

Temporal characteristics of contextual effects in sound localization

Luboš Hládek,^{a)} Beáta Tomoriová, and Norbert Kopčo^{b)}

Institute of Computer Science, Faculty of Science, P. J. Šafárik University, Jesenná 5, Košice, 04001, Slovakia

(Received 30 November 2016; revised 6 October 2017; accepted 6 November 2017; published online 28 November 2017)

Two experiments examined plasticity induced by context in a simple target localization task. The context was represented by interleaved localization trials with the target preceded by a distractor. In a previous study, the context induced large response shifts when the target and distractor stimuli were identical 2-ms-noise clicks [Kopčo, Best, and Shinn-Cunningham (2007). *J. Acoust. Soc. Am.* **121**, 420–432]. Here, the temporal characteristics of the contextual effect were examined for the same stimuli. Experiment 1 manipulated the context presentation rate and the distractor-target inter-stimulus interval (ISI). Experiment 2 manipulated the temporal structure of the context stimulus, replacing the one-click distractor either by a distractor consisting of eight sequentially presented clicks or by a noise burst with total energy and duration identical to the eight-click distractor. In experiment 1, the contextual shift size increased with increasing context rate while being largely independent of ISI. In experiment 2, the eight-click-distractor induced a stronger shift than the one-click-distractor context, while the noise-distractor context induced a very small shift. These results suggest that contextual plasticity is an adaptation driven both by low-level factors like spatiotemporal context distribution and higher-level factors like perceptual similarity between the stimuli, possibly related to precedence buildup. © 2017 Acoustical Society of America.

<https://doi.org/10.1121/1.5012746>

[GCS]

Pages: 3288–3296

I. INTRODUCTION

Acoustic scenes often contain multiple sources emitting sounds in an arbitrary temporal sequence. The listener's ability to localize a sound of interest in such scenes is influenced by many factors, including the temporal distribution of the stimuli (Simpson *et al.*, 2007) and their perceptual similarity (Best *et al.*, 2007). However, which factors are the most important is not well understood, especially when the sounds are presented without temporal overlap, i.e., when non-energetic factors dominate how the preceding stimuli influence target localization.

The primary goal of the current study was to investigate a recently reported form of spatial auditory plasticity, called *contextual plasticity* (CP; Kopčo *et al.*, 2007). In Kopčo *et al.* (2007), CP was observed as biases of up to 10° in localization of single-click targets, induced by *contextual* stimuli preceding the targets on the time scale of seconds to tens of seconds. The contextual stimuli were distractor-target click pairs, randomly *interleaved* with the single-click target-alone trials. Each distractor-target click pair consisted of clicks identical to the single-click target. The distractor click location was fixed within a block while the target location varied, and the distractor preceded the target by up to 400 ms.

Several previous studies have shown that target localization can be influenced by stimuli preceding the target on the

time scale of seconds. These studies examined various phenomena and their underlying mechanisms, including adaptation of the neural representation due to prolonged exposure to preceding stimuli (Carlile *et al.*, 2001; Dahmen *et al.*, 2010; Furukawa, 2005; Kashino and Nishida, 1998; Phillips and Hall, 2005), attentional cuing (Maddox *et al.*, 2014; Maier *et al.*, 2010; Sach *et al.*, 2000), buildup of the precedence effect (Freyman *et al.*, 1991), auditory grouping (Best *et al.*, 2007; Kopčo *et al.*, 2017), spatial memory (Martin *et al.*, 2011), and attentional effects utilizing *a priori* information about the position of the target and maskers (Kopčo *et al.*, 2010). The current study examines which of these phenomena and their underlying mechanisms might be related to CP. It tests the prediction that CP is related to the spatial distribution-sensitive adaptation reported, e.g., in Dahmen *et al.* (2010), and to precedence buildup (Freyman *et al.*, 1991).

Dahmen *et al.* (2010) studied how lateralization of a brief noise stimulus is systematically affected by the statistics of the lateral distribution of the preceding adaptor stimulus lateralized using interaural level differences (ILDs). They observed biases in responses due to a change in the mean and/or the variance of the distribution in the preceding adaptor sound, consistent with the idea that the auditory representation adapts to the distribution statistics to optimally encode the currently presented stimuli. It is possible that CP reflects a similar neural adaptation mechanism sensitive to the distribution of the previously presented sounds, as the distractor click location changed between blocks of Kopčo *et al.* (2007).

In precedence buildup studies, a repeated presentation of lead-lag click pairs from two fixed locations with a short

^{a)}Current address: Medical Research Council/Chief Scientist Office Institute of Hearing Research—Scottish Section, Glasgow Royal Infirmary, 10-16 Alexandra Parade, Glasgow, G31 2ER, UK.

^{b)}Electronic mail: norbert.kopco@upjs.sk

delay between the lead and lag stimuli results in an increased suppression of the spatial sensitivity to the lag click in comparison to a single lead-lag click pair (e.g., Freyman *et al.*, 1991; Djelani and Blauert, 2001). The precedence buildup paradigm differs from the paradigm of Kopčo *et al.* (2007) in several important ways. For instance, there is no lead stimulus in the CP single-click target, while the precedence buildup shift is assessed on a lead-lag click pair. Despite that, it is possible that the repeated presentation of the contextual distractor-target click pairs in CP studies results in a precedence-buildup-like adaptation in processing of the subsequent single-click targets, causing the observed target localization biases.

This paper is the third in a series that examines CP. The previous studies showed that the effect is robust, observed in both anechoic and reverberant rooms (Kopčo *et al.*, 2007), and not sensitive to the response method or the availability of vision (Kopčo *et al.*, 2015). Finally, the effect was shown to be unaffected by the contextual task difficulty, as it persists even when the contextual task was made very easy by playing the distractor after the target on the contextual trials (Kopčo *et al.*, 2015). The results of the studies of Dahmen *et al.* (2010) and Freyman *et al.* (1991) suggested that the spatiotemporal distribution of stimuli may play a role in CP. However, that distribution was not manipulated in the previous CP investigations. Therefore, the current study aims to systematically examine it by manipulating the rate at which the distractor-target vs the target-alone trials are presented. Also, the influences of other temporal properties of the experimental design are manipulated. Finally, in the previous CP studies the distractor and target stimuli always consisted of one or more 2-ms clicks. Here, we tested whether CP is sensitive to perceptual similarity between the stimuli by using distractors identical to the target as well as dissimilar ones.

Two experiments were performed to explore the effect of the temporal characteristics of the context and the distractor-target stimulus similarity on CP. Experiment 1 systematically varied two temporal parameters: the *context rate*, i.e., the ratio of the number of distractor-target contextual trials vs the number of target-alone trials, and the inter-stimulus interval (ISI) between the distractor and target stimuli in contextual trials. It was hypothesized that increasing the context rate would result in a larger contextual bias, as more distractor clicks would be presented, shifting the stimulus distribution mean toward the fixed distractor location. The sensitivity of CP to changes in the ISI was expected to be small, as the overall stimulus distribution does not vary with varying ISI.

In experiment 2, the change in the temporal profile of the context was introduced by replacing the one-click distractor by a distractor consisting of a sequence of eight clicks. This manipulation resulted in an eightfold increase in the number of clicks coming from the distractor location, a change that shifted the stimulus distribution mean toward the distractor more dramatically than the context rate manipulations in experiment 1. Thus, a larger contextual bias was expected in experiment 2. Experiment 2 also tested whether CP is sensitive to high-level factors like perceptual similarity

or grouping (Best *et al.*, 2007) between the distractor and target stimuli or whether the spatiotemporal stimulus energy distribution is the main important factor. To test this, a condition was included in which the distractor was a single continuous noise burst with the total duration, spectrum, and power equal to the eight-click distractor. This noise distractor was expected to produce the same amount of CP, if the high-level factors like similarity between stimuli were not important.

This study also analyzes the effect of immediately preceding distractor on target localization on the contextual trials. This analysis is performed to confirm that the effect of distractor operates on a much shorter time scale than CP (up to hundreds of milliseconds; Kopčo *et al.*, 2007). Thus, it is expected not to be influenced by the context rate manipulations in experiment 1 while being influenced by the distractor-target similarity manipulations in experiment 2.

II. METHODS

Two experiments were performed using setup, stimuli, and procedures similar to the previous CP studies (Kopčo *et al.*, 2007; Kopčo *et al.*, 2015; Kopčo *et al.*, 2017).

A. Subjects

Eleven subjects (ten males and one female), participated in experiment 1. All subjects had normal hearing (by self-report), with ages ranging from 23 to 35 yr. Authors B.T. and N.K. participated in the study. Most of the other participants had no previous experience with psychoacoustic experiments. The data of one subject were discarded due to a technical error. Eight subjects from experiment 1 also participated in experiment 2. All listeners gave informed consent. All experimental procedures were approved by the institutional ethical review board.

B. Setup and listening environment

Listeners were seated in a dimly lit, single-walled, sound-proof booth [broadband $T_{60} = 415$ ms, background noise 32 dBA, internal dimensions 3 m (h) \times 2 m \times 3 m] with their head supported by a head rest. The presentation system consisted of a personal computer, RME Fireface 400 audio processor (RME Audio, Haimhausen, Germany), Crown D-75 A amplifier (Crown Audio, Elkhart, IN), and a custom-made switching relay routing the two-channel audio output to different combinations of the loudspeakers. Nine loudspeakers (Bose Acoustimass cube speakers, Bose, Framingham, MA) were positioned on an arc with radius of 1.1 m at the level of listener's ears, spanning 90° around the listener (angular separation of 11.25°). The listener sat in the center of the arc and faced either the left- or rightmost speaker (see Fig. 1). The loudspeakers were matched so that the maximum deviation of the response in 1/3-octave frequency bands from 0.5 to 10 kHz did not exceed 3 dB from the across-speaker average for any of the speakers. The subjects responded by pointing a custom-made handheld video-tracked wand in the perceived target direction. Light-emitting diodes (LEDs) were mounted on the pointing wand,

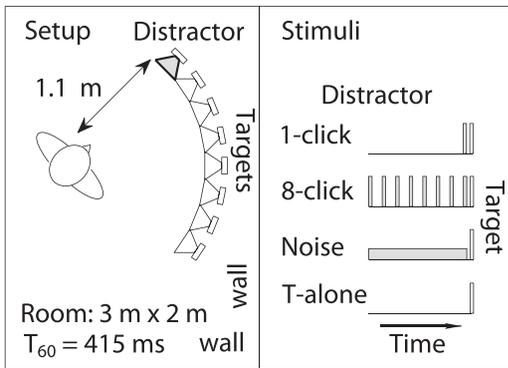


FIG. 1. Diagram of experimental setup and stimuli. (Left) Schematic of the orientation and location of the listener and the loudspeakers in the experimental booth. The loudspeaker depicted by gray color is the distractor (always frontal). White loudspeakers are targets. For half of the runs, the setup was mirror-flipped so that the subject faced the rightmost speaker (which served as the distractor) and the targets were on the left. (Right) Temporal profile of the distractor (gray) and target (white) stimuli. ISI is the interval between the offset of the distractor and onset of the target.

the loudspeakers, and on top of the listener's head. A camera was fixed ~ 2 m above the listener. The listener's responses were extracted from images showing the position of the LEDs, captured when the listener pressed a response button mounted on the wand. The loudspeakers were hidden behind an acoustically transparent cloth to minimize information about possible target locations. During experimental runs, subjects were instructed to keep their eyes closed.

C. Stimuli and procedures

Both experiments were divided into four sessions, each consisting of multiple runs of trials. A trial consisted of the presentation of the target stimulus and a subject's response. The target was presented either *alone* in target-alone trials or preceded by a *distractor* in distractor-target trials (see Fig. 1). The target could be presented from one of the seven loudspeakers at azimuths of 11° – 79° , where the azimuth was defined as laterality with respect to the frontal speaker, positive for the right quadrant if the loudspeakers were on the subject's right and for the left quadrant if the loudspeakers were on the left. The distractor was always presented from the frontal loudspeaker (azimuth 0°). In experiment 1, both the target and the distractor stimuli were identical 2-ms frozen noise bursts (a "click" sound) presented at 67 dBA. In experiment 2, other types of distractors were also tested, and are described in the last paragraph of this section. In the target-alone trials, the distractor was replaced by silence so that the trial duration was comparable to the distractor trial duration.

The subjects were instructed to keep their eyes closed during the experimental runs and to localize the target by pointing to it while disregarding the distractor. The subjects were encouraged to take a short break after each run and were free to have their eyes open during the break. No feedback was provided.

Each session included several *experimental runs* and one *baseline run*. The *experimental runs* consisted of three parts: pre-adaptation (14 trials), adaptation (140 trials), and post-adaptation (35 trials in experiment 1 and 21 trials in

experiment 2). In the pre- and post-adaptation, only target-alone trials were included. During the adaptation, the target-alone trials were randomly interleaved with the distractor-target trials representing the "context" whose characteristics were experimentally manipulated and fixed within a run. Only the adaptation portion of each experimental run (and the corresponding portion of baseline runs) was used in the analysis of contextual biases. The pre- and post-adaptation portions were used to examine the buildup and decay of CP. The *baseline runs* consisted of target-alone trials and had the same overall number of trials as the experimental runs. The order of the runs within a session was random and differed from session to session. The subjects changed their orientation from facing the leftmost speaker to facing the rightmost speaker (or vice versa) after each run. One session took ~ 2.5 h in experiment 1 and 1.5 h in experiment 2.

In experiment 1, the distractor was a single click identical to the target (one-click). Experiment 1 manipulated two parameters, the context rate, i.e., the ratio of the distractor-target contextual trials to the target-alone trials, and the distractor-target ISI. The experimental conditions included all combinations of the context rates of 50% (50% contextual trials, 50% target-alone trials) and 75% (75% contextual trials, 25% target-alone trials) and ISIs of 23, 98, and 398 ms. In addition, a condition with 90% contextual trials and ISI of 98 ms was included. The runs with higher context rates were repeated multiple times to obtain a constant number of target-alone trials for each condition. Specifically, each session included one 50%-rate run per ISI, two 75%-rate runs per ISI, and five 90%-rate runs for the 98-ms ISI. In total, 1 experimental session contained 14 experimental runs and 1 baseline run.

Experiment 2 manipulated the number of clicks in the distractor and the similarity between the distractor and the target in the contextual trials (see right-hand panel of Fig. 1). The distractor was either a single click like in experiment 1 (one-click condition), a train of eight clicks identical to the target and presented at a constant rate with peak-to-peak period of 125 ms (eight-click condition), or a single white noise burst with duration and root-mean-square (RMS) power identical to the eight-click stimulus (noise condition). As a consequence of the RMS equalization, the instantaneous level of the noise stimulus was 17.3 dB lower than that of the click sound. The context rate was fixed at 75% and the distractor-target ISI was always 23 ms for all conditions. Thus, the one-click condition in experiment 2 was identical to the (ISI 23 ms, context rate 75%) condition from experiment 1. Notably, introduction of the eight-click condition resulted in an eightfold increase in the overall rate at which clicks were presented from the distractor location. Thus, even though the context rate was fixed at 75%, the rate at which distractor clicks were presented was much higher, on average six distractor-clicks per trial. A new measure, the *effective context rate*, was introduced to express the rate of distractor click presentation. This rate gives values identical to the context rate for experiment 1, but for the eight-click condition of experiment 2 its value is 600%, capturing the rate at which the distractor clicks were presented. One

experiment 2 session consisted of six experimental runs (two runs per distractor type) and one baseline run.

D. Data analysis

The images captured during each response in the experiments were analyzed (online in experiment 1 and offline in experiment 2) to extract the response angle. The angle was defined as the horizontal angle subtended by the wand and the frontal loudspeaker at the center of the loudspeaker arc. A response was discarded if the angle could not be extracted due to technical problems associated with identification of the pointer coordinates (1.5% of trials were discarded in experiment 1, 0.01% in experiment 2). The mean responses were computed separately for each combination of condition, target location, and trial type (target-alone or distractor-target). To parallel the analysis of the previous CP studies, data were collapsed across the two listener orientations and repetitions, and analyzed as if the subject was always facing the leftmost loudspeaker. All reported statistical analyses were performed as multi-way repeated measures analyses of variance (ANOVAs), using CLEAVE software (Herron, 2005). The reported p values were corrected for potential violations of sphericity using the Greenhouse-Geisser epsilon.

III. RESULTS

The following analysis of the CP effects focuses on the biases in mean localization responses relative to the baseline. First, the data were averaged across time within a run and evaluated as a function of the target location. Then, the temporal profile of the contextual bias within a run was analyzed. Finally, the effect of the immediately preceding distractor on the target localization in the contextual trials is presented.

A. Contextual effects in mean responses

1. Effect of frequency of occurrence of the context and of ISI

Experiment 1 examined how the contextual effect is influenced by the temporal distribution of the stimuli. The context rate and the ISI were manipulated. It was expected

that the contextual effect will be primarily influenced by the context rate.

Figure 2(A) shows the localization bias in the baseline condition for the two experiments. The responses were consistently biased toward the middle of the speaker array. As a result, the response range is compressed by as much as 25%, more in experiment 1 than in experiment 2, and slightly more than reported in the previous studies (Kopčo *et al.*, 2007; Kopčo *et al.*, 2015). Also, the experiment 2 responses are overall shifted more frontally than the experiment 1 responses. These biases are mainly due to the response method, which was used here even though it was previously shown to cause such biases (Kopčo *et al.*, 2015). The method was kept to be consistent with the previous studies and because it mainly affects the absolute responses [as shown in Fig. 2(A)], not the relative measures of the CP bias with respect to baseline and the effect of immediately preceding distractor (Kopčo *et al.*, 2015).

The left-hand portion of Fig. 2(B) plots the contextual bias in experiment 1, evaluated as the shift of the target-alone localization responses from the experimental runs re. baseline [from Fig. 2(A)]. The contextual bias is plotted as a function of the target lateral angle, separately for each combination of the ISI (separate panels) and the context rate (different symbols within each panel).

The context produced a bias of 1° – 6° away from the distractor (i.e., laterally) that tended to decrease with target laterality. This bias was influenced by the context rate, increasing for the 75% rate vs the 50% rate (squares fall above triangles in each panel) but not for the 90% rate vs the 75% rate at the 98-ms ISI (crosses and squares are aligned in the middle panel). On the other hand, ISI only showed a very weak trend: at the 23-ms ISI the bias slightly decreased with increasing target lateral angle, while at the larger ISIs this dependence disappeared (the approximately linear trends shown in the rightmost panel are shallower than the ones in the leftmost panel). These results were confirmed by a three-way repeated measures ANOVA with the factors of ISI (23 ms, 98 ms, 398 ms), context rate (50%, 75%), and target laterality (seven target locations from 11° to 79°), which found a significant main effect of context rate [$F(1,9) = 25.19$; $p < 0.01$; $\eta^2 = 0.038$], and a

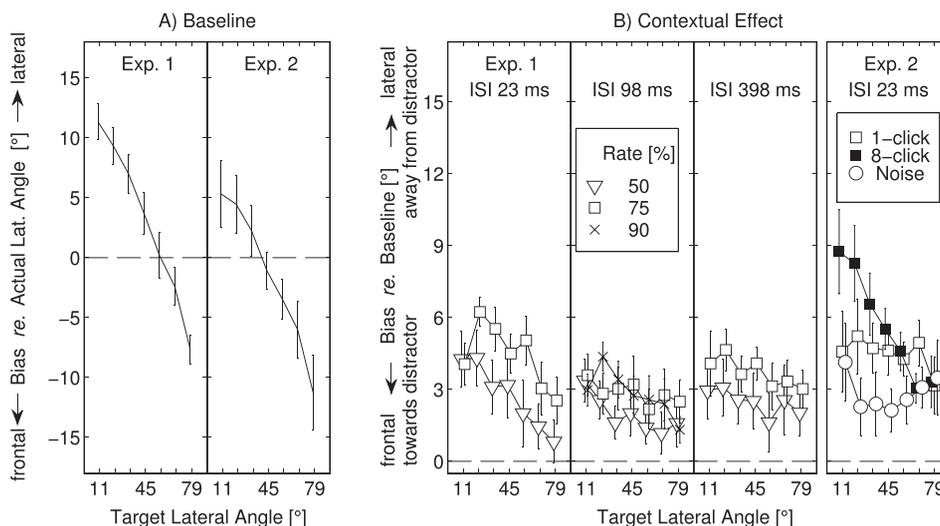


FIG. 2. Across-subject mean (\pm SEM, standard error of the mean) localization response bias in target-alone trials as a function of the target azimuth during the adaptation portion of experimental runs (or corresponding portion of baseline runs) in experiments 1 and 2. (A) Bias re. actual target location in the no-distractor baseline runs. (B) Contextual bias in experiments 1 and 2, defined as bias re. baseline condition from (A). Symbols and subplots correspond to different experimental conditions. Note that the y axis scales are different in the two panels.

trend toward an interaction of ISI \times target laterality [$F(12,108) = 2.30$; $p = 0.059$; $\eta^2 = 0.0184$].

In summary, the contextual biases were influenced by the frequency of occurrence of the contextual trials, reaching up to 6° at 75% rate and 4° at 50% rate. This result is consistent with the hypothesis that the CP is caused by a process that is sensitive to the overall spatiotemporal distribution of the click stimuli. Specifically, the higher context rate means that a click was presented more frequently from the distractor location, shifting the click distribution mean toward the distractor location (Dahmen *et al.*, 2010; Kopčo *et al.*, 2015), which in turn caused a stronger CP away from the distractor. In line with our expectations, the change in the ISI did not significantly influence the contextual effect, consistent with an interpretation that CP operates on a time scale that is much larger than the ISIs explored here.

2. Effect of stimulus spatiotemporal distribution and target-distractor similarity

The main result of experiment 1 was that the contextual bias grew when the context rate was increased from 50% to 75%, independent of the target location or ISI. Experiment 2 aimed to test the idea that the number of presented distractor clicks, not the context rate *per se*, was the factor determining the CP size. Experiment 2 introduced a distractor with eight clicks on each contextual trial, a manipulation that shifted the mean of the distribution of the clicks more dramatically toward the distractor, while keeping the context rate unchanged. To better capture the effect of this manipulation, a new measure called the *effective context rate* was introduced, defined as the average number of distractor clicks per trial (in percent). Then, keeping the original context rate at 75% and replacing the one-click distractor by an eight-click distractor results in an increase in the *effective* context rate from 75% to 600% (i.e., eightfold). The effective context rate of 600% means that, on average, six clicks were presented from the distractor location on each trial for the eight-click distractor, while the effective context rate 75% means that, on average, 0.75 clicks were presented from the distractor location on each trial for the one-click distractor. Note that the *effective* context rate is identical to the context rate for experiment 1 and only differs for experiment 2.

In addition, experiment 2 aimed to test whether the distractor has to be perceptually similar to the target in order to induce contextual bias, as is the case when perceptual grouping influences performance (Best *et al.*, 2007; Kopčo *et al.*, 2017), or whether any distractor stimulus with a similar spatiotemporal distribution of acoustic energy would suffice to induce it. Therefore, a condition with a noise distractor dissimilar from the target was added. The noise distractor was designed to induce bias similar to the eight-click distractor if the contextual bias depends on the spatiotemporal distribution of acoustic energy, not on factors like the similarity between the distractor and target stimuli. On the other hand, if the target-distractor stimulus similarity is important, then the noise distractor would induce bias smaller than the one-click distractor.

The rightmost portion of Fig. 2(B) plots the contextual bias as a function of target laterality, with the symbols representing different distractor types used in experiment 2. Note that open squares in the experiment 2 sub-panel present a replication of the same condition as the open squares for the 23-ms ISI in the experiment 1 sub-panel.

The one-click distractor induced a bias with maximum of 5° that slightly decreased with target laterality (open squares). The eight-click distractor induced a bias that reached 9° near the distractor and approximately linearly decreased to 3° for the targets far from the distractor (filled squares). The noise distractor had the weakest effect, only around 3° at all target locations. Repeated measures ANOVA with factors of target laterality and distractor type found a main effect of distractor type [$F(2,14) = 7.85$, $p < 0.01$; $\eta^2 = 0.1012$] and an interaction of distractor type and target laterality [$F(12,84) = 5.88$, $p < 0.01$; $\eta^2 = 0.0769$].

The difference between the eight-click and the noise conditions indicates that the contextual bias is caused by a mechanism that is sensitive both to the similarity between the distractor and target and to the distribution of the click stimuli in the contextual trials. The experiment 2 results are not completely consistent with the results of experiment 1. CP in experiment 1 was increased from 50% to 75% of the effective context rate but not from 75% to 90%, while in experiment 2 a large increase was observed for the 600% effective rate. A possible explanation for the disparity between the experiments is the examined range of the *effective* context rates. In experiment 1, the range was much smaller (50% vs 75% vs 90% compared to 600% in experiment 2), which might have been insufficient to make some of the effects observed in experiment 2 visible. However, other differences between the experiments (e.g., the temporal distribution of the distractor clicks) might also be important. Finally, the eight-click distractor produced shifts that were location dependent, concentrating on the target locations near the distractor. This is consistent with the trend of experiment 1 in which a slightly stronger bias was observed near the distractor. However, in experiment 1 the increase in context rate from 50% to 75% resulted in a uniform increase in the CP shift across the targets, whereas in experiment 2 the increase was limited to the targets near the distractor.

B. Temporal profile of contextual bias

This section presents an analysis of the CP buildup and decay during the experimental runs. Only experiment 2 data are presented. Experiment 1 analysis is omitted because the 23-ms condition is identical to the one-click condition of experiment 2, and the 398-ms condition is similar to data presented previously (Kopčo *et al.*, 2015).

Each row in Fig. 3 shows the temporal profile for one distractor type of experiment 2. The two columns show the profiles at locations near the distractor (pooled across azimuths of 11° – 23° , left-hand column) and far from the distractor (pooled across azimuths of 68° – 79° , right-hand column). The time course of the run is represented by sub-runs, each corresponding to seven trials. The data from the adaptation part of each run were pooled across pairs of

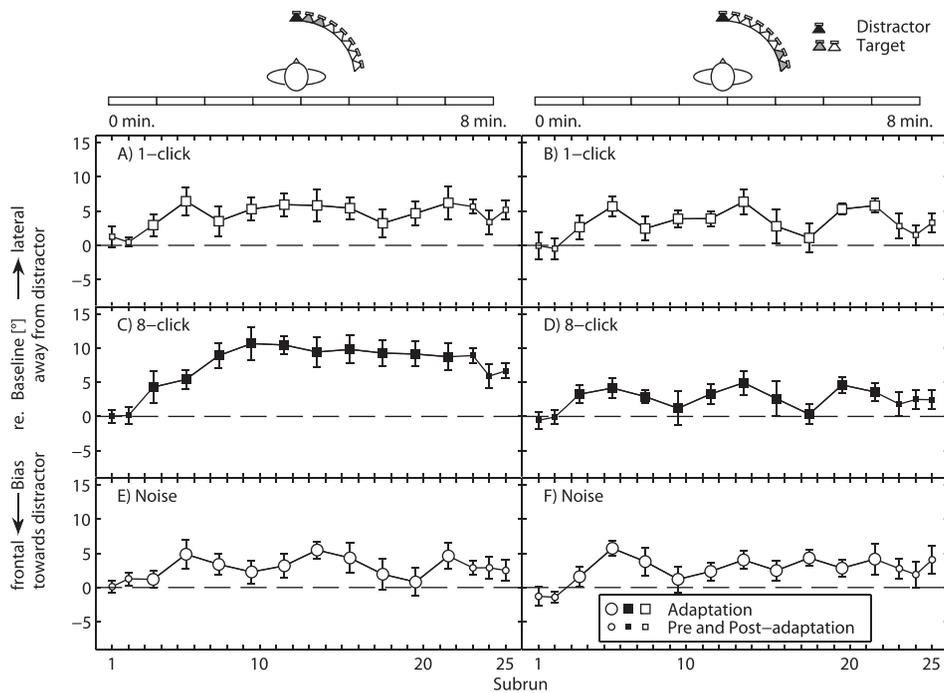


FIG. 3. Across-subject mean (\pm SEM) contextual bias as a function of subrun in target-alone trials of experiment 2. Each row shows data for one distractor type. Frontal-target data (A),(C),(E) were pooled across azimuths of 11° – 23° (see gray loudspeakers in the diagram), lateral-target data (B),(D),(F) were pooled across azimuths of 68° – 79° (see the diagram). Data in the adaptation part were pooled across two consecutive subruns.

consecutive subruns to reduce noise and are plotted using larger symbols, while the pre- and post-adaptation data are plotted using smaller symbols and separately for each sub-run. For the one-click and noise distractors at both target azimuths [Figs. 3(A), 3(B), 3(E), and 3(F)], the contextual bias had a very fast buildup, reaching the maximum by subrun number 6, i.e., just four subruns after the onset of the context in subrun number 3. On the other hand, the eight-click distractor caused a much longer buildup that continued until subrun number 10 and reached the maximum of 10° for the frontal targets [Fig. 3(C)]. The temporal profiles for the eight-click distractor and the lateral targets [Fig. 3(D)] were similar to those in Figs. 3(A), 3(B), 3(E), and 3(F). After the offset of the contextual trials [i.e., in last three data points in Figs. 3(A)–3(F)] the contextual bias decayed although it did not disappear completely.

The temporal axis along the upper edge of Fig. 3 indicates approximate timing of the runs in minutes instead of subruns. The one-click distractor data [Fig. 3(A)] show that the contextual adaptation reaches its asymptote within ~ 1 min. On the other hand, the buildup caused by the eight-click distractor took 2–3 min to reach the maximum [Fig. 3(C)]. This result allows two interpretations. One is that the adaptation rate is approximately constant in the units of degrees per subrun for these two distractor types, thus, requiring longer time to reach the larger asymptotic value for the eight-click distractor. Alternatively, it is possible that an additional/different slower adaptation mechanism is activated for the eight-click distractor.

C. Effect of preceding distractor

This section evaluates the effect of the immediately preceding distractor on target localization in the contextual trials. Several predictions were formulated for the experimental manipulations used in experiments 1 and 2. For experiment 1,

the results were expected to be unaffected by the changes in the context rate while being dependent on the distractor-target ISI since the effect of preceding distractor was expected to operate on time scales much shorter than CP. In experiment 2, the effect of distractor on target localization was expected to be reduced compared to the 1-click distractor when the distractor was a noise dissimilar from the target. When the distractor was an eight-click train, the target was expected to be processed in a stream separate from the distractor stream since the target did not fall into the timing predicted by the periodicity of the distractor click train. Specifically, the eight-click distractor had ISI of 125 ms while the distractor-target ISI was 23 ms. Thus, while the similarity of the distractor and target was expected to result in grouping of these two objects into one stream for the one-click distractor, the temporal mismatch in the eight-click distractor was expected to result in the target being processed in a separate stream, possibly reducing any bias due to the immediately preceding one-click distractor. However, as shown in Kopčo *et al.* (2017), in which very similar distractor conditions were tested, the eight-click distractor can actually cause an increase in the response bias.

The three left sub-panels of Fig. 4 show the experiment 1 biases in the target responses induced by the immediately preceding distractor. The biases were computed as the difference between the mean distractor-target trial responses and the mean target-alone trial responses in the experimental runs [shown in Fig. 2(B)]. Similar to Fig. 2(B), the biases are plotted as a function of the target lateral angle, separately for each combination of the ISI (separate panels) and the context rate (different symbols within each panel).

At the ISIs of 23 ms and 98 ms (left-hand and central sub-panel, respectively), the target localization responses were biased toward the distractor (opposite to the direction of CP). Very little bias was observed at the 398-ms ISI (right-hand sub-panel). At all ISIs, the bias peaked in the middle of the response range (reaching a maximum of 5° at

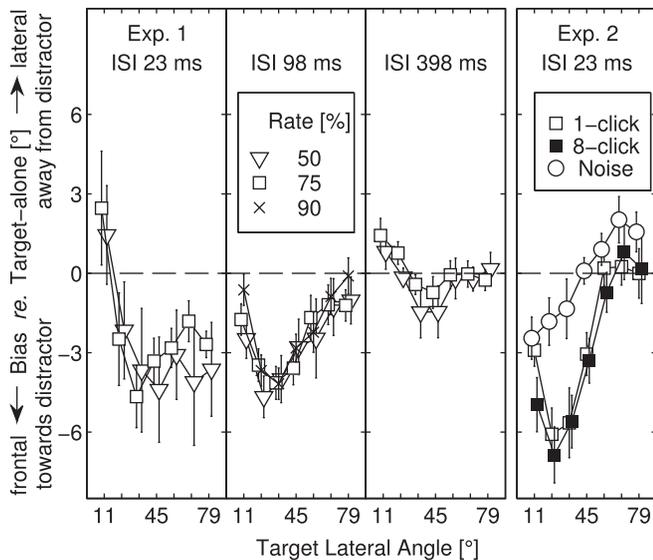


FIG. 4. Across-subject mean (\pm SEM) bias in target localization induced by the immediately preceding distractor. The bias is computed as the difference between the distractor-target trial responses and the target-alone trial responses in the experimental runs [from Fig. 2(B)], and plotted as a function of target azimuth. Three left sub-panels plot experiment 1 data separately for all combinations of context rate (symbols) and ISI (sub-panels). The rightmost panel plots experiment 2 data separately for each distractor type.

23-ms ISI) and decreased near the edges (reversing to a bias 2° away from the distractor at 398 -ms ISI). Importantly, these results were independent of the context rate (the different symbols are aligned in all three experiment 1 panels of Fig. 4). A three-way repeated measures ANOVA with the factors of target laterality, ISI, and context rate confirmed these results, showing an interaction of ISI and target laterality [$F(12,108) = 4.23$; $p < 0.05$; $\eta^2 = 0.0516$] and a significant main effect of target laterality [$F(6,54) = 6.13$, $p < 0.05$; $\eta^2 = 0.0682$].

The rightmost panel of the Fig. 4 plots the effect of preceding distractor in experiment 2 in a format similar to that used for experiment 1. The one-click distractor elicited a bias toward the distractor location that peaked at 6° for the 23° – 34° targets and decreased away from those locations (open squares). The eight-click distractor had an almost identical effect, except for a slight increase in the bias at the 11° target laterality (solid squares). The noise distractor induced very little bias that was directed toward the distractor at the 11° target location and gradually changed to a bias away from the distractor at the 79° location (circles). A two-way repeated measures ANOVA with factors of distractor type and target laterality confirmed these results, showing significant main effects of target laterality [$F(6,42) = 20.30$, $p < 0.01$; $\eta^2 = 0.3939$] and distractor type [$F(2,14) = 10.11$, $p < 0.01$; $\eta^2 = 0.114$], as well as their interaction [$F(12,84) = 2.76$, $p < 0.05$; $\eta^2 = 0.0411$].

The most important result of experiment 1 is that the effect of preceding distractor is independent of the context rate, the only factor that was significantly influencing CP. Thus, CP and the effect of preceding distractor seem to be independent of each other for the conditions explored in experiment 1. The eight-click distractor in experiment 2 had

an effect similar to the one-click distractors. Thus, it is again unlikely that this short-term effect would be interacting with CP, which was much larger for the eight-click distractor. Finally, the noise distractor had the smallest effect, as was expected due to its dissimilarity from the target. These results are discussed in more detail in Sec. IV.

IV. DISCUSSION

A. Contextual effects

CP is a phenomenon likely to affect sound localization in many everyday situations in which the sounds of interest are presented in the context of other auditory stimuli. This study examines a specific situation in which the target of interest is a brief noise click. The main finding of the current study is that the size of induced CP depends on the distractor click presentation rate when the distractor consists of clicks identical to the target.

Overall, the context induced biases away from the distractor that were large near the frontal distractor and became smaller with increasing target laterality. The size of the bias depended on the distractor click rate for the nearby targets (9° for eight-click context vs 3° – 4° for the one-click context), but not for the distant ones (1° – 3° for all contexts). Therefore, the temporal density of the distractor clicks appears to be the main factor influencing CP. On the other hand, the exact timing of the distractor clicks relative to the target clicks in the contextual trials (i.e., the ISI in experiment 1) had a negligible effect on response biases. Finally, the noise distractor context induced a very weak bias, on average less than 3° .

Temporal analysis showed that the CP buildup takes between less than 1 min and 3 min to asymptote. The duration of the buildup depended on the type of distractor (longer with eight-click than one-click distractors). While it is not clear why the temporal arrangements of the context have such a large effect on the buildup duration, one possibility is that multiple adaptive mechanisms might be activated for the eight-click distractor.

Several previous studies of short-term adaptation observed repulsive effects of a distractor or adaptor similar to the ones observed here (Carlile *et al.*, 2001; Dahmen *et al.*, 2010; Kashino and Nishida, 1998). It is likely that CP is driven by a low-level spatiotemporal-distribution-sensitive mechanism, similar to the ones discussed in those previous studies and elsewhere (Brown *et al.*, 2012; Getzmann, 2004; Grothe *et al.*, 2010; Maier *et al.*, 2010; Stange *et al.*, 2013). However, importantly, CP is dependent on the type of stimulus used as the distractor, not just on the presence of distractor energy, as the noise distractor in experiment 2 induced smaller shifts in responses than a single click distractor, even though its power was eight times larger. Several mechanisms might be responsible for such selectivity, discussed in the following.

First, it could be a low-level automatic mechanism, as mentioned in the previous paragraph, but specific to clicks or other transient stimuli. This is in line with the current result that little CP was induced by the noise distractor, which only has one onset that is smaller because of the reduced

instantaneous level of the noise. This explanation would predict that similarly small CP would result even if the target and distractor were identical but without strong transients.

Second, CP might be an effect related to precedence buildup, a phenomenon in which a repeated presentation of two consecutive clicks from two different locations leads to a progressively increasing fusion of the two clicks even at ISIs at which they are initially perceived as two objects (Djelani and Blauert, 2001; Freyman *et al.*, 1991). Typically, studies of precedence buildup estimate the echo threshold, i.e., the ISI at which the second click becomes perceptible (Keen and Freyman, 2009; Litovsky *et al.*, 1999); however, the effect is also exhibited by a lateralization bias of the perceptually fused clicks for ISIs above 23 ms (Brown and Stecker, 2013). Therefore, mechanisms similar to those underlying precedence buildup (or breakdown) might also underlie CP. One way of assessing whether the phenomena have a common underlying mechanism could be to compare the effect of different acoustic environments on them. Similar amounts of CP were previously observed in anechoic and moderately reverberant environments (Kopčo *et al.*, 2007; Kopčo *et al.*, 2017). Therefore, it would be interesting to know whether precedence buildup occurs, and is similar, for an identical setup in anechoic and reverberant environments. Also, note that while the neural nature of precedence buildup is still a subject of debate, changes in cortical ERPs indicate that higher stages of auditory processing are possibly involved (Sanders *et al.*, 2011).

A third option is that CP is a high-level effect related to streaming that builds up on relatively slow temporal scales (Boehnke and Phillips, 2005; Micheyl *et al.*, 2013; Shamma *et al.*, 2011; Weintraub *et al.*, 2014) or is due to factors like expectation (Kopčo *et al.*, 2010). This interpretation is supported by the observation that the contextual bias was reduced when the distractor clicks were replaced by noise, suggesting that the context stimuli must be perceptually similar to the target in order to induce CP. However, note that the noise and eight-click distractors differed also in other aspects, e.g., their instantaneous level, that likely resulted in differences even at a relatively low level of processing.

Further studies need to be performed to determine the mechanism underlying CP, as well as the extent to which this effect might be influencing sound localization in everyday situations, for target stimuli other than clicks, and for various distractor-target spatial configurations and different reverberant environments.

B. Effect of preceding distractor

In the current study, the immediately preceding frontal distractor had a complex effect on the perceived location of the following target click. Consistent with the previous results (Kopčo *et al.*, 2007; Kopčo *et al.*, 2015), the one-click distractor induced attractive bias with maximum of 6° at the ISI of 23 ms and at target laterality of around 30°. The bias decreased both with increasing ISI and with laterality changing away from 30°. The experiment 1 results show that these effects are largely independent of the context rate (i.e., of temporal factors on scales of tens of seconds), supporting

the hypothesis that the results are mainly caused by low-level adaptation mechanisms operating on much shorter time scales, as suggested by Kopčo *et al.* (2007).

The eight-click distractor in experiment 2 induced a bias that was similar or larger than that induced by the one-click distractor. It was originally expected that the eight-click distractor would induce streaming, reducing the effect of the distractor on the target. The observed slight increase in biases suggests that streaming has, if anything, an opposite effect in this condition. While the reason for this unexpected result is not clear, it is consistent with a recent study (Kopčo *et al.*, 2017) in which a similar manipulation also produced an increased bias for some spatial configurations while, for most configurations, it resulted in a reduction of bias and variance. A possible explanation is that there was an additional mechanism active, such as inhibition of return (Spence and Driver, 1998), canceling the beneficial effect of the streaming distractor.

The noise distractor had a small effect on the localization of the immediately following target click. However, interestingly, compared to the target-alone trials, the effect was always in the direction toward the actual target locations: frontal bias of up to 3° for frontal targets and lateral bias of up to 2° for lateral targets. This suggests that perceptual organization is an important factor here, as the dissimilarity between this distractor and the target improved the listener's ability to process the target spatial information. This might be achieved, e.g., by the distractor acting as an anchor for relative target location estimation (Kopčo *et al.*, 2010).

Overall, these results indicate that both low- and high-level mechanisms influence the listener's ability to localize a transient target stimulus presented after an *a priori* known distractor, and the effect of the distractor can also be beneficial in these conditions in which the distractor precedes the target. The results also confirm that CP and the effect of preceding distractor can be assumed to be independent of each other, in particular for the conditions explored in experiment 1.

V. SUMMARY AND CONCLUSIONS

The current study showed a complex effect of the temporal characteristics of context on sound localization, likely driven by processing at multiple stages in the auditory pathway. The main results are as follows:

- (1) CP induces biases of 3°–9° away from the distractor location that grow with increasing distractor click presentation rate and decreasing angular distance from the distractor to the target. The CP bias is also sensitive to similarity of the temporal profile of the distractor and the target, not only to the amount of acoustic energy in the distractor. Specifically, a noise distractor that has the same spectrum, overall energy, and duration as an eight-click distractor induces a much smaller CP, presumably because it contains fewer onsets and offsets. (Note that the analysis of CP in this study is limited by the chosen response method that introduced response biases even in the baseline and had a relatively large response standard deviation of up to 6°; Kopčo *et al.*, 2015.)
- (2) CP operates on the time scale of tens of seconds to minutes, and the time required to achieve maximum

adaptation varies from 1 to 3 min, depending on the rate at which the distractor clicks are presented (longer for eight-click than for one-click distractors).

- (3) Processing at multiple levels of the auditory pathway is likely to influence CP. CP is likely influenced by factors and phenomena like perceptual organization, expectation, and precedence buildup. These are assumed to originate later in the auditory pathway. The main early-stage mechanism is likely to be adaptation in the subcortical representations of space.
- (4) The immediately preceding distractor affected target localization on the contextual trials by inducing localization biases. Such effects are on the time scale of less than a second, likely due to short-term adaptation, but are also influenced by factors like the distractor-target similarity. While it is possible that CP and the effect of preceding distractor are caused by the same neural mechanisms on different time scales, the effects appear to be largely independent of each other.
- (5) These data provide a challenging test for the models of the spatial auditory perception (Bernstein and Trahiotis, 2002; Breebaart *et al.*, 2001; Dietz *et al.*, 2011), which usually do not consider processing over a wide range of time scales or multiple forms of adaptation operating at the same time. A multi-level model that considers both high- and low-level processing and various time scales might be necessary to describe the data.

ACKNOWLEDGMENTS

This work was supported by the SRDA, Contract No. APVV-0452-12, EU H2020-MSCA-RISE-2015 Grant No. 691229, and by the EU RDP projects TECHNICOM I, ITMS: 26220220182, and TECHNICOM II, ITMS2014+: 313011D232. Rudolf Andoga and Michal Barto assisted with the study. We thank Virginia Best and Barbara Shinn-Cunningham who consulted on this study, and we also thank two anonymous reviewers for insightful comments.

Bernstein, L. R., and Trahiotis, C. (2002). "Enhancing sensitivity to interaural delays at high frequencies by using 'transposed stimuli,'" *J. Acoust. Soc. Am.* **112**, 1026–1036.

Best, V., Gallun, F. J., Carlile, S., and Shinn-Cunningham, B. G. (2007). "Binaural interference and auditory grouping," *J. Acoust. Soc. Am.* **121**, 1070–1076.

Boehnke, S. E., and Phillips, D. P. (2005). "The relation between auditory temporal interval processing and sequential stream segregation examined with stimulus laterality differences," *Percept. Psychophys.* **67**, 1088–1101.

Breebaart, J., van de Par, S., and Kohlrausch, a (2001). "Binaural processing model based on contralateral inhibition. I. Model structure," *J. Acoust. Soc. Am.* **110**, 1074–1088.

Brown, A. D., Kuznetsova, M. S., Spain, W. J., and Stecker, G. C. (2012). "Frequency-specific, location-nonspecific adaptation of interaural time difference sensitivity," *Hear. Res.* **291**, 52–56.

Brown, A. D., and Stecker, G. C. (2013). "The precedence effect: Fusion and lateralization measures for headphone stimuli lateralized by interaural time and level differences," *J. Acoust. Soc. Am.* **133**, 2883–2898.

Carlile, S., Hyams, S., and Delaney, S. (2001). "Systematic distortions of auditory space perception following prolonged exposure to broadband noise," *J. Acoust. Soc. Am.* **110**, 416–424.

Dahmen, J. C., Keating, P., Nodal, F. R., Schulz, A. L., and King, A. J. (2010). "Adaptation to stimulus statistics in the perception and neural representation of auditory space," *Neuron* **66**, 937–948.

Dietz, M., Ewert, S. D., and Hohmann, V. (2011). "Auditory model based direction estimation of concurrent speakers from binaural signals," *Speech Commun.* **53**, 592–605.

Djelani, T., and Blauert, J. (2001). "Investigations into the build-up and breakdown of the precedence effect," *Acta Acust. united Acust.* **87**, 253–261, available at <http://www.ingentaconnect.com/content/dav/aaau/2001/00000087/00000002/art00012>.

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

Furukawa, S. (2005). "Dependency of the interaural phase difference sensitivities of inferior collicular neurons on a preceding tone and its implications in neural population coding," *J. Neurophysiol.* **93**, 3313–3326.

Getzmann, S. (2004). "Spatial discrimination of sound sources in the horizontal plane following an adapter sound," *Hear. Res.* **191**, 14–20.

Grothe, B., Pecka, M., and McAlpine, D. (2010). "Mechanisms of sound localization in mammals," *Physiol. Rev.* **90**, 983–1012.

Herron, T. (2005). "C language exploratory analysis of variance with enhancements," available at <http://www.ebire.org/hcnlab/software/cleave.html> (Last viewed November 16, 2017).

Kashino, M., and Nishida, S. (1998). "Adaptation in the processing of interaural time differences revealed by the auditory localization aftereffect," *J. Acoust. Soc. Am.* **103**, 3597–3604.

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

Kopčo, N., Andrejková, G., Best, V., and Shinn-Cunningham, B. (2017). "Streaming and sound localization with a preceding distractor," *J. Acoust. Soc. Am.* **141**, EL331–EL337.

Kopčo, N., Best, V., and Carlile, S. (2010). "Speech localization in a multi-talker mixture," *J. Acoust. Soc. Am.* **127**, 1450–1457.

Kopčo, N., Best, V., and Shinn-Cunningham, B. G. (2007). "Sound localization with a preceding distractor," *J. Acoust. Soc. Am.* **121**, 420–432.

Kopčo, N., Marcinek, L., Tomoriová, B., and Hládek, L. (2015). "Contextual plasticity, top-down, and non-auditory factors in sound localization with a distractor," *J. Acoust. Soc. Am.* **137**, EL281–EL287.

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

Maddox, R. K., Pospisil, D. A., Stecker, G. C., and Lee, A. K. C. (2014). "Directing eye gaze enhances auditory spatial cue discrimination," *Curr. Biol.* **24**, 748–752.

Maier, J. K., McAlpine, D., Klump, G. M., and Pressnitzer, D. (2010). "Context effects in the discriminability of spatial cues," *J. Assoc. Res. Otolaryngol.* **11**, 319–328.

Martin, R. L., Flanagan, P., McAnally, K. I., and Eberle, G. (2011). "Memory for the locations of environmental sounds," *J. Acoust. Soc. Am.* **129**, 3873–3883.

Micheyl, C., Hanson, C., Demany, L., Shamma, S., and Oxenham, A. J. (2013). "Auditory stream segregation for alternating and synchronous tones," *J. Exp. Psychol. Hum. Percept. Perform.* **39**, 1568–1580.

Phillips, D. P., and Hall, S. E. (2005). "Psychophysical evidence for adaptation of central auditory processors for interaural differences in time and level," *Hear. Res.* **202**, 188–199.

Sach, A. J., Hill, N. I., and Bailey, P. J. (2000). "Auditory spatial attention using interaural time differences," *J. Exp. Psychol. Hum. Percept. Perform.* **26**, 717–729.

Sanders, L. D., Zobel, B. H., Freyman, R. L., and Keen, R. (2011). "Manipulations of listeners' echo perception are reflected in event-related potentials," *J. Acoust. Soc. Am.* **129**, 301–309.

Shamma, S. A., Elhilali, M., and Micheyl, C. (2011). "Temporal coherence and attention in auditory scene analysis," *Trends Neurosci.* **34**, 114–123.

Simpson, B. D., Brungart, D. S., Gilkey, R. H., Iyer, N., and Hamil, J. T. (2007). "Localization in multiple source environments: Localizing the missing source," in *Proc. 13th Int. Conf. Audit. Disp.*, June 26–29, Montreal, Canada, pp. 280–284.

Spence, C., and Driver, J. (1998). "Inhibition of return following an auditory cue the role of central reorienting events," *Exp. Brain Res.* **118**, 352–360.

Stange, A., Myoga, M. H., Lingner, A., Ford, M. C., Alexandrova, O., Felmy, F., Pecka, M., Siveke, I., and Grothe, B. (2013). "Adaptation in sound localization: From GABA(B) receptor-mediated synaptic modulation to perception," *Nat. Neurosci.* **16**, 1840–1847.

Weintraub, D. M., Metzger, B. A., and Snyder, J. S. (2014). "Effects of attention to and awareness of preceding context tones on auditory streaming," *J. Exp. Psychol. Hum. Percept. Perform.* **40**, 685–701.