

Streaming and sound localization with a preceding distractor^{a)}

Norbert Kopčo,^{1,b)} Gabriela Andrejková,¹ Virginia Best,²
and Barbara Shinn-Cunningham³

¹*Institute of Computer Science, Faculty of Science, P. J. Šafárik University, Jesenná 5,
04001 Košice, Slovakia*

²*Department of Speech, Language and Hearing Sciences, Boston University, Boston,
Massachusetts 02215, USA*

³*Hearing Research Center, Boston University, Boston, Massachusetts 02215, USA
norbert.kopco@upjs.sk, gabriela.andrejko@upjs.sk, ginbest@bu.edu, shinn@cns.bu.edu*

Abstract: Localization of a 2-ms click target was previously shown to be influenced by a preceding identical distractor for inter-click-intervals up to 400 ms [Kopčo, Best, and Shinn-Cunningham (2007). *J. Acoust. Soc. Am.* **121**, 420–432]. Here, two experiments examined whether perceptual organization plays a role in this effect. In the experiments, the distractor was designed either to be grouped with the target (a single-click distractor) or to be processed in a separate stream (an 8-click train). The two distractors affected performance differently, both in terms of bias and variance, suggesting that grouping and streaming play a role in localization in multisource environments.

© 2017 Acoustical Society of America

[DDO]

Date Received: August 25, 2016 **Date Accepted:** March 10, 2017

1. Introduction

Localization of a target sound can be dramatically influenced by the presence of a preceding distractor, even if the two sounds do not overlap in time. A previous study, in which listeners were asked to indicate the perceived lateral position of a target, found that when an identical click was used as both distractor and target, the distractor biased localization and increased response variability for inter-stimulus intervals (ISIs) of up to 400 ms (Kopčo *et al.*, 2007). The previous study was performed in both anechoic and reverberant space using a setup in which the distractor location was fixed either in front of the listener or on his/her side [Fig. 1(A)]. The study identified two distinct types of biases that were interpreted as being caused by central neural mechanisms: (1) a strong attractive bias toward the lateral distractor for frontal targets at ISIs of 25 to 100 ms, which occurred in reverberant, but not in anechoic space, likely due to a central mechanism of adaptation to room reverberation (e.g., Clifton *et al.*, 2002) and (2) a bias away from the lateral distractor for nearby targets, largely independent of the ISI, likely caused by a change in response strategy (Kopčo *et al.*, 2010). While these effects were observed with the lateral, but not with the frontal, distractor, both distractors also induced increases in response variability that were stronger in reverberation than in anechoic space.

The current study tested the hypothesis that the effects observed in the previous study were caused in part by perceptual organization (e.g., Elhilali *et al.*, 2009). Specifically, given that the distractor and target were identical clicks in the previous study, they may have been processed as a single auditory object or stream, which could explain some of the observed interactions (e.g., perceptual integration of the target and distractor could explain attractive biases). To test this hypothesis, the current study replicated two conditions of the previous study, and included two additional conditions in which the distractor was modified to reduce the likelihood of it being grouped with the target. The new distractor consisted of eight clicks identical to the target, presented in an isochronous sequence with a peak-to-peak period that differed from the ISI [Fig. 1(A)]. Thus, while the final distractor click was identical to the 1-click distractor, the preceding context was designed to capture the final distractor click into a stream distinct from the target (e.g., Rajendran *et al.*, 2013). Given this, we predicted that the 8-click distractor would interact less with the target than the 1-click distractor, mitigating some of the effects observed in

^{a)}Portions of this work were presented at the 169th Meeting of the Acoustical Society of America.

^{b)}Author to whom correspondence should be addressed. Also at: Hearing Research Center, Boston University, Boston, MA 02215, USA.

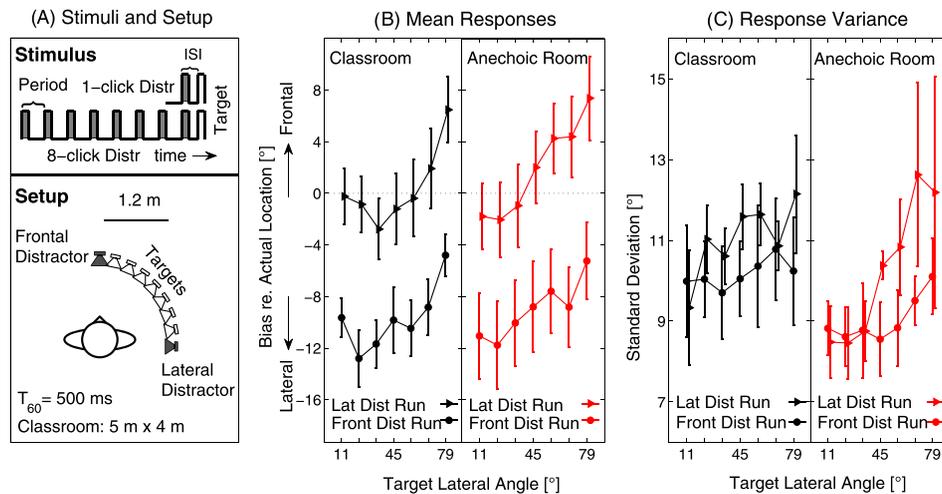


Fig. 1. (Color online) (A) Experimental design. (B) and (C) Performance in no-distractor (target-alone) baseline trials from the frontal-distractor runs and lateral-distractor runs, plotted as response bias re. actual target location (B) and response standard deviation (C) as a function of the actual target location. Across-subject mean (\pm standard error) data from the Classroom experiment (Exp. 1) and the Anechoic Room experiment (Exp. 2) are shown in separate subpanels and by different colors. Triangles and circles show, respectively, the data from runs in which the interleaved distractor-target trials had the distractor fixed in front and to the side.

the previous study. In particular, we expected the effects of the lateral distractor, which were interpreted as centrally mediated in Kopčo *et al.* (2007), to be weakened, including the response bias due to reverberation suppression, the response bias due to a change in response strategy, and the increases in response variability. On the other hand, we expected little change in the effect of the frontal distractor as none of the biases caused by this distractor appeared to be centrally mediated in the previous study.

Two experiments were performed, which were identical except for the environment in which they were conducted (one in a classroom and one in an anechoic room), using a design very similar to that of Kopčo *et al.* (2007). In this design, frontal and lateral distractor locations were tested in separate blocks, and baseline (no-distractor) trials were interleaved with the distractor trials. Importantly, we expected that these baseline responses would be shifted differently in the frontal-distractor vs lateral-distractor runs, creating a contextual effect, as reported in Kopčo *et al.* (2007). However, this contextual effect, operating on the time scale of tens of seconds to minutes, was not expected to interact with the effect of the preceding distractor, which operates on time scales of up to 400 ms (as confirmed in Kopčo *et al.*, 2015). We tested performance for one short ISI (50 ms) and one long (200 ms) ISI, predicting that effects of grouping would be visible in results for the shorter ISI, and weaker for the longer ISI.

2. Methods

2.1 Subjects

Seven listeners (three females) with ages ranging from 23 to 32 yrs, including authors N.K. and V.B., participated in experiment 1 (Classroom), and four of these listeners also participated in experiment 2 (Anechoic Room). All listeners reported normal hearing, gave informed consent, and were paid for their participation. The listeners had previously participated in experiments 1 and 2 of Kopčo *et al.* (2007).

2.2 Listening environment

Experiment 1 was conducted in an empty, quiet rectangular classroom. The reverberation times in octave bands centered at 500, 1000, 2000, and 4000 Hz were 613, 508, 512, and 478 ms, respectively. The background acoustic noise was at a level of approximately 39 dBA. Experiment 2 was conducted in an anechoic room. Both rooms were the same as those used in Kopčo *et al.* (2007). The listener was seated approximately in the center of either room with his/her head held stable by a head rest. Nine loudspeakers (Bose Acoustimass, Bose, Framingham, MA) were positioned on an arc with radius of 1.2 m spanning 90°. The listener sat in the center of the arc and faced either the left-most loudspeaker [so that the targets occurred on his/her right, see Fig. 1(A)] or the right-most loudspeaker [setup mirror-flipped compared to Fig. 1(A)]. In the following, 0° azimuth always represents the location directly ahead of the listener, and 90° is the location of the left- or right-most speaker (depending on the listener

orientation). The loudspeakers were not hidden, but the listeners kept their eyes closed during runs to minimize responses clustering at the loudspeaker locations. Digital stimuli were generated by a TDT System 3 audio interface and passed through power amplifiers (Crown D-75 A, Crown Audio, Elkhart, IN) to the loudspeakers. The listener held a pointer in one hand for indicating the perceived direction of each target. A Polhemus FastTrak electromagnetic tracker was used to measure the location of the listener's head, the approximate location of the loudspeakers, and the listener's responses.

2.3 Stimuli and task

A single 2-ms frozen noise burst presented at 67 dBA (maximum root-mean-square value in a 2-ms running window for continuous noise at the location of the listener's head) was used as both target and distractor in the 1-click distractor condition. Eight such clicks presented at the rate of 10/s (peak-to-peak period of 100 ms) were used as the 8-click distractor. The distractor was presented from the frontal or the lateral speaker (fixed within a run). On each trial, the target location was randomly selected from one of the seven central loudspeakers (spanning approximately 11° – 79° azimuth). The distractor-target ISI, measured from the onset of the final distractor click to the target click, was either 50 or 200 ms. Note that in the local context of a preceding 8-click distractor, the target occurs earlier than “expected” for the 50-ms ISI, and later than “expected” for the 200-ms ISI. Runs consisted of an equiprobable mixture of five trial types: target-alone (no-distractor), 1-click-distractor 50-ms-ISI, 1-click-distractor 200-ms-ISI, 8-click-distractor 50-ms-ISI, and 8-click-distractor 200-ms-ISI. Every combination of the five trial types and seven target locations was presented four times in random order within a run. The subject changed his/her orientation after each run to face either the left-most loudspeaker or the right-most loudspeaker by rotating his/her whole body. Experiments 1 and 2 each comprised four sessions. Each session, which took approximately 30 min, contained four runs, one for each combination of subject orientation (facing the left-most speaker, facing the right-most speaker) and distractor location (frontal, lateral).

2.4 Data analysis

Negligible left–right differences were observed, so the data were collapsed across the two listener orientations prior to statistical analysis. All reported statistical analyses were performed as repeated-measures analysis of variance (ANOVA). Main effects or interactions that did not reach $p < 0.05$ significance are not reported.

3. Results

3.1 Baseline performance

Figure 1 shows the mean response bias [i.e., the difference between the response location and the true target location, Fig. 1(B)] and standard deviation [Fig. 1(C)] in the no-distractor baseline trials that were randomly interleaved with the distractor-target trials during the experimental runs. The data are shown separately for the two experiments (plotted in different subpanels and by a different color) and for the two different distractor locations (triangle vs circle symbols).

In both rooms the responses in the frontal-distractor runs were shifted laterally compared with the responses in the lateral-distractor runs [the circles fall below the triangles in Fig. 1(B)]. The response standard deviations increased with target laterality in both rooms, especially in the lateral-distractor runs, where the effects are greater than in the frontal-distractor runs [the circles tend to fall below the triangles in Fig. 1(C)]. A similar contextual effect, where the distribution of stimuli heard in a given run (not just within a trial) affects mean localization responses, was also observed in Kopčo *et al.* (2007) and has been further explored in Kopčo *et al.* (2015). These previous studies show that contextual effects operate on the time scale of tens of seconds to minutes and do not influence the effects of an immediately preceding distractor that occurs within hundreds of milliseconds (the focus of the current study). More importantly, the contextual bias is common to both 1- and 8-click distractor trials and thus should not affect comparisons across these key conditions (see Sec. 3.2).

3.2 Effect of distractor

We expected that certain effects of a preceding distractor on target localization would be reduced if the target and distractor were perceived as distinct auditory streams. In particular, we expected reductions in the centrally mediated effects seen for the lateral distractors and short ISIs. To evaluate this hypothesis, Figs. 2 (for the lateral distractor)

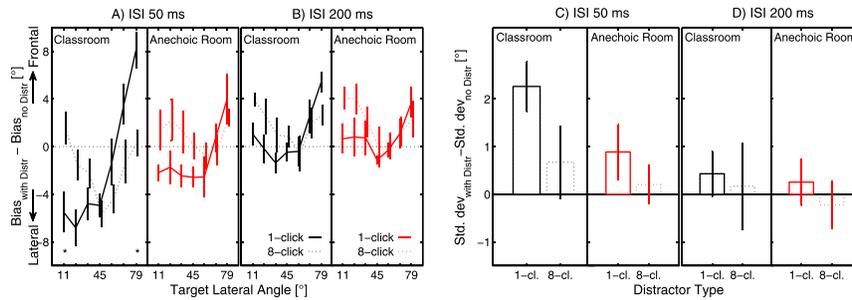


Fig. 2. (Color online) Effect of lateral distractor on the response bias [(A) and (B)] and standard deviation [(C) and (D)] re. baseline from Fig. 1. Across-subject mean (\pm standard error) data from the Classroom experiment (Exp. 1) and the Anechoic Room experiment (Exp. 2) are shown in separate subpanels and by different colors, plotted as a function of target lateral angle in (A) and (B), and collapsed across the lateral angle in (C) and (D). Solid and dotted lines/bars show, respectively, the responses when the target is preceded by 1- and 8-click distractor. Asterisks along the x axis indicate target locations at which a Bonferroni-corrected paired t -test found a significant difference between the two distractor types ($p < 0.05$).

and 3 (for the frontal distractor) compare performance with the 1- and 8-click distractor in terms of response bias and variance in the two acoustic environments.

3.2.1 Lateral distractor

Figures 2(A) and 2(B) show the response bias induced by the preceding lateral distractor [re. no-distractor baseline from Fig. 1(B)], separately for each combination of ISI and room. In each panel, results for the 1-click distractor are plotted with solid lines and results for the 8-click distractor are plotted with dotted lines. At the 50-ms ISI [Fig. 2(A)], the 1-click distractor induced an attractive bias of up to 6° for frontal targets in the classroom, an effect that was reduced to 3° in the anechoic room. When the 1-click distractor was replaced by an 8-click distractor, this attractive bias was reduced or eliminated in both rooms. At the 200-ms ISI [Fig. 2(B)], the 1-click distractor did not induce any bias for frontal targets. However, the 8-click distractor induced a repulsive bias of up to 4° in both rooms. At the other end of the target range, the most lateral target was perceived with a frontal bias of 4° – 8° when preceded by the 1-click distractor in both rooms and at both ISIs [solid lines in Figs. 2(A) and 2(B)]. The 8-click distractor (dotted lines) eliminated this bias in the 50-ms classroom condition (in which it was largest for the 1-click distractor), reduced it in the 200-ms classroom condition, and had a tendency to reduce it in both anechoic conditions.

These broad observations were confirmed by three-way repeated-measures ANOVAs, performed separately for each combination of distractor location and room, which are summarized in Table 1. In the classroom, a significant 3-way interaction was found. To interpret the interaction, partial ANOVAs were run separately for the two ISIs. Both of these ANOVAs found significant interactions between target location and distractor type ($p < 0.0005$), confirming that streaming had location-dependent effects at both ISIs. Finally, a set of paired t -tests with Bonferroni corrections was performed, comparing the 1- and 8-click data separately for each target location [asterisks in Figs. 2(A) and 2(B)], confirming that the effect was highly significant at the two extreme locations but only with the ISI of 50 ms. In the anechoic room, no significant

Table 1. Results of three-way repeated measures ANOVA performed on the response bias data separately for each combination of distractor location and room. F values and significance values are listed only for the effects and interactions with significance level $p < 0.05$. Significance levels are as follows: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.005$.

Factor	df	Classroom		Anechoic room	
		Lat Dist	Front Dist	Lat Dist	Front Dist
Target location (7 Spkrs)	6, 36	14.28***	4.5***	6, 18	10.21***
Distractor Type (1-cl., 8-cl.)	1, 6			1, 3	
ISI (50 ms, 200 ms)	1, 6	15.86**		1, 3	
Target \times Dist Type	6, 36	5.58***	3.38***	6, 18	8.22***
Target \times ISI	6, 36	14.62***	2.61*	6, 18	
Dist Type \times ISI	1, 6	4.73*		1, 3	
Target \times Dist Type \times ISI	6, 36	5.58***	3.58**	6, 18	

Table 2. Results of three-way repeated measures ANOVA performed on the response standard deviation data separately for each combination of distractor location and room. F values and significance values are listed only for the effects and interactions with significance level $p < 0.05$. Significance levels are as follows: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.005$.

Factor	df	Classroom		df	Anechoic room	
		Lat Dist	Front Dist		Lat Dist	Front Dist
		F Signif.	F Signif.		F Signif.	F Signif.
Target location (7 Spkrs)	6, 36			6, 18		2.86*
Distractor Type (1-cl., 8-cl.)	1, 6	15.99**		1, 3		
ISI (50 ms, 200 ms)	1, 6	37.03***		1, 3		
Target \times Dist Type	6, 36			6, 18		3.68*
Target \times ISI	6, 36			6, 18		
Dist Type \times ISI	1, 6	15.96**		1, 3		
Target \times Dist Type \times ISI	6, 36			6, 18		

interaction involving ISI was found, while a significant interaction between target location and distractor type (for both distractor locations) confirmed that the effect of streaming varied with location. None of the paired t -tests found significant differences. Thus, it is not possible to state conclusively which locations drove the significant interaction.

Figures 2(C) and 2(D) plot the change in response standard deviation induced by the preceding lateral distractor [re. no-distractor baseline from Fig. 1(C)] using a layout similar to Figs. 2(A) and 2(B), but with the data collapsed across target location. The largest increase in response standard deviations, on average more than 2° , was observed with the lateral 1-click distractor for the 50-ms ISI in the classroom [black solid bar in Fig. 2(C)]. The increase was reduced in the anechoic room to about 1° [red solid bar in Fig. 2(C)]. In both cases, the increases in variance were reduced or eliminated with the 8-click distractor (corresponding dotted bars). The distractor caused no consistent effects on response variability for the ISI of 200 ms [Fig. 2(D)]. These broad observations were confirmed by three-way repeated-measures ANOVAs, performed separately for each room, which are summarized in Table 2. Specifically, these ANOVAs showed a significant interaction between distractor type and ISI for the classroom data; however, this interaction was not significant for the anechoic data.

3.2.2 Frontal distractor

Figure 3 shows the frontal distractor data using a layout identical to that of Fig. 2; Tables 1 and 2 show the corresponding statistical analyses. The bias data [Figs. 3(A) and 3(B)] show that the frontal distractor tended to attract frontal targets, with the largest shifts occurring for targets at 20° – 30° . For the most lateral targets the bias was reduced (in the classroom experiment) or even reversed (in the anechoic experiment). This pattern was largely independent of the ISI [Figs. 3(A) and 3(B) are similar] and the distractor type (solid and dotted lines are similar within each panel, with one exception: for the 11° target in the classroom at 50-ms ISI, the 8-click distractor induced a larger bias than the 1-click distractor). Three-way repeated-measures ANOVAs performed on these data found a significant 3-way interaction in the classroom (Table 1). To analyze this interaction, additional partial ANOVAs were performed separately for the two ISIs. A

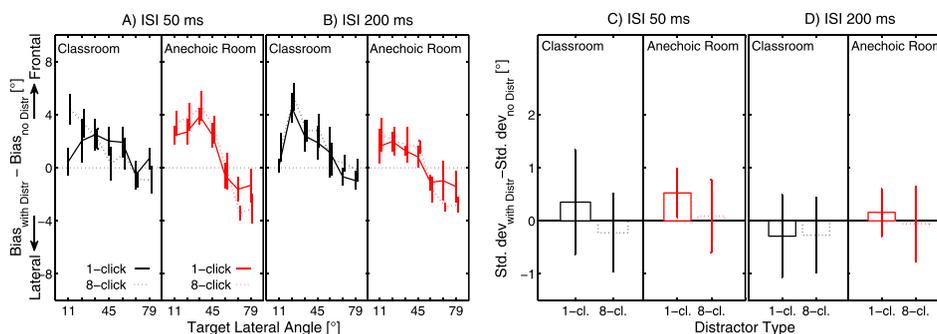


Fig. 3. (Color online) Effect of frontal distractor on the response bias [(A) and (B)] and standard deviation [(C) and (D)] (re. baseline from Fig. 1). Layout of the figure is identical to Fig. 2.

significant interaction between target location and distractor type was found for the 50-ms ISI ($p < 0.0005$) but not for the 200-ms ISI, suggesting that the 3-way interaction was driven by the single point [11° target in Fig. 3(A)] where the distractor type affected the bias (however, paired t -tests found no significant differences). In the anechoic room, a significant interaction between target location and distractor type was found, suggesting that there was an effect of streaming for some target locations. However, as seen in Figs. 3(A) and 3(B), the differences are small (never larger than 1° – 2°) and well within the error bars (paired t -tests found no significant differences).

Figures 3(C) and 3(D) plot the response standard deviations. This plot shows that the response variability was not affected by the presence of the frontal distractor for any of the conditions. Three-way repeated measures ANOVAs performed on these data (Table 2) found no significant effects in the classroom and a marginally significant interaction between target location and distractor type in the anechoic room. This again suggests that there might be a small reduction in response variability for the 8-click distractor at some locations.

4. Discussion and conclusions

The current study replicated the results of Kopčo *et al.* (2007), showing that a preceding distractor click can affect the localization of a target click, inducing different types of localization biases, as well as increasing response variance. Here, some of these effects were reduced when the distractor was replaced by an 8-click distractor that was designed to have similar low-level effects on the processing of the target stimulus but to be perceived as a distinct auditory stream. Specifically, the effects of the preceding distractor that we speculated were “central,” or related to perceptual organization, were mostly reduced when the distractor and target were unlikely to be perceptually grouped together. These reductions were greatest in conditions where the original effects were most dramatic (for the lateral distractor condition and at the shorter ISI).

A lateral, 1-click distractor caused an attractive bias for the frontal targets at the short ISI [Fig. 2(A)], an effect that was stronger in the classroom than in anechoic space. This effect was eliminated by streaming, which suggests that the underlying mechanism causing the attractive bias (possibly adaptation to room reverberation related to the precedence effect; Clifton *et al.*, 2002; Freyman *et al.*, 1991) is not activated if the two stimuli are processed as separate objects. However, at the long ISI, an unexpected effect of streaming was observed for frontal targets: the 8-click lateral distractor induced a small frontal bias, while there was no bias with the 1-click lateral distractor. It is possible that a different mechanism, e.g., related to inhibition of return (Spence and Driver, 1998) was activated at this longer ISI, expanding the perceived spatial separation between the two streams and giving rise to these repulsive shifts.

The lateral 1-click distractor caused a repulsive bias for lateral targets in both environments and at all ISIs [Figs. 2(A) and 2(B)]. This effect was again reduced or eliminated by streaming in the classroom (where the 1-click effect was strong), and it showed a tendency to be reduced in the anechoic environment (where the 1-click effect was weaker) confirming that its origin was not peripheral. We originally speculated that subjects may have adopted a *relative* rather than *absolute* localization strategy when the 1-click distractor and target were in the same vicinity (Kopčo *et al.*, 2010; Recanzone *et al.*, 1998). These new data suggest that this strategy change was less likely when the distractor and target were very dissimilar.

Finally, the lateral 8-click distractor caused smaller increases in response variance compared to the 1-click distractor, suggesting that perceptual organization plays a role in this effect. It is worth noting that this change in response variability was similar for all target locations; in contrast, differences in how the 8- and 1-click distractors affected localization bias depended on target location. Thus, the effects of streaming on these two aspects of performance may arise from independent mechanisms.

Overall, these results support our hypothesis that perceptual grouping contributes to the effects of a preceding distractor on localization like those observed in our previous study. However, contrary to our original hypothesis, we found here that perceptually segregating the distractor and target into different streams did not always lead to more accurate target localization, as the 8-click distractor induced additional bias in several conditions.

It is not clear why the effects of grouping (i.e., the effects of the 1-click distractor, and the reduction in these effects with the 8-click distractor) were more dramatic when the distractor was located laterally compared to when it was located frontally. One possibility is that the image of the lateral distractor is broader and less salient than the image of the frontal distractor, resulting in a stronger tendency for it to group

with the target. In that case, it might be that the additional clicks allow for the formation of a more distinct object with a tighter spatial representation (Best *et al.*, 2008), which further enhances its segregation from the target.

Finally, it is important to note the limitations of this study. First, the 8-click distractor differed from the 1-click distractor in more than just the tendency to stream with the target. For example, given that the 8-click distractor was longer in duration and contained more clicks, it is possible (but unlikely) that some of the effects observed here were a result of other processes, e.g., related to adaptation or attention. Also, the contextual effect observed here might have interacted differentially with the 1- and 8-click distractors, even though it has been shown that the contextual and preceding-distractor effects are largely independent of each other for the 1-click distractor (Kopčo *et al.*, 2015).

These results illustrate that complex adaptive mechanisms active at multiple processing levels and multiple temporal scales interact when localization is examined, even for a relatively simple setup consisting of only two temporally non-overlapping sources. Future studies will need to investigate these different mechanisms and their interactions, as well as the slow-time-scale contextual effects observed here and in Kopčo *et al.* (2007).

Acknowledgments

This work was supported by the SRDA, Contract No. APVV-0452-12, EU H2020-MSCA-RISE-2015 Grant No. 69122, and by the TECHNICOM project, ITMS: 26220220182, of the EU RDP. V.B. was supported by NIH-NIDCD Grant No. R01 DC04545. B.S.-C. was supported by NIH-NIDCD Grant No. R01 DC013825.

References and links

- Best, V., Ozmeral, E. J., Kopčo, N., and Shinn-Cunningham, B. C. (2008). "Object continuity enhances selective auditory attention," *Proc. Natl. Acad. Sci. U.S.A.* **105**(35), 13174–13178.
- Clifton, R. K., Freyman, R. L., and Meo, J. (2002). "What the precedence effect tells us about room acoustics," *Percept. Psychophys.* **64**, 180–188.
- Elhilali, M., Ma, L., Micheyl, C., Oxenham, A. J., and Shamma, S. A. (2009). "Temporal coherence in the perceptual organization and cortical representation of auditory scenes," *Neuron* **61**, 317–329.
- Freyman, R. L., Clifton, R. K., and Litovsky, R. Y. (1991). "Dynamic processes in the precedence effect," *J. Acoust. Soc. Am.* **90**, 874–884.
- Kopčo, N., Best, V., and Carlile, S. (2010). "Speech localization in a multitalker 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.
- Rajendran, V. G., Harper, N. S., Willmore, B. D., Hartmann, W. M., and Schnupp, J. W. H. (2013). "Temporal predictability as a grouping cue in the perception of auditory streams," *J. Acoust. Soc. Am.* **134**, EL98–E104.
- Recanzone, G. H., Makhama, S. D., and Guard, D. C. (1998). "Comparison of relative and absolute sound localization ability in humans," *J. Acoust. Soc. Am.* **103**(2), 1085–1097.
- Spence, C., and Driver, J. (1998). "Auditory and audiovisual inhibition of return," *Percept. Psychophys.* **60**, 125–139.