Manipulations of listeners' echo perception are reflected in event-related potentials

Lisa D. Sanders[a] and Benjamin H. Zobel
Neuroscience and Behavior Program, Department of Psychology, University of Massachusetts, Amherst, Massachusetts 01003

Richard L. Freyman
Department of Communication Disorders, University of Massachusetts, Amherst, Massachusetts 01003

Rachel Keen
Department of Psychology, University of Virginia, P.O. Box 400400, Charlottesville, Virginia 22904

(Received 3 June 2010; revised 12 October 2010; accepted 13 October 2010)

To gain information from complex auditory scenes, it is necessary to determine which of the many loudness, pitch, and timbre changes originate from a single source. Grouping sound into sources based on spatial information is complicated by reverberant energy bouncing off multiple surfaces and reaching the ears from directions other than the source's location. The ability to localize sounds despite these echoes has been explored with the precedence effect: Identical sounds presented from two locations with a short stimulus onset asynchrony (e.g., 1–5 ms) are perceived as a single source with a location dominated by the lead sound. Importantly, echo thresholds, the shortest onset asynchrony at which a listener reports hearing the lag sound as a separate source about half of the time, can be manipulated by presenting sound pairs in contexts. Event-related brain potentials elicited by physically identical sounds in contexts that resulted in listeners reporting either one or two sources were compared. Sound pairs perceived as two sources elicited a larger anterior negativity 100–250 ms after onset, previously termed the object-related negativity, and a larger posterior positivity 250–500 ms. These results indicate that the models of room acoustics listeners form based on recent experience with the spatiotemporal properties of sound modulate perceptual as well as later higher-level processing. © *2011 Acoustical Society of America*. [DOI: 10.1121/1.3514518]

PACS number(s): 43.66.Qp, 43.66.Lj, 43.64.Ri [BLM] Pages: 301–309

I. INTRODUCTION

In many typical listening environments the majority of sound reaching the two ears is reflected energy, approaching from directions other than that of the original source. The ability to localize sounds in such complex settings is aided by a phenomenon known as the precedence effect, through which the direct and reverberant energy are gathered into a single object percept near the position of the sound source (for review, see Blauert, 1997). Although obviously critical for successful listening in almost every natural environment, a complete understanding of the peripheral and central processes that contribute to the precedence effect remains elusive. Many researchers have employed simple sound pairs, a lead sound presented at one location and a single simulated reflection presented later at a different location, to achieve the stimulus control required to approach such an understanding (e.g., Wallach *et al.*, 1949; Gardner, 1968; Zurek, 1980). These studies have shown that there are both lower and upper temporal limits to the precedence effect. On the low end, when reflection delays are less than 1 ms, localization of the unified object depends strongly on both the lead and lag sounds. At the upper end, delays are reached at which the object begins to break apart and an "echo" is heard at the location of the lag sound. This upper limit, called echo threshold, varies in length from just a few milliseconds to around 50 ms depending on the individual listener and type of stimulus (Saberi and Perrott, 1990; Fitzpatrick *et al.*, 1999; Litovsky *et al.*, 1999).

Even for a specific stimulus and individual, echo threshold is not a fixed value. It is malleable and depends on what sounds the listener is exposed to in the moments just before the test stimulus that requires a judgment (Clifton, 1987; Freyman *et al.*, 1991; Blauert and Col, 1992; Clifton *et al.*, 1994; Grantham, 1996; Yost and Guzman, 1996). Specifically, echo threshold can be built up through repeated presentation of identical lead–lag pairs, and the built-up threshold can be broken down through repeated presentation of aberrant lead–lag configurations. Buildup of the precedence effect is defined as the increase in echo threshold after repeated presentations, compared to echo threshold for an isolated sound pair. It has been suggested that the effects of recent context on echo threshold reflect listeners forming a model of auditory space through repeated presentation of identical lead–lag configurations and discarding that model quickly when exposed to sounds that disconfirm it (Clifton *et al.*, 2002; Freyman and Keen, 2006; Keen and Freyman, 2009).

The purpose of the present study was to gain further insight into the nature of the buildup of echo threshold by measuring neurophysiological responses. There are several advantages to augmenting psychophysical measurements of

[a]Author to whom correspondence should be addressed. Electronic mail: lsanders@psych.umass.edu

the buildup of echo threshold with physiological ones. First, objective evidence that repeated presentation of an identical lead–lag pair changes a perceptual threshold would provide important confirmation of the conclusions drawn from more subjective measurements. Second, physiological measurements could provide information about the levels of cortical processing that are involved in the buildup phenomenon. Specifically, effects of recent auditory experience on evoked potentials in the first 250 ms after sound onset can be interpreted as modulation of perception, whereas differences in later portions of the waveforms more likely reflect higher-level processing. Finally, although attention was not manipulated in the current study, neurophysiological measures would ultimately make it possible to answer questions that cannot be addressed with behavioral responses alone, including whether attention is necessary for the buildup of echo threshold to occur.

Several previous studies have addressed the physiological underpinnings of the precedence effect in humans. An initial study that involved recording auditory brainstem responses (ABRs) and event-related potentials (ERPs) showed that the earliest potential affected by the presence of a preceding sound peaked around 40 ms after onset in the middle latency range (Liebenthal and Pratt, 1999). This positive peak (Pa) was reduced in amplitude when a sound was presented as a simulated reflection rather than in isolation. A second study reported psychophysical, ABR, and ERP measurements and concluded that the earliest evidence of the precedence effect in the ERP waveforms was the mismatch negativity (MMN) which peaked around 150 ms after the onset of the lag sound (Damaschke *et al.*, 2005). However, neither of these studies recorded participant responses during or under precisely the same conditions as electrophysiological measurements. Since the relationship between stimulus onset asynchrony (SOA) (i.e., the delay between the onsets of the lead and lag sounds) and a listener reporting hearing the lag sound as a separate source varies widely among individuals and, as reviewed above, depends on the immediately preceding auditory context, it is difficult to make direct connections between perception and the underlying neurosensory processing without recording behavioral and electrophysiological responses on the same trials.

By measuring behavior and ERPs on the same trials from the same participants, Sanders *et al.* (2008) showed that physically identical click pairs presented near the individual's echo threshold elicited a larger negativity 100–250 ms over anterior and medial regions on trials when listeners reported hearing compared to not hearing the lag sound as a separate source. Unlike previous precedence effect studies, this negativity was unrelated to the frequency with which participants reported hearing an echo and so could not be interpreted as an MMN. Instead, the difference in ERPs when listeners heard two compared to one sound source was interpreted as an object-related negativity (ORN), an effect that has primarily been explored in the domain of pitch perception. Specifically, mistuning a single frequency in a harmonic complex to the extent that listeners report hearing two simultaneously presented distinct pitches (one based on the fundamental frequency of the complex and one on the frequency of the mistuned harmonic), elicited a larger negativity 100–250 ms over anterior and medial regions (Alain *et al.*, 2001; Dyson *et al.*, 2005). The ORN has also been reported in studies that induced two-sound perception through the use of dichotic pitch stimuli (Johnson *et al.*, 2003; Hautus and Johnson, 2005; Chait *et al.*, 2006). Further, when the amount of harmonic mistuning was not enough to result in perception of two distinct pitches, presenting the mistuned harmonic from a different location elicited the ORN (McDonald and Alain, 2005). The link between the ORN and reported perception of two compared to one sound source in the precedence effect (Sanders *et al.*, 2008) provides a tool for indexing the buildup of echo threshold. Specifically, when a test lead–lag sound pair follows a repeated train of identical sounds with the same SOA and locations (i.e., a buildup train), fewer echoes should be reported and the ORN should be reduced in amplitude.

However, there are methodological hurdles that must be overcome to use the ORN to investigate buildup of echo threshold. If ERPs elicited by test clicks on trials that do and do not include a buildup train are compared directly, echo perception will be confounded with recovery cycle effects. Sounds recently preceded by other similar sounds elicit smaller amplitude auditory evoked potentials (Bess and Ruhm, 1972; Erwin and Buchwald, 1986; Budd *et al.*, 1998; Coch *et al.*, 2005) such that buildup trains would result in both fewer trials on which listeners report hearing the lag sound as a separate source and more refractory ERPs. One solution to this problem is to include a train of clicks on every trial. When the lead–lag configurations of the train and test clicks match, echo threshold will be built up; a switch in the lead and lag locations between the train and the test sounds effectively negates the buildup (Clifton and Freyman, 1989; Clifton *et al.*, 1994; Yost and Guzman, 1996; Litovsky *et al.*, 1999). Unfortunately, the lead–lag location reversal would be expected to elicit an MMN as a result of the sudden change in the stimulus (Sams *et al.*, 1984; Paavilainen *et al.*, 1989; Pekkonen *et al.*, 1995; Picton *et al.*, 2000; Näätänen *et al.*, 2004). The differences in ERPs could be attributed either to the number of perceived sound sources or to the new location of the lead sound.

To study the buildup of the precedence effect with physiological measures, it is important to measure behavioral and electrophysiological responses in the same participants on the same trials, and to avoid confounds with both the recovery cycle and the MMN. Two earlier studies that addressed the topic met some, but not all, of these conditions (Dimitrijevic and Stapells, 2006; Spierer *et al.*, 2009). Unambiguous interpretation of ERP results requires that both the test sounds and the immediately preceding sounds be identical on trials when the listener does and does not report hearing the lag sound as a separate source. But how can the number of sound sources listeners perceive be manipulated for physically identical trains and test stimuli? One solution is to exploit the fact that echo threshold can actually be lowered below that for isolated lead–lag pairs by presenting sounds from only one side (single-source) prior to the test pair (Freyman *et al.*, 1991; Keen and Freyman, 2009). In buildup trains, each repetition raises echo threshold by a small amount until it reaches asymptote after about nine presentations (Freyman *et al.*, 1991). If echo threshold is first lowered by presenting single-source sounds, a subsequent buildup train may not raise echo threshold to the same level as that buildup train alone. To give a specific example, an individual might have an echo threshold of 8 ms for isolated click pairs, which increases to 12 ms when that pair is preceded by a buildup train, and which decreases to 5 ms when that pair is immediately preceded by single-source clicks. If those two trains are combined for the same individual, echo threshold might first be lowered to 5 ms such that the 4-ms increase induced by the buildup train results in a final echo threshold of only 9 ms. This manipulation might result in listeners reporting a single sound source for a particular lead–lag onset asynchrony after a buildup train, but reporting two sources when the buildup to this critical test pair has been preceded by a single-source train. Data reported by Keen and Freyman (2009) are consistent with this possibility. When a train of single-source clicks from the lead loudspeaker preceded a train of lead–lag clicks, echo thresholds increased during the buildup phase but did not rise above the threshold measured with the test clicks presented in isolation. This phenomenon, referred to as "depressed buildup" in the current paper, could provide a crucial control condition for buildup trials. ERPs elicited by test sounds following buildup and depressed buildup contexts would be equally refractory and there would be no mismatch response in either condition since the test sounds are identical to the lead–lag pairs that precede them.

The present study was designed to investigate the physiology of echo threshold buildup by verifying and then exploiting depressed buildup as a new comparison condition. Electroencephalogram (EEG) was recorded while participants listened to click pairs with a range of SOAs and indicated whether they heard the lag sound as a separate source or not. The click pairs were presented in three contexts designed to manipulate echo threshold, as depicted in Fig. 1: In isolation, following a train of seven identical pairs that typically result in the buildup of echo threshold, and following a train of five single-source clicks that typically depresses echo threshold

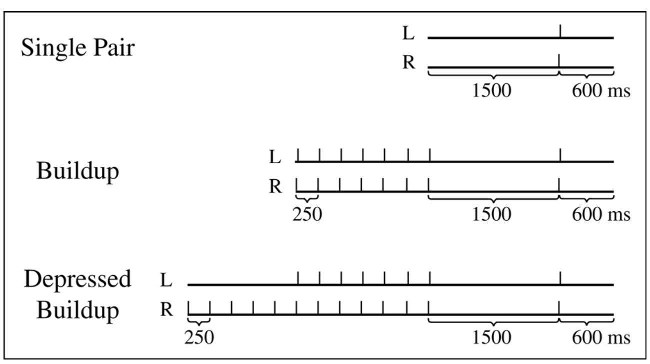

FIG. 1. Depiction of the three contexts in which test click pairs were presented. All trials began with the onset of a fixation point. Test pairs were a click from the right speaker followed 1–14 ms later by an identical click from the left speaker. All trials included 1500 ms of silence before the test pair and 600 ms after the test pair before participants were prompted to indicate whether they had heard one sound or two. Click pairs in the buildup and depressed buildup trains were delivered at a rate of 4/s.

and then the buildup train. To the extent that the depressed buildup condition yields lower echo thresholds than the buildup condition, the effects of recent auditory context on auditory perception can be isolated by comparing ERPs elicited by physically identical stimuli when listeners report hearing a single sound source on buildup trials and hearing two sound sources on depressed buildup trials. This comparison is expected to yield both an ORN and a later posterior positivity reflecting the effects of recent auditory experience on both perceptual and higher-level processing.

II. METHOD

A. Participants

Eighteen adults (four females, aged 20–30 yr) provided the data included in the analysis. Data from four participants were excluded because of excessive noise in the EEG (excessive blinking and/or low frequency drift likely caused by skin potentials). An additional six adults completed a screening task but did not participate in the experiment since their pattern of behavioral responses made it difficult to determine the SOAs that should be employed for test trials. All participants reported being right-handed, having no neurological problems, and not taking psychoactive medication within a year of the study. Further, all were found to have normal hearing thresholds (< 20 dB HL) at 1, 2, 4, and 8 kHz. Participants provided informed consent and were compensated for their time at a rate of $10/h.

B. Procedure

Stimuli were 181-μs positive rectangular pulses of four sample points (at 22.05 kHz sampling rate) that listeners typically describe as clicks. Click pairs were constructed by pasting the lead click into the right channel and, 1–14 ms later (in 1-ms steps), an identical lag click into the left channel of a stereo WAV file. All sounds were presented over M-Audio StudioPro3 loudspeakers (Irwindale, CA) with EPRIME software running on a PC with a Creative Audigy 2 ZS sound card (Singapore). Each loudspeaker was located 1.4 m from the participants and 55° from midline. Sound level over the two loudspeakers was equated immediately before each participant's arrival by measuring the output level of click trains (ten clicks per second) presented from a single loudspeaker and adjusting the channel-specific volume on the PC to produce a reading of 70 dBA. Sounds were presented in a 2.5 m $\times$ 3.5 m acoustically shielded and sound-deadened room.

Participants were first introduced to their task by listening to isolated click pairs assumed to be well below (SOA = 1 ms) and well above (SOA = 25 ms) echo threshold so they could experience what it was like to hear and to not hear the lag sound as a separate source. They were then asked to complete a screening task in which click pairs with seven SOAs (in 2-ms steps) were presented as single pairs (140 trials with the first 35 excluded as practice) and following the buildup context. On the 28 buildup trials, test clicks were preceded by a train of seven identical pairs presented at a rate of 4/s and 1.5 s of silence. Responses on this screening were used to select listeners who seemed to have a higher echo threshold in the buildup compared to the single pair condition and to determine the six SOAs (in 1-ms steps) for each participant that were most likely to encompass echo thresholds in the three conditions for the test trials. Specifically, test SOAs were selected such that the shortest delay fell 1 ms below the estimated echo threshold for single pairs in the screening task.

On test trials, as with screening trials, participants were told in separate blocks that they would hear either one test click or a series of sounds followed by a pause (1.5 s) and one test click. They were told that their task was to make a judgment for the test click only. Additionally, they were informed that the test click would always include a sound presented from the right loudspeaker. The task on each trial was to report if they also heard a sound from the left side by pressing one of two buttons on a response box. The lead sound was always presented on the right side to reduce the number of trials and length of the data collection sessions while retaining enough power to detect consistent, small amplitude ERP effects. Trials were presented in four single pair blocks of 96 trials each (16 at each of six SOAs in random order), as well as eight buildup blocks and eight depressed buildup blocks of 48 trials each (eight at each of six SOAs in random order). On single pair trials (see Fig. 1), a fixation point appeared on the screen indicating participants were to look straight ahead and avoid blinking and other motion. The test click pair with one of six SOAs selected on the basis of the screening task was presented 1.5 s later. The fixation point remained on the screen for another 600 ms and was followed by a prompt to respond. Participants controlled when the next trial in the same block began by pressing a button following another prompt. On buildup trials, the fixation point appeared on the screen at the onset of a series of seven click pairs with the same SOA presented at a rate of 4/s and followed by 1.5 s of silence. The test pair with the same SOA employed in the buildup context was then presented and the fixation point remained on the screen for another 600 ms. On depressed buildup trials, the fixation point appeared at the onset of a series of five single-source clicks from the right loudspeaker only and then the seven click pairs of the same SOA. Again, a test pair was presented after 1.5 s of silence and the fixation point remained on the screen for an additional 600 ms. With this design, all trial types were physically identical from 1.5 s before the test click pair through the participant's response; the buildup and depressed buildup trials were physically identical beginning 3.25 s before the test click pair.

C. EEG recording and analysis

EEG was recorded continuously throughout all of the test trials (250 Hz sampling rate, 0.01–100 Hz bandpass) from 128 electrodes (EGI, Eugene, OR). Scalp impedances at all electrode sites were maintained under 50 kΩs which has been shown to be sufficient to reduce electrical noise from sources other than the brain with high input-impedance bio-amplifiers (Ferree *et al.*, 2001). An SOA at which a participant reported hearing the lag sound as a separate source on most single pair and depressed buildup trials, and failing to hear the lag sound as a separate source on most buildup trials, was selected as the between-thresholds SOA for that individual. Segments of EEG beginning 100 ms before and continuing 500 ms after the onset of test click pairs with the between-thresholds SOA were isolated for three conditions: Single pair when the listener reported hearing the lag sound, depressed buildup when the listener reported hearing the lag sound, and buildup when the listener reported not hearing the lag sound. A different SOA for which an individual reported hearing the lag sound as a separate source on about half of depressed buildup trials (40%–60%) was also identified. Trials with artifacts from muscle tension defined by large amplitude deflections near 40 Hz at electrodes near the face and ears, blinks or eye-movements recognized by large-amplitude low-frequency changes in the EEG of opposite polarity around the eyes, or motion marked by large shifts in the EEG at many or all electrodes across the scalp were excluded from analysis. EEG from remaining trials was averaged together by condition. The 100 ms before test pair onset was used as a baseline and ERPs were re-referenced to the averaged mastoid recording.

Peak latency and average amplitude measurements were taken on the waveforms in the following time windows after lead sound onset: 30–80 ms (P1), 80–130 ms (N1), and 170–220 ms (P2). Additional average amplitude measurements were made at 100–250 ms and 250–500 ms. Measurements were made on data collected at 81 of the 128 electrode sites across the scalp such that electrode position could be included as multiple factors in statistical analyses. Specifically, measurements from nine electrodes were grouped together in a 3 (anterior, central, posterior or ACP) $\times$ 3 (left, medial, right or LMR) grid and electrode position within those groups of nine was ignored. Data from pairs of conditions designed to index perception of the lag sound as a separate source (lag sound not reported for the buildup context and reported for the depressed buildup context), recovery cycle effects (lag sound reported for both the single pair and depressed buildup conditions), and the combination of perception of the lag sound and recovery cycle effects (lag sound not reported for the buildup context and reported for the single pair condition) were entered in 2 (Condition) $\times$ 3 (ACP electrode position) $\times$ 3 (LMR electrode position) repeated-measures analysis of variances (ANOVAs) (Greenhouse–Geisser corrected). All significant ($p < 0.05$) interactions of Condition and electrode position factors were followed by ANOVAs conducted separately for each level of the relevant electrode position factor.

III. RESULTS

A. Behavioral responses

An SOA was selected for each participant as falling between echo thresholds for the depressed buildup and buildup contexts. At this SOA, listeners reported hearing the lag sound as a separate source on a high percentage of single pair trials [mean (M) = 83.5%, standard error (SE) = 3.4] and depressed buildup trials (M = 75.5%, SE = 3.6) as shown in Fig. 2. There was no statistical difference in the percentage of reported echoes for these two conditions. In

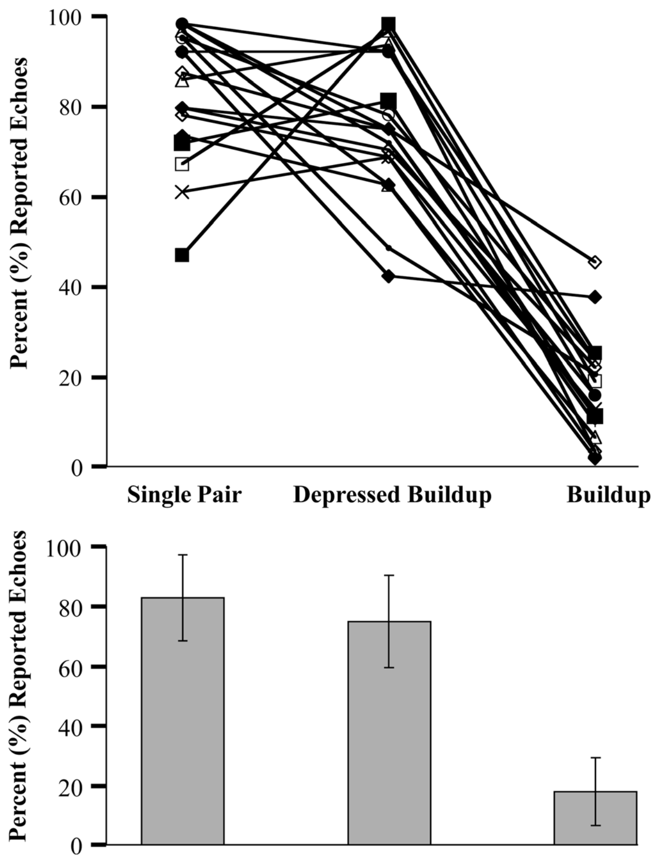

FIG. 2. Percent of trials (out of 64) on which each participant reported hearing the lag sound as a separate source for test pairs at a single SOA selected for that individual. Participants were more likely to report hearing two auditory objects (reflecting lower echo thresholds) in the single pair and depressed buildup conditions. Group averages and standard error bars are shown in the lower panel.

contrast, listeners reported hearing the lag sound on a much lower percentage of buildup trials ($M = 18.2\%$, $SE = 2.7$). These responses differed from those of both the single pair [$t(17) = 15.5$, $p < 0.001$] and the depressed buildup conditions [$t(17) = 12.1$, $p < 0.001$]. There was a great deal of inter-subject variability in the SOA at which these differences in responses were observed. On average, the SOA that fell most clearly between depressed buildup echo threshold and buildup echo threshold was 8 ms, but this value ranged from 3 to 14 ms across participants.

B. ERPs

The comparison of ERPs elicited by a test pair at a single SOA for each participant such that the lag sound was not reported for the buildup context and was reported for the depressed buildup context indexes perception of the lag sound as a separate source for physically identical stimuli. For both conditions, the average latency of the first positive peak (P1) was 63 ms, the first negative peak (N1) was 108 ms, and the second positive peak (P2) was 197 ms. All of these peaks were largest at anterior and central electrodes and, as expected, did not differ in latency across condition ($p > 0.7$).

There were no differences in P1 amplitude (30–80 ms) between the buildup and depressed buildup conditions. In contrast, mean amplitude between 100 and 250 ms did differ at some electrode positions [Condition × ACP, $F(2,34) = 5.83$, $p < 0.05$]. Specifically, as shown in Fig. 3, at anterior and central electrodes test clicks in the depressed buildup condition for which listeners reported hearing the lag sound elicited a larger negativity than identical test clicks in the buildup trials for which listeners reported not hearing the lag sound [$F(1,17) = 7.13$, $p < 0.01$]. There was no evidence of a correlation between the size of the effect of context on behavioral responses (difference in number of trials on which the lag sound was reported following the depressed buildup and buildup contexts) and on ERP amplitude between 100 and 250 ms ($p > 0.4$). Between 250 and 500 ms, trials on which listeners reported hearing the lag sound in the depressed buildup condition resulted in a larger positivity over central and posterior electrodes compared to when they did not hear it in the buildup condition [Condition × ACP, $F(2,34) = 5.81$, $p < 0.05$; Condition × ACP × LMR, $F(4,68) = 3.80$, $p < 0.05$; main effect of Condition at posterior and medial electrodes, $F(1,17) = 9.33$, $p < 0.01$). Participants who showed a larger effect of context on behavioral responses also tended to show a larger amplitude ERP effect in this time window (Spearman's $r = 0.48$, $p < 0.05$). However, these results must be interpreted with caution since participants who evidenced larger effects of context on behavioral responses necessarily had more trials averaged into each condition for ERP measurements.

Perception of two sound sources compared to one source can also be indexed with ERPs elicited by single pairs of clicks when listeners reported hearing the lag sound and the test pairs with the same SOA in the buildup context when listeners reported not hearing the lag sound. However, in this comparison the effects of perceiving one or two sound sources are confounded with the recovery cycle. The single pairs compared to the same sound following the buildup context elicited a larger negativity 80–130 ms [$F(1,17) = 12.73$, $p < 0.01$] over anterior and central regions [Condition × ACP, $F(2,34) = 8.17$, $p < 0.01$]. The larger negativity for the single pair condition continued into the longer time range of 100–250 ms [$F(1,17) = 5.37$, $p < 0.05$] and was again larger over anterior and central regions [Condition × ACP, $F(2,34) = 10.75$, $p < 0.001$]. This negativity was followed by a larger positivity 250–500 ms for the single pair condition [$F(1,17) = 8.89$, $p < 0.01$] over posterior and central regions [Condition × ACP × LMR, $F(4,68) = 3.19$, $p < 0.05$].

The refractory effects alone were indexed by comparing ERPs elicited in the single pair and depressed buildup conditions when perception of the test pair was identical (listeners reported hearing the lag sound as a separate source), but the sounds preceding the test pair differed (several clicks in depressed buildup condition, silence in the single pair condition). The difference waves for this comparison are shown in Fig. 4. This comparison yielded a larger negativity for the single pair condition between 80 and 130 ms [$F(1,17) = 11.71$, $p < 0.01$], which was larger over anterior and central regions [Condition × ACP, $F(2,34) = 4.96$, $p < 0.05$]. Unlike the comparison that includes perception of the lag sound as a separate source and refractory effects, this difference did not continue into the longer 100–250 ms time range.

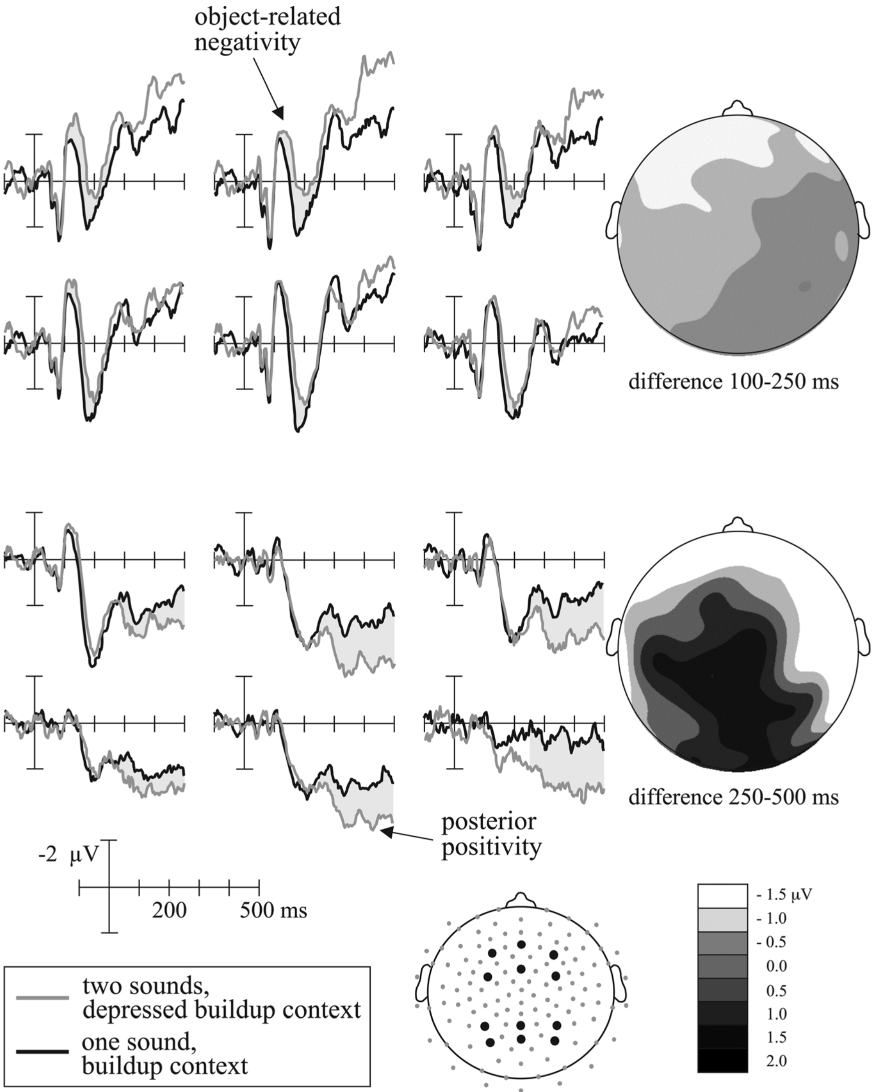

FIG. 3. Event-related potentials time-locked to the onset of the lead sound for a single stimulus onset asynchrony for each participant when they reported hearing the lag sound on depressed buildup trials (gray) and when they reported not hearing the lag sound on buildup trials (black). Waveforms are shown for the 12 recording sites depicted on the electrode map. Topographic plots show the difference in amplitude (reported lag sound minus did not report lag sound) across all electrodes. The difference in amplitude between 100 and 250 ms over anterior regions is the ORN and between 250 and 500 ms over posterior regions is the late posterior positivity.

The refractory effect comparison also evidenced a larger positivity for the single pair condition between 250 and 500 ms [$F(1,17) = 7.16$, $p < 0.05$]. However, unlike the positivity following the ORN, this positivity was largest at anterior electrodes [Condition × ACP, $F(2,34) = 5.82$, $p < 0.05$].

The above analyses compared electrophysiological responses to sounds presented in different contexts. Previous evidence showed differences in ERPs elicited by isolated click pairs near echo threshold in *identical* contexts when listeners report hearing and not hearing the lag sound (Sanders *et al*., 2008). In the current study, the context manipulations were effective enough that for the selected SOA, there were very few trials on which listeners reported hearing the lag sound as a separate source following the buildup context or not reporting the lag sound after the depressed buildup context. However, the range of SOAs employed for test trials did include sound pairs near depressed buildup echo threshold for 16 of the 18 participants. That is, it was possible to identify an SOA at which an individual reported hearing the lag sound on about half of the depressed buildup trials (range 40%–60%). As expected, this SOA was consistently shorter than that selected for the comparisons across contexts. Between 100 and 250 ms over anterior and central regions, test clicks in the depressed buildup condition elicited a larger negativity when participants reported hearing the lag sound as a separate source [Condition × ACP, $F(2,30) = 8.03$, $p < 0.01$; main effect of Condition at anterior and central electrodes, $F(1,15) = 4.52$, $p < 0.05$]. However, there was no evidence of differences in ERPs between 250 and 500 ms for trials on which listeners did and did not report hearing the lag sound (p's > 0.60).

IV. DISCUSSION

These results provide evidence that both perceptual and later higher-level auditory processing are modulated by recent auditory experience that also influences the subjective perception of the number of sound sources in the precedence effect. As lead–lag pairs of clicks with a fixed onset asynchrony are repeated in a train, the perception of the lag sound fades away and the single remaining source is localized near the lead sound. In other words, the delay at which the lag sound is reported to be a separate source increases, which has been called the buildup of echo threshold (Freyman *et al*., 1991; Blauert and Col, 1992; Clifton *et al*., 1994; Grantham, 1996; Yost and Guzman, 1996). To measure the electrophysiological

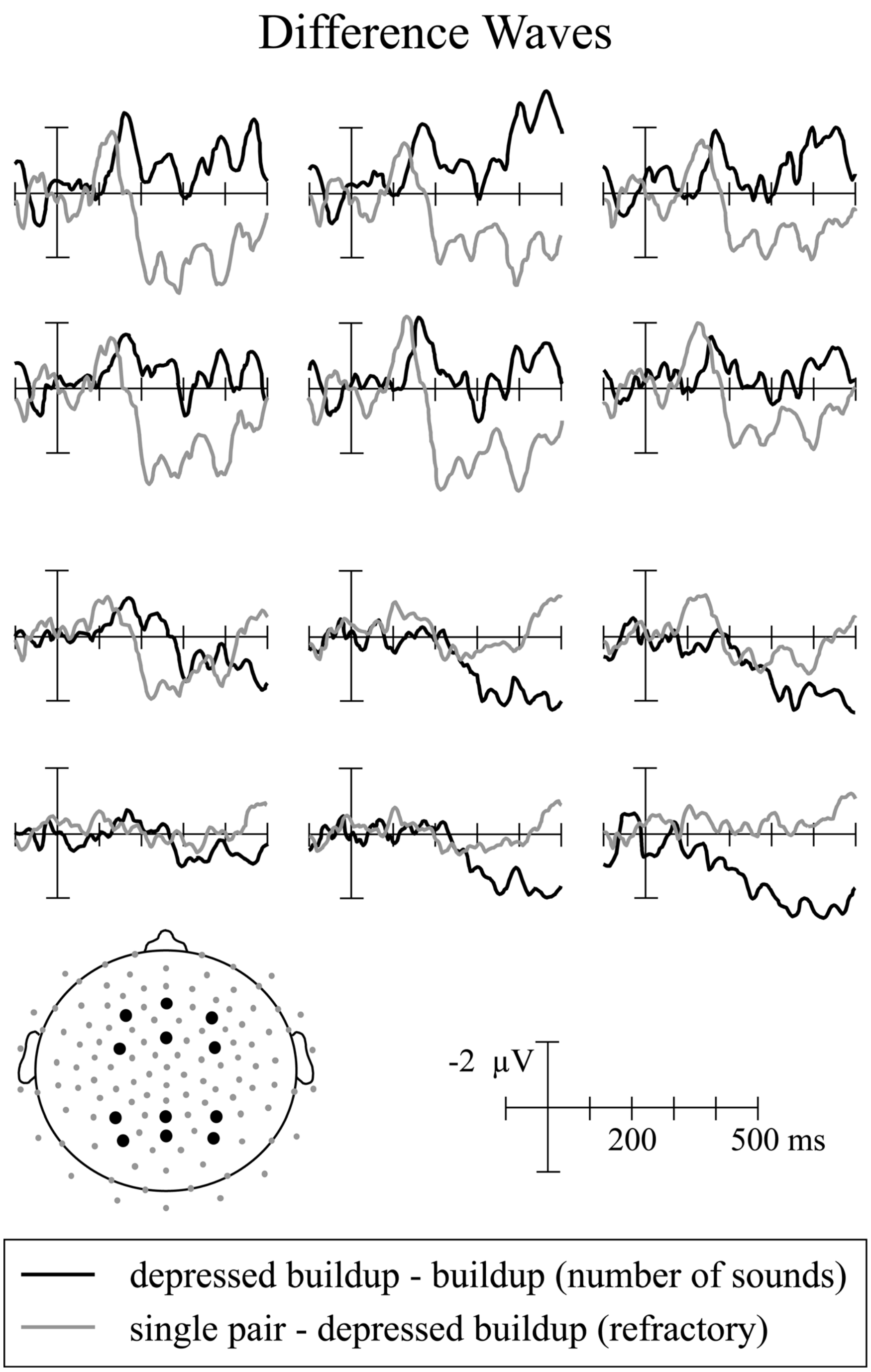

FIG. 4. Difference waves for two pairs of conditions. Perception of the lag sound as a separate source is indexed by the difference between trials on which listeners reported hearing the lag sound following the depressed buildup context and not hearing the lag sound following the buildup context (two sound sources – one sound source, black line). Refractory effects are indexed by the difference between trials on which listeners reported hearing two sound sources on the single-pair and depressed buildup trials (single pair minus preceding context, gray line). Waveforms are shown for the 12 recording sites depicted on the electrode map.

responses associated with the buildup of echo threshold, ERPs elicited by physically identical click pairs in conditions referred to as buildup and depressed buildup were compared. Both of these conditions included seven repetitions of the same lead–lag click pair (i.e., the buildup train) followed by an identical test pair. The depressed buildup trials began with an additional train of five single-source clicks presented from the leading loudspeaker only, prior to the buildup train. The initial series of single-source clicks lowered echo threshold (Freyman and Keen, 2006; Keen and Freyman, 2009) and continued to exert an influence over echo thresholds even after the buildup train had been presented. The reduction in echo threshold for the depressed buildup condition relative to the buildup condition provided an opportunity to study the physiology associated with the buildup of echo threshold without the confound of refractory effects because the stimuli presented in the last several seconds before the test pairs were identical for the buildup and depressed buildup trials. It was possible to identify a single lead–lag delay for each participant that was reported to be one sound source following the buildup context and two sound sources following the depressed buildup context.

Differences in ERPs elicited by physically identical stimuli following the buildup and depressed buildup context were evident in two distinct time ranges. The earliest difference was an anterior and medial negativity 100–250 ms after sound onset when the test pair was presented following the depressed buildup context and listeners reported hearing the lag sound as a separate source compared to following the buildup context when listeners reported not hearing a sound from the lag side. The negative difference was also evident in response to test pairs near depressed buildup echo threshold when listeners reported hearing compared to not hearing the lag sound. This ERP effect is similar in timing, distribution, and amplitude to that observed when participants report hearing two sound sources compared to one sound source for clicks near single-pair echo threshold in the precedence effect (Sanders *et al.*, 2008), when a mistuned harmonic which is not sufficient to produce perception of two simultaneous pitches is presented from a distinct location (McDonald and Alain, 2005), and various acoustic manipulations that result in listeners hearing two simultaneous pitches compared to a single pitch (Alain *et al.*, 2001; Alain *et al.*, 2002; Alain and Izenbert, 2003; Dyson and Alain, 2004; Hautus and Johnson, 2005; Hiraumi *et al.*, 2005; Chait *et al.*, 2006). This difference in ERPs observed when listeners perceive two compared to one auditory event, regardless of the feature that separates the two events (location or pitch) and whether perception is driven by physical differences in the stimuli or not, has been termed the ORN.

The disappearance of the ORN in the buildup of echo threshold provides objective electrophysiological data to augment the behavioral finding of raised echo thresholds following trains of repeated clicks. The modulation of echo threshold by previous experience has been interpreted as suggesting that listeners form models of an acoustic space when exposed to sound and make predictions about how direct and reverberant energy should behave in such a space (Freyman and Keen, 2006; Keen and Freyman, 2009). To the extent that this hypothesis is correct, the electrophysiological data from the current study indicate that those models of acoustic space based on recent auditory experience also modulate perceptual processing as indexed by the ORN.

In addition to the ORN, click pairs presented in contexts that led participants to report hearing two compared to one sound source elicited a later posterior positivity. The same pattern of data was reported in studies that compared ERPs for sounds perceived as two compared to one simultaneously presented pitch (Alain *et al.*, 2001; Alain and Izenbert, 2003; Dyson *et al.*, 2005). In these studies, the ORN was observed regardless of the task in which participants were engaged. In contrast, when attentional resources were fully occupied by a difficult visual n-back task, there was no evidence of the later positivity. In the current study, listeners were always asked to attend to the auditory stimuli and to make judgments about what they heard. This approach has the advantage of allowing for ERP analysis based on what individuals report hearing on each trial. However, it also means no direct attention manipulation can be made. The parallels in the current and previous studies predict that the ORN would be present regardless of task, consistent with the idea that it indexes automatic auditory object processing. In contrast, attention may be required in order for the later positivity to be evident, suggesting that this waveform deflection is more closely related to conscious perception of multiple sound sources or awareness of the second source when the task is to report hearing one or two sounds.

The timing and distribution of the posterior positivity further suggest that it may be related to the previously described P3 component (Squires *et al.*, 1975; Roth *et al.*, 1976; Donald and Little, 1981; Vogel *et al.*, 1998). Specifically, a P3 is typically only observed when participants are highly confident in their behavioral responses as measured by explicit confidence ratings or highly consistent performance (Squires *et al.*, 1975). In studies that manipulated the number of pitches participants reported by mistuning a harmonic (Alain *et al.*, 2001; Dyson *et al.*, 2005; McDonald and Alain, 2005), the amount of mistuning was specifically selected based on consistent behavioral responses and the posterior positivity was evident. In the current study, the context manipulation was powerful enough that participants consistently (around 80%) reported hearing the lag sound as a separate source in the single pair and depressed buildup conditions and not hearing the lag sound in the buildup condition, and the posterior positivity was evident. Further, there were individual differences in the size of the context effect on behavioral responses and these differences were correlated with the amplitude of the late posterior effect. The same high level of consistency in behavioral responses was not observed for test clicks following the depressed buildup context at SOAs selected to be near each participant's echo threshold for that condition. When listeners reported hearing two sound sources, click pairs elicited an ORN but no late posterior positivity. This result parallels that reported previously (Sanders *et al.*, 2008); single pairs near echo threshold elicit an ORN, but the amplitude differences between 250 and 500 ms are not statistically significant when listeners report hearing the lag sound on around half of the trials.

Auditory evoked potentials elicited by sounds that have been recently preceded by other similar sounds are smaller than those elicited by sounds presented in isolation (Woods and Courchesne, 1986; Coch *et al.*, 2005). These refractory effects last several seconds and are evident in the depressed buildup condition in which the test pair was preceded by seven repetitions of the same sound compared to the single-pair condition in which the test pair was presented in isolation. In the current study, the earliest refractory effects evident in the comparison of the depressed buildup and single pair conditions (80–130 ms) appear somewhat earlier in the waveform than the effects of perceiving the lag sound as a separate source evident in the comparison of the depressed buildup and buildup conditions (100–250 ms). The different but overlapping timing of these effects may reflect a distinction in the levels of auditory processing modulated by stimulus and top-down factors.

In summary, recent auditory experience that increases echo threshold and induces listeners to subjectively report hearing a single auditory event more frequently, the buildup of the precedence effect, also modulates auditory perception. Parallels with previous research suggest the processing differences indexed by the ORN are similar regardless of whether the perception of two sound sources is defined by location or pitch and whether the differences in perception are driven by subtle manipulations of the stimuli or complex models of room acoustics based on recent experience. Click pairs in contexts that led participants to report hearing two sound sources compared to one also elicited a later posterior positivity, which has been shown to be modulated by auditory attention in other studies. These findings are consistent with the idea that people use experience with sound to alter their acoustic model of a spatial environment. These models include predictions about how reflected energy should behave relative to direct energy from the sound source. The acoustic models and predictions not only influence listeners' subjective experience and reports of hearing the delayed energy as originating from a distinct location but also modulate neuroperceptual processing as indexed by ERPs.

ACKNOWLEDGMENTS

The research reported here was partially funded by NIH R01 DC01625 awarded to REK and RLF. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIDCD or NIH.

Alain, C., Arnott, S., and Picton, T. (**2001**). "Bottom-up and top-down influences on auditory scene analysis: Evidence from event-related brain potentials," J. Exp. Psychol. Hum. Percept. Perform. **27**, 1072–1089.

Alain, C., and Izenbert, A. (**2003**). "Effects of attentional load on auditory scene analysis," J. Cogn. Neurosci. **15**, 1063–1073.

Alain, C., Schuler, B., and McDonald, K. (**2002**). "Neural activity associated with distinguishing concurrent auditory objects," J. Acoust. Soc. Am. **111**, 990–995.

Bess, J., and Ruhm, H. (**1972**). "Recovery cycle of the acoustically evoked potential," J. Speech Hear. Res. **15**, 507–517.

Blauert, J. (**1997**). *Spatial Hearing: The Psychophysics of Human Sound Localization* (MIT Press, Cambridge, MA), pp. 1–480.

Blauert, J., and Col, J. (**1992**). "A study of temporal effects in spatial hearing," in *Auditory Physiology and Perception*, edited by Y. Cazal, L. Demany, and K. Korner (Pergamon, Oxford), pp. 531–538.

Budd, T., Barry, R., Gordon, E., Rennie, C., and Michie, P. (**1998**). "Decrement of the N1 auditory event-related potential with stimulus repetition: Habituation vs refractoriness," Int. J. Psychophysiol. **31**, 51–68.

Chait, M., Poeppel, D., and Simon, J. (**2006**). "Neural response correlates of detection of monaurally and binaurally created pitches in humans," Cereb. Cortex **16**, 835–848.

Clifton, R. K. (**1987**). "Breakdown of echo suppression in the precedence effect," J. Acoust. Soc. Am. **82**, 1834–1835.

Clifton, R. K., and Freyman, R. (**1989**). "Effect of click rate and delay on breakdown of the precedence effect," Percept. Psychophys. **46**, 139–145.

Clifton, R. K., Freyman, R., Litovsky, R., and McCall, D. (**1994**). "Listeners' expectations about echoes can raise or lower echo thresholds," J. Acoust. Soc. Am. **95**, 1525–1533.

Clifton, R. K., Freyman, R., and Meo, J. (**2002**). "What the precedence effect tells us about room acoustics," Percept. Psychophys. **64**, 180–188.

Coch, D., Skendzel, W., and Neville, H. (**2005**). "Auditory and visual refractory period effects in children and adults: An ERP study," Clin. Neurophysiol. **116**, 2184–2203.

Damaschke, J., Riedel, H., and Kollmeier, G. (**2005**). "Neural correlates of the precedence effect in auditory evoked potentials," Hear. Res. **205**, 157–171.

Dimitrijevic, A., and Stapells, D. (**2006**). "Human electrophysiological examination of buildup of the precedence effect," NeuroReport **17**, 1133–1137.

Donald, M., and Little, R. (**1981**). "The analysis of stimulus probability inside and outside the focus of attention as reflected by the auditory N1 and P3 components," Can. J. Psychol. **35**, 175–187.

Dyson, B., and Alain, C. (**2004**). "Representation of concurrent acoustic objects in primary auditory cortex," J. Acoust. Soc. Am. **115**, 280–288.

Dyson, B., Alain, C., and He, Y. (**2005**). "Effects of visual attentional load on low-level auditory scene analysis," Cogn. Affect. Behav. Neurosci. **5**, 319–338.

Erwin, R., and Buchwald, J. (**1986**). "Midlatency auditory evoked responses: Differential recovery cycle characteristics," Electroenceph. Clin. Neurophysiol. **64**, 417–423.

Ferree, T., Luu, P., Russell, G., and Tucker, D. (**2001**). "Scalp electrode impedance, infection risk, and EEG data quality," Clin. Neurophysiol. **112**, 536–544.

Fitzpatrick, D., Kuwada, S., Kim, D., Parham, K., and Batra, R. (**1999**). "Responses of neurons to click-pairs as simulated echoes: Auditory nerve to auditory cortex," J. Acoust. Soc. Am. **106**, 3460–3472.

Freyman, R., Clifton, R. K., and Litovsky, R. (**1991**). "Dynamic processes in the precedence effect," J. Acoust. Soc. Am. **90**, 874–884.

Freyman, R., and Keen, R. (**2006**). "Constructing and disrupting listeners' models of auditory space," J. Acoust. Soc. Am. **106**, 3957–3965.

Gardner, M. (**1968**). "Historical background of the Haas and/or precedence effect," J. Acoust. Soc. Am. **43**, 1243–1248.

Grantham, D. (**1996**). "Left-right asymmetry in the buildup of echo suppression in normal-hearing adults," J. Acoust. Soc. Am. **99**, 1118–1123.

Hautus, M., and Johnson, B. (**2005**). "Object-related brain potentials associated with the perceptual segregation of a dichotically embedded pitch," J. Acoust. Soc. Am. **117**, 275–280.

Hiraumi, H., Nagamine, T., Morita, T., Naito, Y., Fukuyama, H., and Ito, J. (**2005**). "Right hemispheric predominance in the segregation of mistuned partials," Europ. J. Neurosci. **22**, 1821–1824.

Johnson, B., Hautus, M., and Clapp, W. (**2003**). "Neural activity associated with binaural processes for the perceptual segregation of pitch," Clin. Neurophysiol. **114**, 2245–2250.

Keen, R., and Freyman, R. (**2009**). "Release and re-buildup of listeners' models of auditory space," J. Acoust. Soc. Am. **125**, 3243–3252.

Liebenthal, E., and Pratt, H. (**1999**). "Human auditory cortex electrophysiological correlates of the precedence effect: Binaural echo lateralization suppression," J. Acoust. Soc. Am. **106**, 291–303.

Litovsky, R., Colburn, H., Yost, W., and Guzman, S. (**1999**). "The precedence effect," J. Acoust. Soc. Am. **106**, 1633–1654.

McDonald, K., and Alain, C. (**2005**). "Contribution of harmonicity and location to auditory object formation in free field: Evidence from event-related brain potentials," J. Acoust. Soc. Am. **118**, 1593–1604.

Näätänen, R., Pakarinen, S., Rinne, T., and Takegata, R. (**2004**). "The mismatch negativity (MMN): Towards the optimal paradigm," Clin. Neurophysiol. **115**, 140–144.

Paavilainen, P., Karlsson, M., Reinikainen, K., and Näätänen, R. (**1989**). "Mismatch negativity to change in spatial location of an auditory stimulus," Electroenceph. Clin. Neurophysiol. **73**, 129–141.

Pekkonen, E., Rinne, T., and Näätänen, R. (**1995**). "Variability and replicability of the mismatch negativity," Electroenceph. Clin. Neurophysiol. **96**, 546–554.

Picton, T., Alain, C., Otten, L., Ritter, W., and Achim, A. (**2000**). "Mismatch negativity: Different water in the same river," Audiol. Neuro-Otol. **5**, 111–139.

Roth, W., Ford, J., Lewis, S., and Kopell, B. (**1976**). "Effects of stimulus probability and task-relevance on event-related potentials," Psychophysiology **13**, 311–317.

Saberi, K., and Perrott, D. (**1990**). "Lateralization thresholds obtained under conditions in which the precedence effect is assumed to operate," J. Acoust. Soc. Am. **87**, 1732–1737.

Sams, M., Alho, K., and Näätänen, R. (**1984**). "Short-term habituations and dishabituation of the mismatch negativity of the ERP," Psychophysiology **21**, 434–441.

Sanders, L., Joh, A., Keen, R., and Freyman, R. (**2008**). "One sound or two? Object-related negativity indexes echo perception," Percept. Psychophys. **70**, 1558–1570.

Spierer, L., Bourquin, N., Tardif, E., Murray, M., and Clarke, S. (**2009**). "Right hemispheric dominance for echo suppression," Neuropsychologia **47**, 465–472.

Squires, K., Squires, N., and Hillyard, S. (**1975**). "Decision-related cortical potentials during an auditory signal detection task with cued observation intervals," J. Exp. Psychol. Hum. Percept. Perform. **1**, 268–279.

Vogel, E., Luck, S., and Shapiro, K. (**1998**). "Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink," J. Exp. Psychol. Hum. Percept. Perform. **24**, 1656–1674.

Wallach, H., Newman, E., and Rosenzweig, M. (**1949**). "The precedence effect in sound localization," Am. J. Psychol. **62**, 315–336.

Woods, D., and Courchesne, E. (**1986**). "The recovery functions of auditory event-related potentials during split-second discriminations," Electroencephalogr. Clin. Neurophysiol. **65**, 304–315.

Yost, W., and Guzman, S. (**1996**). "Auditory processing of sound sources: Is there an echo in here?" Curr. Dir. Psychol. Sci. **5**, 125–131.

Zurek, P. (**1980**). "The precedence effect and its possible role in the avoidance of interaural ambiguities," J. Acoust. Soc. Am. **67**, 952–964.