Contextual plasticity, top-down, and non-auditory factors in sound localization with a distractor[a)]

Norbert Kopčo,[b)] Ľuboš Marcinek, Beáta Tomoriová, and Ľuboš Hládek
Institute of Computer Science, Faculty of Science, P. J. Šafárik University, Jesenná 5, 04001 Košice, Slovakia
norbert.kopco@upjs.sk, marcineklubos@yahoo.com, beata.tomoriova@upjs.sk, lubos.hladek@upjs.sk

Abstract: Localization of a 2-ms-click target was previously shown to be influenced by interleaved localization trials in which the target was preceded by an identical distractor [Kopčo, Best, and Shinn-Cunningham (2007). J. Acoust. Soc. Am. **121**, 420–432]. Here, two experiments were conducted to explore this contextual effect. Results show that context-related bias is not eliminated (1) when the response method is changed so that vision is available or that no hand-pointing is required; or (2) when the distractor-target order is reversed. Additionally, a keyboard-based localization response method is introduced and shown to be more accurate than traditional pointer-based methods.

[QJF]
Date Received: October 30, 2014 **Date Accepted:** March 1, 2015

1. Introduction

A previous study of short-term auditory localization plasticity found an unexpected form of adaptation that occurred on a time scale of seconds to tens of seconds (Kopčo *et al.*, 2007). In that study, the experimental trials, in which the target was preceded by a distractor, were interleaved with reference trials, in which the target was presented alone. The unexpected adaptation, called contextual plasticity, was observed in the reference trials. There, the responses shifted by as much as 10° depending on whether the interleaved experimental trials had the distractor in front or to the side of the listener. While the previous study showed that the effect is observed in both anechoic and reverberant environments, many questions about the effect were left unanswered. For example, in the previous study the distractor-target stimulus onset asynchrony, SOA, was varied between 25 and 400 ms from trial to trial. Thus, it is not clear whether the easy, 400-ms-SOA trials contributed to the contextual effect as much as did the difficult, 25-ms-SOA trials.

This paper is the first in a series that examines contextual plasticity. We present two follow-up experiments. Experiment 1 kept SOA fixed at 25 ms and explored whether the effect is caused by non-auditory factors like the response method used (Brungart *et al.*, 2000). In the original study, the subjects provided responses using a hand-held pointer while keeping their eyes closed. In Exp. 1, the performance obtained with this response method was compared to two other methods. The first one used an identical pointing method while the subjects kept their eyes open. The second one used an array of alphanumeric labels along the speaker array and the subject responded by reading out and typing the letter–number combination from the label nearest to the perceived target location, thus eliminating the sensory-motor transformation required to direct the hand-held pointer. If contextual plasticity is related to the unavailability

[a)]Portions of this work were presented at the 32nd meeting of the Association for Research in Otolaryngology and at Forum Acusticum 2014.
[b)]Author to whom correspondence should be addressed.

of vision or the pointing-related sensory-motor transformations, it would be eliminated with these response methods.

Experiment 2 examined whether contextual plasticity is related to the contextual task difficulty and the strategies adopted by the subjects to cope with it (Snyder and Weintraub, 2013). In the previous study, the distractor location was fixed within a run such that the target array was always to the left or to the right of it. Given that the task was rather difficult on the distractor-target trials, in particular, when SOA was 25 ms, the listeners might have tried to focus away from the distractor toward the target array. Employing this strategy could have resulted in shifts in localization away from the distractor on reference target-alone trials, as was observed in that study. To test this hypothesis, the experimental task was modified to be very easy by fixing the SOA to 400 ms, which also reduced any acoustic or short-term peripheral interference due to the distractor, or by presenting the target prior to the distractor (and keeping the 400-ms SOA), thus allowing the listener to only pay attention to the initial stimulus and ignore the rest. It was expected that if, in the previous study, contextual plasticity was related to the difficulty of the contextual trials, the effect would be eliminated in the new conditions (Maier *et al.*, 2010). On the other hand, if the effect was related to the peripheral processing and stimulus distribution (Dahmen *et al.*, 2010), it would persist even in these conditions.

The current study also examined whether contextual plasticity is observed in response variance. Finally, the effect of preceding or following distractor on target localization was measured in several new conditions, extending the results of Kopčo *et al.* (2007).

2. Methods

2.1 Subjects

Ten listeners participated in Experiment 1, seven in Experiment 2. All listeners reported normal hearing and gave informed consent.

2.2 Listening environment

The environment was similar to that used in the reverberant experiment of the previous study. Listeners were seated in a darkened single-walled booth with their head supported by a head rest. Nine loudspeakers (Bose Acoustimass cube speakers, Bose, Framingham, MA) were positioned on an arc with diameter of 1.1 m, spanning 90° around the listener who sat in the center of the arc and faced either the left-most or the right-most speaker [see inset in Fig. 1(c)]. Speakers were covered by an acoustically transparent cloth. Labels with random letter–number combinations (e.g., X3) were attached to the cloth above the speakers, covering a 110° range in 1° steps. Digital stimuli were generated by an RME Fireface 400 (RME Audio, Haimhausen, Germany) audio interface and passed through power amplifiers (Crown D-75A, Crown Audio, Elkhart, IN) to the loudspeakers. The subjects responded by either pointing a hand-held video-tracked pointer or by entering a letter–number combination on a computer keyboard.

2.3 Stimuli and task

The stimuli and procedures were very similar to those used in the previous study. A single 2-ms frozen noise burst presented at 69 dBA was used as both target and distractor throughout the study. The distractor was always presented from the frontal speaker. The target location was on each trial randomly selected from one of the 7 loudspeakers at 11°–79°. The target-distractor SOA was 25 ms in Exp. 1 and 400 ms in Exp. 2. Runs of 175 trials (Exp. 1) or 203 trials (Exp. 2) consisted of either only target-alone trials, in baseline runs, or a mixture of 75% distractor + target (contextual) trials and 25% of target-alone trials, in the adaptation part of experimental runs. The experimental runs started with 14 target-alone pre-adaptation trials and ended with 21 target-alone post-adaptation trials. The subject changed his/her orientation after each run to face either the left-most speaker, so that the target speakers were on his/her right [as shown in Fig. 1(c)], or the right-most speaker [setup mirror-flipped compared

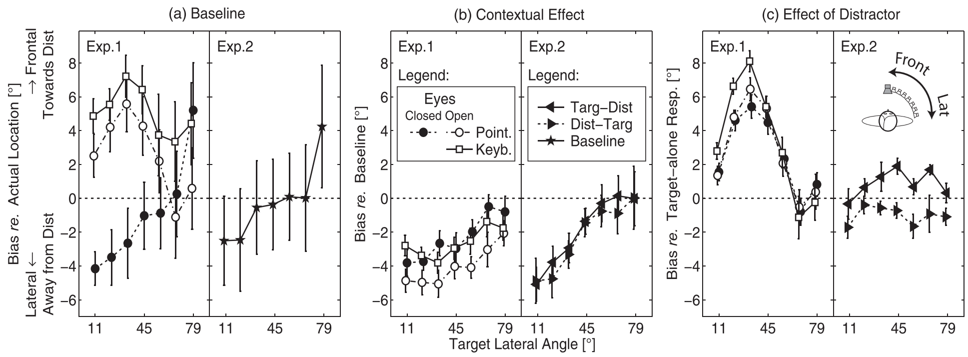

Fig. 1. Across-subject mean (± standard error of the mean, SEM) response bias as a function of the actual target location. The left-hand portion of each panel shows data from Exp. 1, the right-hand portion from Exp. 2. The experimental setup is shown as the inset in panel (c), which also shows the directions corresponding to frontal and lateral biases. (a) Bias in responses re. actual target location in the baseline runs. (b) Contextual effect defined as the bias in target-alone responses in experimental runs re. baseline runs [from panel (a)]. (c) Bias due to the presence of distractor computed as the difference between the responses to distractor + target stimuli and the target-alone stimuli [from panel (b)] in the experimental runs.

to the Fig. 1(c) inset]. Experiment 1 consisted of 4 sessions, each containing 15 runs, 5 runs per condition (Closed-Eyes-Pointer, Open-Eyes-Pointer, Keyboard), 4 of which were experimental and 1 baseline. Experiment 2 consisted of 4 sessions, each containing 5 runs (2 Distractor-Target, 2 Target-Distractor, 1 baseline).

2.4 Data analysis

Data were collapsed across the two listener orientations prior to analysis. All reported statistical analyses were performed as two-way repeated-measures analysis of variances (ANOVAs) with factors of target location and condition. Main effects or interactions that did not reach $p < 0.05$ significance are omitted.

3. Results and discussion: Response biases

3.1 Baseline performance

Figure 1(a) plots the baseline performance measured in separate runs consisting of target-alone trials. Each graph represents the mean response bias (re. actual target location) as a function of the target location. In Exp. 1, a separate baseline was measured for each response method (Closed-Eyes-Pointer, Open-Eyes-Pointer, Keyboard). In the Closed-Eyes-Pointer condition the responses were consistently biased toward the middle of the speaker range, causing the response range to be compressed by approximately 13% (filled circles for Exp. 1 and asterisks for Exp. 2 show lateral bias for frontal targets and frontal bias for lateral targets). In the Open-Eyes-Pointer condition, the responses to the targets at 11° to 45° were frontally biased while the two most lateral locations showed no bias (open circles). In the Keyboard condition the responses at all locations were frontally biased by roughly 5° (open squares). ANOVA performed on Exp. 1 data confirmed that response method interacted with target location ($p < 0.001$).

These results illustrate that the availability of visual signals and the involvement of sensory-motor transformations (to respond using a hand-held pointer) do affect the measured localization performance. Overall, the Keyboard response method was the most accurate, introducing only an approximate constant frontal bias. This bias is likely related to the fact that the subjects tended to look straight ahead before the stimulus was presented in each Keyboard-run trial, as previous studies reported that sound localization can be biased toward the gaze direction (Razavi *et al.*, 2007). The pattern of results observed in the Open-Eyes-Pointer condition was similar, even though it was less consistent, possibly because the subjects did not use vision to guide their responses to the most lateral targets. On the other hand, the compressed response range observed with the Closed-Eyes-Pointer method illustrates that with no visual inputs the subjects were less accurate, biasing their responses toward the middle of the response range. It is likely that a similar compressive bias also influenced the results of Kopčo *et al.* (2007) which used the Closed-Eyes-Pointer method.

3.2 Contextual effect

Figure 1(b) plots the contextual effect, defined as the bias in the responses on the target-alone trials from the experimental runs re. baseline runs [from panel (a)]. In Exp. 1, the contextual bias was observed for all three response methods, exhibited by up to 5° lateral bias for the frontal targets (near the distractor location) and gradually decreasing with increasing target laterality. This effect was slightly modulated by the response method used, with the largest lateral offsets observed for the Open-Eyes-Pointer method (open circles) and slightly smaller shifts shown for the other two methods (filled circles and squares). The main effects of target location ($p < 0.001$) and response method ($p = 0.045$) were significant. In Exp. 2, a very similar significant contextual effect was observed both for the Distractor-Target and Target-Distractor runs (right-pointing triangles and left-pointing triangles, respectively).

The Exp. 1 results for the Closed-Eyes-Pointer vs Open-Eyes-Pointer methods show that the contextual effect observed in Kopčo *et al.* (2007) cannot be explained by the fact that the subjects had their eyes closed while performing the task. In fact, in the current study the contextual bias increased in the Open-Eyes-Pointer condition. However, this small difference might be related to the compressed baseline response range in the Closed-Eyes-Pointer condition [Fig. 1(a)]. There was no significant difference between the Keyboard and Open-Eyes-Pointer conditions (pairwise *t*-test, $p = 0.15$). So, it can be concluded that none or, at best, a small part of the contextual effect is related to the sensory-motor transformations required to respond using a hand-held pointer.

In Exp. 2 contextual plasticity persisted when the contextual task was made easy by increasing the SOA to 400 ms or by presenting the distractor 400 ms after the target on the contextual trials. Thus, it is unlikely that the effect is related to the subjects trying to focus away from the frontal distractor and toward the lateral targets—a strategy that might have been employed in the previous study when the task was difficult due to short SOAs. Instead, the effect can be induced by the mere presence of distractors even if they can be easily ignored. Thus, it is likely that the effect is caused by some bottom-up process, e.g., related to the distribution of the stimuli in the auditory scene (Dahmen *et al.*, 2010). Apparently, when a new stimulus location is added (in our case, the frontal location of the distractor, not present in the baseline), the responses to nearby targets get shifted in the direction away from that new stimulus location.

3.3 Effect of distractor

Figure 1(c) plots the effect of distractor defined as the bias in the responses on the distractor + target trials re. target-alone trials in the experimental runs [from panel (b)]. For all three response methods in Exp. 1, the distractor induced a frontal bias in responses that peaked at the 34° location and decreased with increasing angular distance from that location. The strongest bias was measured in the Keyboard condition (up to 8°, open squares), and a slightly weaker bias was measured in the two Pointer conditions (up to 6°, open and filled circles). These differences were significant (a significant interaction, $p < 0.001$). In Exp. 2 in which the SOA was 400 ms, the effect of distractor was much smaller (only around 1°–2°) and approximately independent of the target location. There was a small effect of condition ($p < 0.005$) such that the preceding distractor tended to repulse the target (lateral bias) while the distractor presented after the target tended to attract responses (frontal bias).

The results of Exp. 1 are consistent with those observed in the 25-ms frontal distractor condition of Kopčo *et al.* (2007) in which the maximum bias was 3°. The small differences between the current conditions are likely related to the differences in the accuracy of the methods. The Keyboard method, which was the most accurate in the baseline, produced the largest bias estimates while the least-accurate Closed-Eyes-Pointer method produced the smallest bias.

Biases in Exp. 2 were small, similar to the 400-ms condition in Kopčo *et al.* (2007). However, the order of stimulus presentation had a significant effect, such that all Target-Distractor responses were approximately 2° more frontal than the corresponding Distractor-Target responses. This result suggests that the listeners were not able to process the target location independent of the distractor even when they were separated by 400 ms and/or the target was presented first. This effect might be related to automatic attentional phenomena like inhibition of return (Spence and Driver, 1998).

3.4 Temporal profile of the contextual effect

Figure 2 shows the temporal profile of the buildup and decay of the contextual bias in Exp. 2. The first panel shows the temporal profile for the 11° target for which the largest contextual bias was observed [right-hand portion of Fig. 1(b)]. The bias had a very sharp onset and offset, reaching maximum/minimum within two to three subruns (a subrun corresponds to seven trials) after the contextual trials were introduced/eliminated. The pattern is similar but weaker for the 34° targets and it vanishes completely for the 56° and more lateral targets. Similar results were obtained for Exp. 1 (data not shown).

An average trial duration in Exp. 2 was 2–3 s. Given that the contextual bias was fully built-up and largely decayed within three subruns, the adaptive process underlying it can be assumed to operate on the time scales of tens of seconds. This time scale is similar to that of the precedence effect buildup phenomenon (Freyman *et al.*, 1991) which is examined under similar conditions. Thus the underlying mechanism might be related for the two effects.

4. Results and discussion: Response standard deviation

4.1 Baseline performance

Figure 3(a) plots the response standard deviations in the baseline runs. Each graph represents the across-subject mean in the response standard deviation within a run as a function of the target location. Separate baselines are shown for the three response methods of Exp. 1 (Closed-Eyes-Pointer, Open-Eyes-Pointer, Keyboard) and for the Closed-Eyes-Pointer method of Exp. 2. In all conditions, the standard deviations increased with target laterality. In the Closed-Eyes-Pointer condition the standard deviations were the largest, ranging from approximately 4.5° to 6°, and they varied the least with the target location (filled circles for Exp. 1 and asterisks for Exp. 2). In the Open-Eyes-Pointer condition, the response standard deviations improved only slightly and only for the targets at 11° to 45° (open circles). In the Keyboard condition the

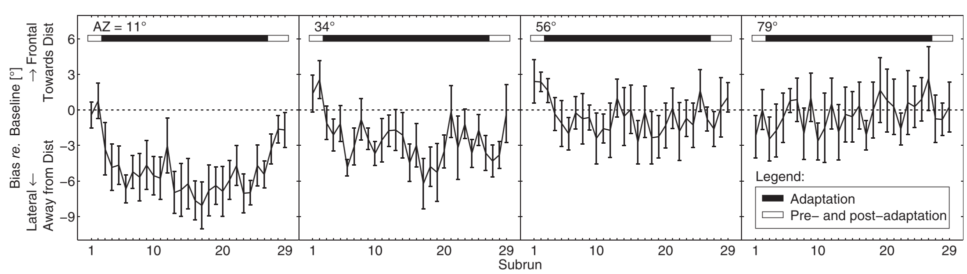

Fig. 2. Across-subject mean (±SEM) response bias as a function of subrun in Exp. 2. Each panel shows data for one target location. Data from the experimental runs were averaged across the Distractor-Target and Target-Distractor conditions, disregarding the trial type, and plotted re. baseline runs.

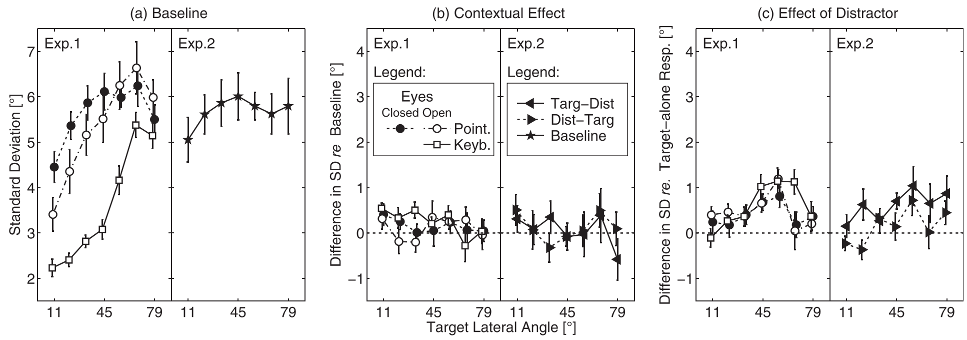

Fig. 3. Across-subject mean (±SEM) of response standard deviations as a function of the actual target location. The left-hand portion of each panel shows data from Exp. 1, the right-hand portion from Exp. 2. (a) Standard deviation in responses in the baseline runs consisting of target-alone trials. (b) Contextual effect defined as the standard deviation in target-alone responses in the experimental runs, re. target-alone responses, in the baseline runs [from panel (a)]. (c) Effect of the presence of distractor on standard deviation, computed as the difference between the response standard deviations for the distractor + target stimuli and the target-alone stimuli [from panel (b)] in the experimental runs.

response variability decreased considerably and the standard deviations ranged from 2° to 5° (open squares). Both main effects and the interaction were significant ($p < 0.001$).

Similar to the bias analysis, the standard deviation results again confirm that the Keyboard response method is superior to the Pointer methods, as the measured response variability decreased by 1° to 2° when this method was employed. The Closed-Eyes-Pointer method showed the largest overall variability for the frontal targets. It is likely that, compared to the Open-Eyes-Pointer method, this variability is strongly influenced by response method noise, as auditory representation-driven noise is expected to grow with target laterality (Perrott and Pacheco, 1989).

4.2 Contextual effect

Figure 3(b) plots the contextual effect defined as the difference in standard deviation in the responses on the target-alone trials from the experimental runs, re. baseline runs [from panel (a)]. No effect or a very small (less than 0.5°) increase due to context is observed in both experiments. These results suggest that the context does not affect response variability, a conclusion that could not be drawn from the Kopčo *et al.* (2007) data since no target-alone baseline was measured in that study.

4.3 Effect of distractor

Figure 3(c) plots the effect of distractor defined as the standard deviations in the responses on the distractor + target trials re. target-alone trials [from panel (b)] in the experimental runs. For all three response methods in Exp. 1, the distractor induced an increase of response standard deviation of up to 1° that reached maximum approximately in the middle of the response range and decreased toward the edges of the range [panel (c)]. In Exp. 2, the effect of distractor was negligible in the Distractor-Target condition while the effect was as large as 1° in the Target-Distractor condition, roughly independent of the target location (a significant main effect of condition, $p < 0.05$).

These results are consistent with the corresponding conditions of Kopčo *et al.* (2007). An unexpected result was the increase in variance caused by the distractor presented 400 ms after the target. Given that a bias toward the distractor was also observed in this condition [Fig. 1(c)], the effect might be caused by sensory memory. Specifically, a distractor presented after a target might interfere with the memory representation of the target, resulting in increased response bias and variation (Keen and Freyman, 2009).

5. Conclusions

The current study showed that the context-related shifts in localization observed in Kopčo *et al.* (2007) cannot be explained by non-auditory factors like absence of vision or the hand-held pointer response method used, or by factors related to the contextual task difficulty and the strategy that the subjects might have employed to cope with it. As an alternative, the effect might be a result of bottom-up adaptation of the auditory representation to the stimulus spatial distribution statistics (Dahmen *et al.*, 2010), integrated over a time window of tens of seconds (Freyman *et al.*, 1991). The results also showed that response variance was not influenced by the context in the tested conditions. Future studies need to be performed to fully characterize the effect.

The current study also introduced a new response method that allows a direct visual readout of the perceived target location, minimizing the influence of sensory-motor transformations involved in providing the response using pointers. Finally, it was shown that target localization can be influenced by a distractor presented 400 ms after the target, likely due to sensory memory interference or automatic attentional shifts.

Acknowledgments

This work was supported by the SRDA, Contract No. APVV-0452-12 and by the TECHNICOM project, ITMS: 26220220182, of the EU RDP. Ivan Surin, Marcel Sopko, and Rudolf Andoga assisted with the study.

References and links

Brungart, D. S., Rabinowitz, W., and Durlach, N. I. (**2000**). "Evaluation of response methods for near-field auditory localization experiments," Percept. Psychophys. **62**, 48–65.

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.

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

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

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.

Perrott, D. R., and Pacheco, S. (**1989**). "Minimum audible angle thresholds for broad-band noise as a function of the delay between the onset of the lead and lag signals," J. Acoust. Soc. Am. **85**, 2669–2672.

Razavi, B., O'Neill, W. E., and Paige, G. D. (**2007**). "Auditory spatial perception dynamically realigns with changing eye position," J. Neurosci. **27**, 10249–10258.

Snyder, J. S., and Weintraub, D. M. (**2013**). "Loss and persistence of implicit memory for sound: Evidence from auditory stream segregation context effects," Atten. Percept. Psych. **75**, 1059–1074.

Spence, C., and Driver, J. (**1998**). "Auditory and audiovisual inhibition of return," Percept. Psychophys. **60**, 125–139.