar, 2010; Morrongiello, Kulig, & Clifton, 1984), suggesting that the precedence effect depends on the maturation of higher-level cortical systems. In adults, studies have shown that the echo threshold can be raised and lowered depending on the sequence of sound directly preceding a lead/lag stimulus, a phenomenon commonly referred to as the “buildup” and “breakdown” of echo threshold (Clifton, 1987; Clifton & Freyman, 1989; Clifton, Freyman, & Litovsky, 1994; Clifton, Freyman, & Meo, 2002; Djelani & Blauert, 2001; Freyman et al., 1991; Freyman & Keen, 2006; Grantham, 1996; Keen & Freyman, 2009). Research examining the conditions under which buildup and

breakdown occur has provided compelling evidence that this contextual modulation of the echo threshold reflects higher-level processes in which listeners construct models of acoustic space based on expectations of how sound typically behaves within reverberant environments (Clifton et al., 1994; Clifton et al., 2002; Freyman & Keen, 2006; Keen & Freyman, 2009).

Although many questions about the precedence effect remain to be answered, it is evident that the grouping of lead and lag sounds into one or two auditory objects is determined by a complex set of mechanisms that are distributed across multiple stages of processing and are sensitive to low- and high-level features of the sounds and the surrounding environment. The precedence effect not only serves to localize sound sources within reverberant scenes, but also to build rich representations of the acoustic space and the auditory objects that occupy it.

ERP indices of auditory object perception

How the auditory system identifies, segments, and groups the components of an auditory scene into representations of unified objects is not well understood. Bregman's (1990) *auditory scene analysis* posits that certain bottom-up processes automatically segregate and integrate elements on the basis of correlations among their spatial and spectrotemporal features. Harmonic grouping is an example of auditory object formation, whereby a set of tuned pure tones are perceptually fused into a single complex sound (Bregman, 1990). Alain et al. (2001) used ERP measures to explore harmonic grouping. Electroencephalography (EEG) was recorded while complex tones were presented and listeners reported whether they heard a single complex sound or a complex sound and a separate pure tone. When one of the harmonics was mistuned to the extent that listeners reported hearing it as a separate tone, the ERPs showed an increased negativity over anterior and central electrodes that extended across the N1–P2 complex 150–250 ms after stimulus onset. The researchers called this negative component the *object-related negativity*, suggesting that it indexed the perception of two auditory objects rather than one. In addition to the ORN, a widely distributed late positivity (LP) was found 350–450 ms after stimulus onset and was attributed to top-down processes involved in recognizing a sound and interpreting its meaning on the basis of prior knowledge. The LP may also reflect processes associated with a P300 ERP effect, including target detection and categorization (for a review, see Polich, 2007). Importantly, ERP measurements from a passive-listening condition, in which listeners read a book while tuned and mistuned stimuli were presented, showed that the ORN was still present, suggesting that attention was not required for harmonic object grouping. The LP, however, was absent, suggesting that attention was important for the top-down processing of the stimuli. A follow-up study by Alain et al. (2002), in which listeners

attended to a silent movie while tuned and mistuned stimuli were presented, replicated the previous results. Alain and Izenberg (2003) showed that the ORN was elicited by mistuned stimuli regardless of whether the stimuli were presented within the attended or unattended channel of a dichotic listening task. Dyson et al. (2005) also found no effect of attention on the ORNs elicited by similar stimuli using zero-back and one-back visual tasks.

In addition to harmonic grouping, the ORN has been elicited in studies of dichotic pitch, a paradigm in which an interaural time difference (ITD; on the order of microseconds) within a narrow frequency band of two identical broadband noise samples presented to the ears produces the perception of two spatially separated sounds, whereas only one sound is perceived when no ITD is present (Hautus & Johnson, 2005; Johnson, Hautus, & Clapp, 2003). Similar to the harmonic-grouping studies above, Hautus and Johnson showed that dichotic pitch perception elicited an ORN regardless of whether listeners were directing their attention to the auditory stimuli or to a silent cartoon, whereas an LP was only elicited by attended sounds. Additionally, a lateralized ORN has been reported in a study by Butcher, Govenlock, and Tata (2011), in which the spatial movement of broadband noise across an array of loudspeakers was manipulated to produce two-sound versus one-sound perception.

Since the precedence effect involves the grouping of sound into an auditory object, one would expect to see an ORN when listeners report hearing the lag sound as a separate source versus when they report hearing a single sound source at the lead location. This is exactly what Sanders et al. (2008) found. ERPs for click pairs at the echo threshold showed an increased negativity 100–250 ms after stimulus onset for trials on which listeners reported hearing the lag click, as compared to trials on which they did not. Unlike the harmonic-grouping and dichotic-pitch studies discussed above, this study compared identical stimuli at the echo threshold, ensuring that the ORN could be attributed to a shift in perceptual grouping rather than to physical differences in the stimuli. Additionally, the ORN was present for trials 1 ms above the echo threshold, in which listeners reported hearing the lag sound on a majority of trials, as compared to 1 ms below the echo threshold, in which case listeners did not hear the lag sound on the majority of trials. Although an LP was numerically noted for these comparisons when listeners heard the lag sound as a separate source, the effect did not reach statistical significance. Since release from the precedence effect occurs gradually as SOAs are increased above the echo threshold (Blauert, 1997), lack of an LP may reflect weak lag-sound recognition and low confidence in responses. Sanders, Zobel, Freyman, and Keen (2011) found similar results when they examined the contextual modulation of echo threshold. ERPs from trials on which click pairs were preceded by a sequence of identical pairs designed to produce buildup of an echo threshold (i.e.,

the lag click was not reported on a majority of trials) were compared to those elicited by click pairs preceded by a sequence of clicks designed to inhibit buildup of an echo threshold (the lag click was reported on the majority of trials). An ORN was observed in response to click pairs following the context that inhibited buildup, relative to identical click pairs following the context that did build up an echo threshold. Additionally, lag-click perception was associated with a significant LP 250–500 ms after stimulus onset, suggesting that the strong contextual manipulation of echo threshold may have produced clearer perceptual distinctions between conditions that listeners could more easily recognize and respond to with confidence.

Attention and auditory object formation

Selective attention is the preferential processing of relevant information within a scene and is an essential mechanism for navigating perceptually challenging conditions. Since the precedence effect involves complex, multilevel processes that are collectively important for the coherent organization of auditory objects in reverberant environments, it is striking to find that nothing is known about the role of attention in the precedence effect. The same challenging reverberant environments in which we depend on the precedence effect for coherency are precisely those in which we are most likely to deploy attention. The present study is the first to have used ERPs to examine the influence of attention on auditory object formation in the precedence effect, by comparing the processing of click pairs across the echo threshold under attended and unattended conditions. Of particular interest was whether attention modulates the grouping of auditory objects as indexed by the ORN. The fact that the ORN persists across attention manipulations in harmonic-grouping and dichotic-pitch studies suggests that object grouping involves only automatic, bottom-up mechanisms under some conditions. As Alain and colleagues have argued, this interpretation is consistent with Bregman's (1990) view, which posits that preattentive processes mediate the grouping of basic features within an auditory scene (Alain et al., 2001; Alain & Izenberg, 2003; Alain et al., 2002).

Even if harmonic grouping and dichotic pitch are preattentive, there is little reason to assume that the processes underlying these phenomena would generalize to object grouping in the precedence effect. Given that frequency is a basic feature by which the representation of sound is organized within the auditory system, and that ITDs are processed to a large extent subcortically, one would expect object grouping based on multiple pitches or ITDs to occur with relative ease and a minimum of attention. Object grouping within the precedence effect, on the other hand, is likely to be more complex, requiring the integration of features across multiple dimensions. One might expect more computationally demanding operations to be facilitated by

attention. Hall, Pastore, Acker, and Huang (2000) and Thompson, Hall, and Pressing (2001) provided evidence that the separable features within an auditory scene, including pitch, duration, and location, can be incorrectly grouped in a manner consistent with the illusory conjunctions predicted by feature-integration theory (Treisman & Gelade, 1980; Treisman & Schmidt, 1982). These findings suggest that attention would be required for the complex object grouping across feature dimensions that we find in the precedence effect.

In the present study, EEG was recorded while participants completed blocks of trials in which their attention was either directed to lead/lag click pairs with SOAs spanning the echo threshold (attend condition) or diverted to a difficult two-back visual task (unattend condition). In both conditions, a stream of letters that alternated in case every two letters was presented on a computer screen in front of the participant. In the attend condition, the participant was instructed to use the letter stream only as a fixation point, to attend instead to click pairs presented from loudspeakers located to the left and right, and to press a button after each presentation, indicating whether or not the lag sound was heard. In the unattend condition, participants were instructed to attend to the letters in the visual stream while the click pairs were presented, and to press a button as soon as they saw a letter appear that alphabetically matched the letter that had been presented two spaces back in the stream. The stimuli presented in both conditions were identical, except for the fact that the attend condition did not contain two-back visual targets. To further motivate our participants, the attend and unattend tasks were presented as a game, with points awarded for responding to every click pair in the attend block, and points awarded for quickly identifying visual targets in the unattend blocks. The present design allowed us to address several important questions: (1) Did the attention manipulation work? (2) Did we replicate ERP indices of the precedence effect (ORN and LP) in the attend condition? (3) Did attention modulate these effects across conditions? and (4) Did attention shift the echo threshold?

Method

Participants

Twenty-three right-handed participants (14 male, nine female) 20–32 years of age ($M = 24$ years, $SD = 3.62$) contributed data to the analysis. Six additional participants completed the initial screening session, but a clear echo threshold could not be estimated from their behavioral responses and they were not asked to return for the experimental session. The data from one participant who completed the experiment were excluded due to excessive low-frequency drift in the EEG recording. All participants reported normal or corrected-to-normal vision, having no known neurological problems, and taking no

psychoactive medication. When participants first arrived, they underwent a preliminary hearing screening with a Beltone audiometer to ensure normal hearing thresholds (≤ 20 dB HL for 1-, 2-, and 4-kHz tones, and ≤ 30 dB HL for 8-kHz tones). All participants provided informed consent and were compensated at a rate of \$10/h.

Stimuli

The auditory stimuli were pairs of “click” sounds composed of identical 181- μ s positive rectangular four-sample pulses (16-bit, 22.050 kHz). Twenty-two right-lead click pairs (SOAs of 1, 1.33, 1.5, 1.75, and 2–19 ms in 1-ms steps) were assembled using Pro Tools audio software by placing a single click in the right channel, followed by an identical click in the left channel of a stereo WAV file. The auditory stimuli were presented at 70 dBA over a matched pair of M-Audio StudioPro3 loudspeakers placed 1.4 m from the participants, 55 deg to the left and right of midline. The visual stimuli consisted of single presentations of white letters (a, b, d, e, f, g, h, j, l, m, n, q, r, and t) on a black background in the center of a computer screen placed 1.5 m directly in front of the participants. One uppercase and one lowercase version of each letter were created, each of which subtended less than 1 deg of visual angle horizontally and vertically. All stimuli were presented from a PC using E-Prime software. Participants were seated in a comfortable chair in the center of an acoustically dampened 2.5 $\times$ 3.5 m room.

Procedure

The experiment consisted of two separately scheduled sessions: a screening session, in which behavioral responses to click pairs across a range of SOAs were collected, to estimate echo thresholds, followed by an experimental session in which behavioral and EEG data were collected simultaneously.

Echo threshold screening session The results from the screening session provided an estimate of each participant’s echo threshold, which was used to determine the seven threshold-centered SOAs for the experimental session. Screening trials consisted of 1,500 ms of silence followed by the presentation of a single click pair. A white fixation cross against a black background appeared in the center of the computer screen for the duration of each trial, followed by a response prompt 600 ms after click-pair offset. Participants were told that on every trial, they would hear a sound from the right side and that their task was to push a button on a button box, when prompted, indicating whether they also heard a sound from the left side. They were instructed to remain centered in the chair and to fixate on the white cross while listening. Before beginning the screening trials, participants practiced

responding to trials that included SOAs of 1 and 25 ms, to provide them with clear examples of hearing and not hearing the lag click as a separate source. Participants were told that the actual screening trials might not always be as clear as these examples, and that they should rely on their best judgment.

The initial screening included 196 trials. Fourteen click pairs with SOAs ranging from 1 to 14 ms in 1-ms steps were repeated 14 times in random order. Odd- and even-numbered SOA trials were presented in separate blocks. The results allowed the experimenter to estimate the echo threshold as the SOA at which the lag sound came closest to being reported on 50 % of trials; reports of the lag sound as a separate source were also required to decrease across lower SOAs and increase across higher SOAs. The echo threshold estimated for each participant was used to determine the seven threshold-centered SOAs to be used in the subsequent screening blocks: threshold (T_0), ± 1 ms (T_{+1} and T_{-1}), ± 2 ms (T_{+2} and T_{-2}), and ± 5 ms (T_{+5} and T_{-5}). Some participants had estimated thresholds that were too low to subtract 2 or 5 ms and still maintain a lowest SOA at which the sound from the right led by at least 1 ms. For thresholds ≤ 5 ms, T_{-5} was set to 1 ms; for thresholds of 3 ms, T_{-2} was set to 1.5 ms; for thresholds of 2 ms, T_{-1} was set to 1.67 ms and T_{-2} to 1.33 ms. Table 1 shows the seven SOAs selected for each estimated echo threshold.

Next, participants received two screening blocks, consisting of five trials at each of the seven selected SOAs presented in random order. Additional blocks were presented in order to check for response consistency, if needed. If the estimated threshold changed with cumulative responses, the experimenter selected seven new SOAs accordingly. Participants with echo thresholds < 2 ms, > 14 ms, or that could not be reliably estimated were not asked to return for the experimental session.

Table 1 The seven threshold-centered stimulus onset asynchronies determined by each estimated echo threshold (in milliseconds)

Estimated Echo Threshold	T_{-5}	T_{-2}	T_{-1}	T_0	T_{+1}	T_{+2}	T_{+5}
2	1	1.33	1.67	2	3	4	7
3	1	1.5	2	3	4	5	8
4	1	2	3	4	5	6	9
5	1	3	4	5	6	7	10
6	1	4	5	6	7	8	11
7	2	5	6	7	8	9	12
8	3	6	7	8	9	10	13
9	4	7	8	9	10	11	14
10	5	8	9	10	11	12	15
11	6	9	10	11	12	13	16
12	7	10	11	12	13	14	17
13	8	11	12	13	14	15	18
14	9	12	13	14	15	16	19

Experimental session The experimental session consisted of two conditions: an *attend* condition, in which participants listened for click pairs selected from the seven threshold-centered SOAs, and an *unattend* condition, in which participants engaged in a visual two-back task while being presented with the click pairs. Both conditions included approximately 2.5-min blocks of simultaneously presented auditory and visual streams. The auditory stream was 35 click pairs, with five repetitions of the seven threshold-centered SOAs, presented in random order. The auditory stream began 900–3,400 ms (in 1-ms steps) after the start of the visual stream, with an interonset interval (IOI) between click pairs that ranged from 2,000 to 7,000 ms in 1-ms steps. The visual stream consisted of single letters presented for 700 ms each, with 900-ms IOIs. The letter case changed after every two letters, such that two uppercase letters were always followed by two lowercase letters and the cases of letters two apart in the stream never matched. In attend blocks, letters were presented in a random order, with the exception that the same letter of the alphabet could not appear two spaces apart. In unattend blocks, the visual stream included 35 two-back targets, with intertarget onset intervals of 900–8,100 ms in 900-ms steps. Targets were defined as the presentation of a letter that matched the letter of the alphabet two spaces back in the stream; targets and their matching letters always differed in case. To keep participants motivated and entertained, the experimental session was presented as a game against the computer. For attend blocks, participants had to respond to every click pair in order to gain five points and avoid having ten points removed from their score. For unattend blocks, participants gained one point for every target that elicited a button press within 3,000 ms; every miss or false alarm gave the computer one point. The correct number of responses in every attend and unattend block was 35. Participants were told at the start of the session that they would have the opportunity to post their final scores to an anonymous leader board displaying their performance rankings among all others.

To begin the experimental session, participants received instructions on the unattend task. They were then presented with a visual-only stream and practiced pressing a button as quickly as possible to targets until they had correctly identified five targets. After receiving instructions on how the unattend task would be scored, they practiced responding to a visual stream containing six targets and viewed the resulting score. Finally, they were presented with an auditory and visual stream to practice responding to four visual targets under the actual conditions of an unattend block. They were reminded that during an unattend block, their goal was to focus on the visual letters and to respond as quickly and accurately as possible.

Next, participants were reacquainted with the auditory stimuli by reviewing the instructions and practice trials that they had completed at the beginning of the previous screening

session. The structure and scoring of the attend blocks were then explained. Participants were told that during an attend block, they should listen for each click and press a button indicating whether they had heard a click from the left side in addition to the one that they would always hear from the right side. They were told that the visual stream would not contain any targets and should be used as a fixation point only. They then practiced responding to five click pairs under the actual conditions of an attend block. Finally, participants received two complete attend blocks as practice with the seven threshold-centered SOAs determined by the screening session. The experimenter examined the responses during the attend practice before determining the seven threshold-centered SOAs to be used in the experimental blocks.

After completing the practice, participants received 32 experimental blocks in random order, evenly divided between attend and unattend conditions, resulting in 560 visual targets across the 16 unattend blocks and 560 click pairs (80 at each of the seven threshold-centered SOAs) across both the 16 attend and 16 unattend blocks. Before each block, the block type and reminder instructions were presented on the computer screen and reinforced by the experimenter. After each block, participants viewed their block and cumulative scores. After all experimental blocks had been completed, participants were given the option to post and view their ranked scores on the leader board.

Behavioral analysis

Participants' performance on the visual two-back task was assessed by comparing the probability of a response being made within a 200- to 1,200-ms time window following the onset of a visual target (hit rate) to the probability of a response being made within any other 1,000-ms time window of the visual stream (false alarm rate). Responses on the auditory task were used to define the conditions for ERP analysis. The proportion of attend trials on which the lag click was reported as a separate sound source was calculated at each of the seven SOAs presented to each participant. A logistic function, free to vary by midpoint and slope, was then fit to each participant's data. The precise echo threshold (as opposed to the whole-number SOAs used for ERP analysis) was defined at the midpoint of the logistic function, predicting the SOA at which the lag click would be heard on exactly 50 % of trials.

EEG recording and analysis

Electrical Geodesics, Inc. (Eugene, OR) hardware and software (Net Station) was used for EEG acquisition and analysis. Vertex-referenced EEG with a 250-Hz sampling rate and a 0.01- to 100-Hz bandpass filter was recorded continuously throughout each experimental block from a 128-electrode HydroCel Geodesic net. A 60-Hz notch filter was applied offline. The EEG time-

locked to auditory stimuli was segmented into 600-ms epochs, beginning 100 ms before stimulus onset. Epochs with changes in voltage exceeding individually set thresholds were excluded from the analysis. ERPs elicited by the onset of the auditory stimuli were created by averaging together artifact-free epochs within the specific SOA categories. The ERPs were re-referenced to the average of the mastoid channels, and the data in the 100-ms prestimulus interval were used as a baseline.

The ERP analyses were guided by four objectives: (1) to confirm that the ERPs were modulated by attention, (2) to replicate the ERP indices of the precedence effect in the attend condition, (3) to examine whether the ERP effects of SOA differed in the attend and unattend conditions, and, if so, (4) to explore whether these differences were consistent with attentional modulation of the echo threshold. The effectiveness of the attention modulation was measured by comparing ERPs to the same sounds in the attend and unattend conditions. To examine the precedence effect in the attend condition, all trials at each of the seven SOAs were averaged together regardless of the behavioral response, rather than including only trials on which two sounds were reported for longer SOAs and on which one sound was reported for shorter SOAs, as had been done in previous studies. This approach was necessary because it was not possible to collect behavioral responses to the auditory stimuli while subjects fully ignored the sounds in the unattend condition, and identical ERP processing in the attend and unattend conditions facilitated comparisons. Central to the analysis of the precedence effect was the comparison of ERPs from trials with SOAs 1 ms above (T_{+1}) a participant's echo threshold (such that the lag sound was typically reported to be a separate source in the attend condition) and from trials with SOAs 1 ms below (T_{-1}) that participant's echo threshold (such that the lag sound was not typically reported to be a separate source in the attend condition). To address the possibility that the echo thresholds might differ for attended and unattended sounds, pairs of SOAs that were both above or both below echo threshold were compared for the unattend condition.

Since echo thresholds vary widely across individuals, and since ERPs were averaged across all trials at each SOA, it was important to identify two SOAs for each participant that differed by 2 ms (or by 1.33 ms if the echo threshold was 2 ms, $N = 1$), such that one SOA was above the echo threshold (T_{+1}) and the other was below the echo threshold (T_{-1}). On attend trials, the participant had to report that the lag click was a separate source on more than 50 % of trials above the echo threshold and on fewer than 50 % of trials below the echo threshold. When more than one pair of SOAs that differed by 2 ms met these criteria for a participant, the SOAs were selected to have the largest difference in the proportions of trials on which the lag click was reported.

To best capture the predicted ERP effects and their distributions, the data from 120 electrodes were included in the

analysis. These electrodes were divided into 15 groups of eight electrodes, designated by their scalp locations within a 3 [left (L), medial (M), right (R)] $\times$ 5 [anterior (A), anterior-central (AC), central (C), central-posterior (CP), posterior (P)] grid, as is shown in Fig. 1. The mean amplitude of each participant's ERPs was measured at two time windows, to assess early and late effects: 85–125 and 250–450 ms after click-pair onset. To assess the effectiveness of the attention manipulation, the mean amplitudes collapsed across T_{+1} and T_{-1} were analyzed in a 2 (Attention: attend, unattend) $\times$ 3 (Left/Right Electrode Position: L, M, R) $\times$ 5 (Anterior/Posterior Electrode Position: A, AC, C, CP, P) repeated measures analysis of variance (ANOVA). To assess the precedence effect within each attention condition, the mean amplitudes of ERPs elicited by T_{+1} and T_{-1} click pairs were analyzed with a 2 (SOA: T_{+1} , T_{-1}) $\times$ 3 (L, M, R) $\times$ 5 (A, AC, C, CP, P) repeated measures ANOVA. Follow-up ANOVAs on the data collected at subsets of electrodes were typically motivated by significant ($p < .05$) Condition $\times$ Electrode Position interactions. Greenhouse–Geisser corrections were applied to all p values; the uncorrected degrees of freedom are reported.

Results

Behavioral results

Performance on the visual task in the unattend condition was characterized by a high hit rate ($M = .85$, $SD = .08$) and a low false alarm rate ($M = .03$, $SD = .02$). A total of 18 participants responded to 100 % of the 560 trials presented in the attend condition, and no participant failed to respond on more than three trials.

A logistic function was successfully fit to each participant's response data (mean pseudo- $R^2 = .89$, $SD = .07$). As is shown in Fig. 2, all of the logistic functions were characterized by

Fig. 1 Approximate locations of the 128 electrodes used for recording electroencephalography. The data from 120 electrodes were included in the analysis (black circles) as 15 regions (dashed lines). Data were averaged across the eight electrodes within each region.

Fig. 2 Logistic functions fit to each participant's behavioral responses to click pairs in the attend condition. The dashed line indicates the stimulus onset asynchrony (SOA) at which each participant was predicted to report hearing the lag click as a separate source on 50 % of trials (i.e., the echo threshold).

positive slopes ($M = .36$, $SD = .15$), indicating more reports of the lag sound as a separate source with longer SOAs. Echo thresholds were defined as the midpoints of the logistic functions ($M = 8.36$ ms, $SD = 3.23$ ms). Both the average echo threshold and large variability across participants (range = 2.37–14.34 ms) were typical for these stimuli and task.

For eight participants, the T_{+1} and T_{-1} SOAs selected for ERP analysis were successfully predicted by responses on the screening trials. For ten participants, the behavioral responses provided during collection of the ERP data indicated that SOAs 1 ms longer than those selected on the basis of screening fit the SOA categories better; for the remaining five participants, SOAs that were 1 ms shorter than predicted were selected. In the resulting T_{+1} condition, listeners reported hearing the lag sound as a separate source on 81.66 % of trials ($SD = 13.55$ %); in the T_{-1} condition, these responses fell to 17.52 % ($SD = 11.23$ %). Figure 3 shows the percentages of trials on which the lag click was reported to be heard across participants at each of the seven SOAs.

Fig. 3 Percentages of trials on which participants reported hearing the lag click as a separate source in the attend condition. The SOA categories were defined for each participant on the basis of behavioral data.

ERP results

The grand-average ERPs showed that click pairs at all SOAs in both attention conditions elicited the positive–negative–positive waveforms that are typical in response to auditory onsets. The first positive-going peak (P1) occurred at around 55 ms. The first negative-going peak (N1) occurred at around 95 ms. The second positive-going peak (P2) occurred at around 180 ms.

ERP indices of attention Figure 4 shows the comparison of the ERPs elicited by attended and unattended click pairs. As expected, attended sounds elicited a larger N1 and a later positivity. Across the entire scalp, we observed interactions between the Attention and Electrode Position factors on the mean amplitudes 85–125 ms after click-pair onset [Attention $\times$ Anterior/Posterior: $F(4, 88) = 16.08$, $p < .001$; Attention $\times$ Left/Right $\times$ Anterior/Posterior: $F(8, 176) = 6.34$, $p < .001$]. At central electrodes (AC, C, CP), sounds elicited a larger N1 in the attend condition than in the unattend condition [$F(1, 22) = 8.27$, $p = .01$, $\eta_p^2 = .27$]. This effect was largest over medial regions [Attention $\times$ Left/Right: $F(2, 44) = 5.18$, $p = .01$]. Attending to the sounds also resulted in a larger positivity 250–450 ms after click-pair onset [$F(1, 22) = 5.77$, $p = .03$, $\eta_p^2 = .21$]. Although this effect was broadly distributed, the differences in mean amplitude were numerically largest over central and posterior electrodes.

ERP indices of the precedence effect Figure 5 shows the comparison of the ERPs elicited by T_{+1} and T_{-1} click pairs in the attend condition. Consistent with previous research, when listeners attended to the sounds, click pairs above the

echo threshold (T_{+1}) elicited a larger negativity 85–125 ms over anterior and central regions, and a later posterior positivity beginning by 250 ms, as compared to click pairs below the echo threshold (T_{-1}). The early effect was evidenced by a main effect of SOA across the scalp [$F(1, 22) = 9.07$, $p = .01$, $\eta_p^2 = .29$] that was numerically larger at anterior and central electrodes. Although the positivity did not result in a main effect of SOA or interactions with the Electrode Position factors on mean amplitudes 250–450 ms after sound onset ($ps \geq .16$), previous studies motivated analysis of the data collected over posterior regions (CP, P). At this subset of electrodes, sounds above the echo threshold elicited a larger positivity than did sounds below the echo threshold [$F(1, 22) = 5.1$, $p = .03$, $\eta_p^2 = .19$].

Unattended sounds Figure 6 shows the comparison of the ERPs elicited by T_{+1} and T_{-1} click pairs in the unattend condition. The early negativity and later posterior positivity observed for the T_{+1} condition when listeners attended to the sounds were not evident when listeners directed their attention to the visual stimuli. Instead, a visual comparison of the T_{+1} and T_{-1} SOAs defined by the behavioral data in the attend condition revealed a larger anterior positivity beginning by 250 ms for the above-threshold sounds. The analyses that showed effects in the attend condition provided no evidence for differences in the responses to the two SOAs in the unattend condition 85–125 ms ($ps \geq .55$) or over posterior regions 250–450 ms after onset ($ps > .20$). However, we saw some indication of a difference at 250–450 ms across the entire scalp [$F(1, 22) = 4.01$, $p = .06$]. Over left and medial anterior regions where this effect appeared to be the largest (L, M and A, AC, C), sounds with SOAs above the echo threshold

Fig. 4 Event-related potentials (ERPs) elicited by click pairs across T_{+1} and T_{-1} in the attend and unattend conditions. The data from all 15 regions are shown in the same positions depicted on the scalp map. The ERPs from two locations are also shown at a larger scale. Shading indicates the time windows in which we observed differences between the conditions.

Fig. 5 ERPs elicited by T_{+1} and T_{-1} click pairs in the attend condition. The data from all 15 regions are shown in the same positions depicted on the scalp map. The ERPs from two locations are also shown at a larger scale. Shading indicates the time windows in which we observed differences between the conditions.

elicited a larger positivity [$F(1, 22) = 4.79, p = .04, \eta_p^2 = .18$]. There was no evidence of a similar effect over the same region in the attend condition ($p > .80$).

To determine whether the effects of SOA were modulated by attention, data from the attend and unattend conditions were included in the same analysis. The interaction of attention and SOA on mean amplitudes 85–125 ms after sound onset was marginally significant for measurements taken across the entire scalp [$F(1, 22) = 3.67, p = .07$]; at the anterior and central regions where the effect of SOA was largest in the attend condition (A, AC, C), an interaction of attention and SOA emerged [$F(1, 22) = 4.29, p = .05, \eta_p^2 = .16$]. In contrast, similar analyses on the mean amplitudes 250–450 ms across the entire scalp, over the posterior regions where above-echo-threshold sounds elicited a larger positivity only in the attend

condition, and over the left and medial anterior regions where above-echo-threshold sounds elicited a larger positivity only in the unattend condition revealed no Attention $\times$ SOA interactions ($ps > .10$).

Shift in echo threshold One possible explanation for the differences in the effects of SOA in the attend and unattend conditions is that attention changed the echo threshold. If so, other pairs of SOAs in the unattend condition might reveal ERP effects similar to those observed in the attend condition. Analysis of the shortest SOA presented to each participant (T_{-5}) and the SOA 2 ms shorter than the echo threshold defined in the attend condition (T_{-2}) included the data from 18 participants; no data were collected in the T_{-2} condition for the other five participants, who had echo thresholds below what was

Fig. 6 ERPs elicited by T_{+1} and T_{-1} click pairs in the unattend condition. The data from all 15 regions are shown in the same positions depicted on the scalp map. The ERPs from two locations are also shown at a larger scale. Shading indicates the time window in which we observed a difference between the conditions.

predicted from screening. As is shown in Fig. 7, consistent with the idea that directing attention to the visual stimuli lowered the echo threshold, the longer of the two SOAs (T_{-2}) elicited a larger negativity 85–125 ms after onset [$F(1, 17) = 8.84$, $p = .01$, $\eta_p^2 = .34$]. This effect was numerically larger over anterior and central electrodes, similar to the difference found for the T_{+1} and T_{-1} comparison in the attend condition. Furthermore, we found no evidence of an Attention $\times$ SOA interaction for these conditions when the analysis was conducted on data collected across the scalp or at anterior and central sites (A, AC, C) ($ps > .25$). Also, no evidence of the posterior positivity that was observed in the attend condition 250–450 ms after sound onset was found in the parallel comparison for the unattend condition (T_{-2} and T_{-5}) across the scalp or at posterior sites (CP, P) ($ps > .10$).

Discussion

The present ERP study was the first to examine the relationship between attention and auditory object formation using the precedence effect. When attention was directed to sounds above the echo threshold, the perception of two auditory objects elicited an ORN and an LP; however, when attention was directed away from the same sounds to a difficult two-back visual task, neither the ORN nor the LP was evident. An effect resembling an ORN was elicited by unattended sounds with lower SOAs, indicating that the threshold for representing the lag click as a separate auditory object may have been lowered in the absence of attention. Previous ERP studies have suggested that auditory object formation occurs preattentively under conditions of harmonic grouping (Alain et al., 2001; Alain & Izenberg, 2003; Alain et al., 2002) and dichotic pitch

(Hautus & Johnson, 2005). Under conditions of the precedence effect, however, attentional modulation of the ORN was evident; within complex reverberant environments, attention is critical for spatial auditory object formation.

To reach these conclusions, the first objective was to determine that attention was indeed manipulated across conditions. The behavioral results indicated that participants had attended to each task as instructed. The behavioral responses to attended sounds for each participant were consistent with the precedence effect, suggesting that participants remained engaged in judging the click pairs. Behavioral performance in the two-back visual task was strong but not at ceiling, suggesting that participants remained focused on identifying visual targets and that the task remained challenging. The comparison of the ERPs elicited by identical attended and unattended sounds provided strong evidence that the tasks were effective at manipulating attention. Attended click pairs elicited a larger N1 and P3. These early and late effects are consistent with well-established ERP indices of attention (Hillyard, Hink, Schwent, & Picton, 1973; Polich, 2007) and provide clear evidence that participants were attending to the click pairs when judging the lag click and ignoring the click pairs while engaged in the two-back visual task. Importantly, this shift in attention across conditions modulated both early perceptual (N1) and late (P3) auditory processing.

The second objective was to replicate the ERP indices of the precedence effect for attended click pairs, the ORN and the LP. Attended click pairs 1 ms above the echo threshold, which typically elicited reports of the lag click as a separate source, elicited a larger negativity 85–125 ms after sound onset than did click pairs 1 ms below the echo threshold, which typically elicited reports of not hearing the lag click. The polarity, scalp distribution, and timing of this early negativity are consistent

Fig. 7 ERPs elicited by T_{-2} and T_{-5} click pairs in the unattend condition. The data from all 15 regions are shown in the same positions depicted on the scalp map. The ERPs from one location are also shown at a larger scale. Shading indicates the time window in which we observed a difference between the conditions.

with the ORNs reported in previous studies when participants reported hearing two sounds rather than one (Alain et al., 2001; Alain & Izenberg, 2003; Alain et al., 2002; Butcher et al., 2011; Dyson et al., 2005; Hautus & Johnson, 2005; Johnson et al., 2003; Sanders et al., 2008; Sanders et al., 2011). One noteworthy difference in the present study was that the ORN did not extend across the P2 time window. The length of the ORN may have been influenced by the visual stimuli or may have reflected a variability in ORN morphology that has not yet been fully described. In addition, the ORN observed in the present study began approximately 15 ms earlier than the ORNs reported in previous studies of the precedence effect (Sanders et al., 2008; Sanders et al., 2011). However, the N1 itself in the present study also peaked 13–25 ms earlier than the N1s reported in those previous studies. This pattern of results suggests that the ERP index of auditory object formation, the ORN, is tied to the timing of other ERP indices of early auditory perceptual processing. Importantly, the presence of the ORN for attended sounds in the present study is consistent with the behavioral data and provides strong evidence that most of the click pairs above the echo threshold were perceived as two distinct auditory objects.

In addition to the ORN, attended click pairs 1 ms above the echo threshold elicited a larger positivity across central and posterior scalp regions 250–450 ms after sound onset. This effect was consistent with the LPs observed in previous studies under conditions in which participants reported hearing two sounds rather than one. The LP may reflect the top-down processes involved in recognizing a sound and its meaning on the basis of prior knowledge (Alain et al., 2001), and may also be related to P300 effects associated with target detection and categorization (Polich, 2007). Although the factors contributing to the LP are difficult to specify, the presence of the LP for attended sounds indicates that at higher levels of processing, participants were able to confidently distinguish between click pairs above and below the echo threshold.

The third objective was to examine whether the ERP effects of SOA differed across attention conditions. In contrast with the ORN observed for attended click pairs, no difference was found for click pairs 1 ms above and below the echo threshold 85–125 ms after sound onset when participants were directing attention to the two-back visual task. Importantly, a significant interaction provided strong evidence that attention modulated the ORN, indicating that attention plays a critical role in the formation of auditory objects in the precedence effect. Furthermore, these results support the hypothesis that the precedence effect represents an active, dynamic system that extends to higher-level processes and is important for constructing perceptual models of acoustic space (Clifton et al., 1994; Clifton et al., 2002; Freyman & Keen, 2006; Keen & Freyman, 2009).

Additionally, when listeners directed attention to the visual stimuli, we found no evidence of a late positive effect for click pairs 1 ms above, as compared to click pairs 1 ms below, the echo threshold. This was expected, since the higher-level processes associated with the LP are necessarily attention-dependent; the LP has been shown to be modulated by attention in previous studies of harmonic grouping (Alain et al., 2001; Alain et al., 2002) and dichotic pitch (Hautus & Johnson, 2005). In the present study, however, no evidence of attentional modulation of the LP was found in testing the Attention Condition $\times$ SOA interaction, although the differences were in the expected direction, such that the positivity was larger for attended than for unattended sounds. If the LP observed for attended sounds does indeed reflect processing associated with detecting and categorizing auditory stimuli, it is probable that the interaction would have reached significance, given a larger sample size.

Interestingly, a larger left and medial anterior positivity was found 250–450 ms after sound onset for unattended click pairs 1 ms above the echo threshold, relative to unattended click pairs 1 ms below the echo threshold. This effect was unexpected and cannot be clearly linked with previously reported ERP effects. However, the fact that any difference emerged in the response to sounds above and below the echo threshold when attention was directed to the visual modality suggests that veridical representations of the sounds were maintained in the absence of attention.

The final objective was to explore whether echo thresholds may have been shifted by attention. With no evidence of an ORN for click pairs 1 ms above and below the echo threshold when listeners directed attention to the visual stimuli, it was possible that the echo threshold was higher in the absence of attention and that both of these sounds were heard as single, fused auditory objects. However, visual inspection of the data revealed no evidence of an ORN for pairs of sounds above the echo threshold, providing no support for this interpretation. Alternatively, it was possible that the echo threshold was lower in the absence of attention, and that sounds both 1 ms above and 1 ms below the echo threshold, as defined by responses to attended click pairs, were heard as two separate sources when attention was directed to the visual modality. Consistent with this idea, when participants directed attention to the visual stimuli, there was evidence of an effect resembling an ORN for click pairs 2 ms below, as compared to 5 ms below, the echo threshold established for attended click pairs. This effect was similar to the ORN measured for attended click pairs in terms of amplitude, distribution, and timing. The exploratory nature of this analysis precludes strong conclusions. However, the possibility that attention to the auditory modality raises the echo threshold is compatible with the claim that attention facilitates feature binding within the precedence effect, perhaps in a manner consistent with feature-integration theory (Hall et al., 2000; Thompson et al., 2001; Treisman & Gelade,

1980; Treisman & Schmidt, 1982). In raising the echo threshold within reverberant environments, auditory attention would provide a reduction in clutter and improvements to comprehension and the localization of actual sound sources.

In conclusion, ERP measures of the precedence effect provided a unique opportunity to examine auditory object processing while participants directed attentional resources toward and away from sounds. In the absence of auditory attention, ERP indices of auditory object perception disappeared. These results suggest that within complex environments, the extent to which direct and reflected sounds are grouped by the precedence effect into coherent auditory objects fluctuates as we attend to different aspects of our surroundings.

Acknowledgments The research reported here was partially funded by NIH Grant No. R01 DC001625, awarded to R.L.F. and L.D.S., and the training of B.H.Z. was funded by NSF GRFP Award No. 1451512. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIDCD, NIH, or NSF. We thank Andrew Cohen for assistance with the behavioral data analysis, and Claire Moore-Cantwell for assistance with the ERP figures.

References

- Alain, C., Arnott, S. R., & Picton, T. W. (2001). Bottom-up and top-down influences on auditory scene analysis: Evidence from event-related brain potentials. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 1072–1089. doi:10.1037/0096-1523.27.5.1072
- Alain, C., & Izenberg, A. (2003). Effects of attentional load on auditory scene analysis. *Journal of Cognitive Neuroscience*, 15, 1063–1073.
- Alain, C., Schuler, B. M., & McDonald, K. L. (2002). Neural activity associated with distinguishing concurrent auditory objects. *Journal of the Acoustical Society of America*, 111, 990–995. doi:10.1121/1.1434942
- Blauert, J. (1997). *Spatial hearing: the psychophysics of human sound localization*. Cambridge, MA: MIT Press.
- Blodgett, H. C., Wilbanks, W. A., & Jeffress, L. A. (1956). Effect of large interaural time differences upon the judgment of sidedness. *Journal of the Acoustical Society of America*, 28, 639–643.
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- Butcher, A., Govenlock, S. W., & Tata, M. S. (2011). A lateralized auditory evoked potential elicited when auditory objects are defined by spatial motion. *Hearing Research*, 272, 58–68.
- Clifton, R. K. (1987). Breakdown of echo suppression in the precedence effect. *Journal of the Acoustical Society of America*, 82, 1834–1835.
- Clifton, R. K., & Freyman, R. L. (1989). Effect of click rate and delay on breakdown of the precedence effect. *Perception & Psychophysics*, 46, 139–145. doi:10.3758/BF03204973
- Clifton, R. K., Freyman, R. L., & Litovsky, R. (1994). Listeners' expectations about echoes can raise or lower echo threshold. *Journal of the Acoustical Society of America*, 95, 1525–1533.
- Clifton, R. K., Freyman, R. L., & Meo, J. (2002). What the precedence effect tells us about room acoustics. *Perception & Psychophysics*, 64, 180–188. doi:10.3758/BF03195784
- Clifton, R. K., Morriongiello, B. A., Kulig, J. W., & Dowd, J. M. (1981). Newborns' orientation toward sound: Possible implications for cortical development. *Child Development*, 52, 833–838.
- Djelani, T., & Blauert, J. (2001). Some new aspects of the buildup and breakdown of the precedence effect. In D. J. Breebart, A. J. Houtsma, V. F. Prijs, & R. Schoonhoven (Eds.), *Physiological and psychophysical bases of auditory function* (pp. 200–207). Maastricht, The Netherlands: Shaker.
- Dyson, B. J., Alain, C., & He, Y. (2005). Effects of visual attentional load on low-level auditory scene analysis. *Cognitive, Affective, & Behavioral Neuroscience*, 5, 319–338. doi:10.3758/CABN.5.3.319
- Freyman, R. L., Clifton, R. K., & Litovsky, R. Y. (1991). Dynamic processes in the precedence effect. *Journal of the Acoustical Society of America*, 90, 874–884.
- Freyman, R. L., & Keen, R. (2006). Constructing and disrupting listeners' models of auditory space. *Journal of the Acoustical Society of America*, 120, 3957–3965. doi:10.1121/1.2354020
- Goverts, S. T., Houtgast, T., & van Beek, H. H. M. (2000). The precedence effect for lateralization at low sensation levels. *Hearing Research*, 148, 88–94.
- Grantham, D. W. (1996). Left-right asymmetry in the buildup of echo suppression in normal-hearing adults. *Journal of the Acoustical Society of America*, 99, 1118–1123.
- Haas, H. (1951). On the influence of a single echo on the intelligibility of speech. *Acustica*, 1, 48–58.
- Hall, M. D., Pastore, R. E., Acker, B. E., & Huang, W. (2000). Evidence for auditory feature integration with spatially distributed items. *Perception & Psychophysics*, 62, 1243–1257. doi:10.3758/BF03212126
- Hartung, K., & Trahiotis, C. (2001). Peripheral auditory processing and investigations of the "precedence effect" which utilize successive transient stimuli. *Journal of the Acoustical Society of America*, 110, 1505–1513.
- Hautus, M. J., & Johnson, B. W. (2005). Object-related brain potentials associated with the perceptual segregation of a dichotically embedded pitch. *Journal of the Acoustical Society of America*, 117, 275–280.
- Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, 182, 177–180.
- Johnson, B. W., Hautus, M., & Clapp, W. C. (2003). Neural activity associated with binaural processes for the perceptual segregation of pitch. *Clinical Neurophysiology*, 114, 2245–2250.
- Keen, R., & Freyman, R. L. (2009). Release and re-buildup of listeners' models of auditory space. *Journal of the Acoustical Society of America*, 125, 3243–3252. doi:10.1121/1.3097472
- Litovsky, R. Y., Colburn, H. S., Yost, W. A., & Guzman, S. J. (1999). The precedence effect. *Journal of the Acoustical Society of America*, 106, 1633–1654.
- Litovsky, R. Y., & Godar, S. P. (2010). Difference in precedence effect between children and adults signifies development of sound localization abilities in complex listening tasks. *Journal of the Acoustical Society of America*, 128, 1979–1991. doi:10.1121/1.3478849
- Miller, S. D., Litovsky, R. Y., & Kluender, K. R. (2009). Predicting echo thresholds from speech onset characteristics. *Journal of the Acoustical Society of America*, 125, 134–140.
- Morriongiello, B. A., Kulig, J. W., & Clifton, R. K. (1984). Developmental changes in auditory temporal perception. *Child Development*, 55, 461–471.
- Muir, D. W., Clifton, R. K., & Clarkson, M. G. (1989). The development of a human auditory localization response: A U-shaped function. *Canadian Journal of Psychology*, 43, 199–216. doi:10.1037/h0084220
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118, 2128–2148. doi:10.1016/j.clinph.2007.04.019
- Saberi, K., & Antonio, J. V. (2003). Precedence-effect thresholds for a population of untrained listeners as a function of stimulus intensity

- and interclick interval. *Journal of the Acoustical Society of America*, 114, 420–429.
- Sanders, L. D., Joh, A. S., Keen, R. E., & Freyman, R. L. (2008). One sound or two? Object-related negativity indexes echo perception. *Perception & Psychophysics*, 70, 1558–1570. doi:[10.3758/PP.70.8.1558](https://doi.org/10.3758/PP.70.8.1558)
- Sanders, L. D., Zobel, B. H., Freyman, R. L., & Keen, R. (2011). Manipulations of listeners' echo perception are reflected in event-related potentials. *Journal of the Acoustical Society of America*, 129, 301–309.
- Schubert, E. D., & Wernick, J. (1969). Envelope versus microstructure in the fusion of dichotic signals. *Journal of the Acoustical Society of America*, 45, 1525–1531.
- Seeber, B. U., & Hafer, E. R. (2011). Failure of the precedence effect with a noise-band vocoder. *Journal of the Acoustical Society of America*, 129, 1509–1521. doi:[10.1121/1.3531836](https://doi.org/10.1121/1.3531836)
- Shinn-Cunningham, B. G., Zurek, P. M., Durlach, N. I., & Clifton, R. K. (1995). Cross-frequency interactions in the precedence effect. *Journal of the Acoustical Society of America*, 98, 164–171.
- Thompson, W. F., Hall, M. D., & Pressing, J. (2001). Illusory conjunctions of pitch and duration in unfamiliar tone sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 128–140. doi:[10.1037/0096-1523.27.1.128](https://doi.org/10.1037/0096-1523.27.1.128)
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97–136. doi:[10.1016/0010-0285\(80\)90005-5](https://doi.org/10.1016/0010-0285(80)90005-5)
- Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, 14, 107–141.
- Wallach, H., Newman, E. B., & Rosenzweig, M. R. (1949). The precedence effect in sound localization. *American Journal of Psychology*, 62, 315–336.