

The effect of velocity on auditory representational momentum

Arne Feinkohl

Cluster of Excellence Hearing4all, Animal Physiology and Behaviour Group, Department
for Neuroscience, School of Medicine and Health Sciences, University of Oldenburg,
D-26111 Oldenburg, Germany
arne.feinkohl@uni-oldenburg.de

Shannon M. Locke, Johahn Leung, and Simon Carlile^{a)}

The Auditory Neuroscience Laboratory, School of Medical Sciences, The Bosch Institute,
The University of Sydney, New South Wales 2006, Australia
sloc5887@uni.sydney.edu.au, Johahn.Leung@gmail.com, simonc@physiol.usyd.edu.au

Abstract: “Representational Momentum” (RM) is a mislocalization of the endpoint of a moving target in the direction of motion. In vision, RM has been shown to increase with target velocity. In audition, however, the effect of target velocity is unclear. Using a perceptual paradigm with moving broadband noise targets in Virtual Auditory Space resulted in a linear increase in RM from 0.9° to 2.3° for an increase in target velocity from 25°/s to 100°/s. Accounting for the effect of eye position also reduced variance. These results suggest that RM may be the result of similar underlying mechanisms in both modalities.

© 2014 Acoustical Society of America

PACS numbers: 43.66.Ba, 43.66.Lj, 43.66.Qp [QJF]

Date Received: February 7, 2014 **Date Accepted:** May 19, 2014

1. Introduction

The localization of moving objects plays an essential role for many species in their interaction with the natural environment. In the visual domain, the perception of a moving target's endpoint tends to be displaced in the direction of motion, a mislocalization effect termed “Representational Momentum” (RM). Over the past 30 years, visual studies have determined a large number of variables influencing RM,¹ and hypothesized several theories underlying this forward displacement.² Evidence that RM might also be present in the auditory domain³ preceded the first findings in the visual domain;⁴ however, the handful of studies investigating auditory RM (Refs. 5–9) have shown varied and conflicting results.

Visual RM has been found with a multitude of paradigms, and the effect of target velocity represents “one of the most robust influences” (Ref. 1, p. 828). For example, Hubbard and Bharucha¹⁰ showed nearly linear increases in visual RM with increases in horizontal and vertical target velocities. This relationship between RM size and target velocity, however, has not yet been clearly demonstrated in audition. Potentially salient differences between paradigms (see Ref. 7), as well as the large variation in the documented size of the auditory RM effect, complicates any comparisons between studies to determine the effect of velocity. Results by Perrott and Musicant³ do not indicate an effect of target velocity on auditory RM, but the large range of RM reported (−11° to 25°) might mask a potential relationship. Getzmann *et al.*⁶ found a smaller RM for the faster target velocity with continuous targets. In Schmiedchen *et al.*,⁹ RM increased with target velocity for targets moving toward the midline, while targets moving toward the periphery resulted in the opposite effect. Such divergent results make it difficult to describe a model underlying auditory RM.

^{a)} Author to whom correspondence should be addressed.

2. Methods

2.1 Participants

Eleven participants (3 female) took part in the experiment, 6 were naive with respect to the goal of the study. The participants' ages ranged from 20 to 55 yrs (median 25). Prior to the experiment, participants were informed of the experimental procedures and signed consent forms in accordance with ethics requirements approved by the University of Sydney Ethics Committee.

2.2 Setup

The experiment was conducted in a 15 m³ sound-attenuated, darkened room with an ambient noise level of less than 40 dB sound pressure level (SPL). Participants sat facing a semicircular wooden frame (radius of 1 m) with their head positioned at the center of the semicircle, and stabilized by a chin rest. A numerical scale (see Fig. 1) with a resolution of 1° was attached to the wooden frame and encompassed -70° to $+70^\circ$ azimuth. The scale was positioned at the eye level of the participants and illuminated from behind by a light-emitting diode (LED) strip. Gray fixation crosses ($3^\circ \text{ W} \times 3^\circ \text{ H}$) were printed on the scale at -25° , 0° , and $+25^\circ$. Below the scale, all tick-marks corresponding to even integers were labeled from 0 to 140.

At the beginning of each session, lasers were used to position the participant's head in the center of the semicircle, with their nose at 0° azimuth. Headphones (DT990, Beyerdynamic, Germany) presented the stimuli produced by the digital to analog interface (Fireface 400, RME, Germany). All stimuli were presented at a comfortable suprathreshold listening level (~ 60 dB SPL). A head-tracker (InertiaCube 4, InterSense, USA) on top of the headphones served to monitor the participant's head position during the experiment. If the head position deviated by more than $\pm 1^\circ$ in yaw, pitch, or roll, the participants were guided by a LED feedback device to return to the correct head position before they could start a trial.

2.3 Stimuli

All stimuli were generated using Virtual Auditory Space (VAS). This technique generates realistic, externalized moving sounds by using individualized Head Related Transfer Functions (HRTFs), which contain the participant's binaural cues (interaural time, phase, and level differences), as well as monaural spectral cues, resulting from spectral filtering by the pinna (for a review, see Ref. 11).

Prior to the experiment, HRTFs of 1° resolution were recorded in an anechoic chamber at a distance of 1 m from the participant as per the method described by Carlile and Blackman.¹² To generate moving VAS stimuli, broadband noise (300 Hz to 16 kHz, 5 ms cosine ramps) was convolved with the transfer functions for successive horizontal positions (e.g., see Ref. 13). Velocity was determined by the duration the sound played at each position in space along the trajectory. Moving targets started either at -90° or at $+90^\circ$ azimuth and moved toward the target eye fixation cross at 25, 50, or 100°/s. The target stopped at the fixation point with a randomized endpoint (jitter) of $\pm 2^\circ$ in 1° steps, which resulted in five possible endpoints per fixation point. Total target duration ranged between 0.63 and 4.68 s. To be able to subtract other

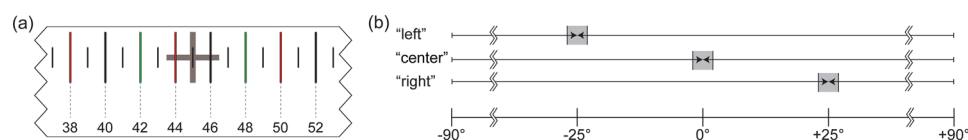


Fig. 1. (Color online) (a) The numerical scale. For purposes of illustration only a portion is shown. The gray cross represents the fixation point at -25° . The bars on the scale alternated between green, red, and black for greater distinguishability. For details, see text. (b) Verbal instruction for fixation point and corresponding stimulus trajectories for moving targets. The gray area represents the randomized jitter of the endpoint in 1° steps.

biases in localization, it was necessary to measure the participants' localization bias for stationary sounds.

Stationary targets had a duration of 0.2 s and were presented at the fixation point with a jitter of $\pm 8^\circ$ in 4° steps, which resulted in five possible positions per fixation point. The jitter was larger than in the trials with moving stimuli to encompass the range of endpoint values expected to be reported for the moving target, as well as to avoid providing participants information about the jittered endpoint locations for the moving sounds.

2.4 Procedure

Initially, all participants conducted a listening task that allowed them to practice attentionally tracking the auditory target while maintaining eye position. Participants then completed the experiment, which was comprised of one training session and five test sessions. Each session presented 46 moving trials: 10 warm-up trials and then 36 test trials in a random order comprised of 2 repeats of each unique combination of motion direction, fixation cross, and velocity ($2 \times 3 \times 3$) for a total of 10 repeats per combination. Forty stationary targets were presented after 36 test trials. To avoid motion after-effects, stationary and moving trials were not mixed.

Participants pressed a button to start a trial and then received a verbal instruction via headphones to fixate on one of the three gray crosses at 0° and $\pm 25^\circ$. The sound stimulus was presented 2 s afterwards. In trials with a moving target, participants were instructed to attend to the auditory motion trajectory. Participants verbally reported the location of the moving sound's endpoint by referring to the numerical scale. In trials with a stationary target, participants verbally reported the point on the numerical scale where the target was presented. No feedback was given during the entire experiment. Each session lasted about 10 min and the recorded verbal responses were then transcribed and matched with the stimulus for analysis. For each session, we first calculated the mean stationary bias for each fixation point and then corrected the responses for the moving stimuli correspondingly (see Ref. 6). Localization errors in the direction of motion are represented by positive values.

3. Results

All participants reported well externalized auditory stimuli and no difficulty in attentionally tracking the movement of the target. In a total of three trials across subjects, the verbal responses were inaudible and were excluded from analysis. One participant's results were bimodally distributed indicating a change in response strategy during the experiment which was confirmed on interview. Those data were not included in the analysis.

Using a Generalized Linear Mixed Model (GLMM), we analyzed the factors that influenced the localization of moving targets. The difference between the actual and perceived endpoint of the moving target was the dependent variable. The participants were treated as a random effect, fixed effects represented the target velocity (25°/s, 50°/s, 100°/s), the target direction (left, right), the fixation point (-25° , 0° , $+25^\circ$), the naivety of the participants (naïve, non-naïve), and all two-way interactions between fixed effects. The size of the RM displacement was only affected by target velocity ($p < 0.001$, GLMM analysis of variance), as shown in Fig. 2(a). Post hoc pairwise comparisons revealed a significantly larger displacement for 100°/s that reached a mean of 2.27° compared to 25°/s that reached a mean of 0.88° ($p < 0.001$, *t*-test, Bonferroni-corrected) and compared to 50°/s that reached a mean of 1.26° ($p = 0.003$, *t*-test, Bonferroni-corrected). The lack of an interaction between target direction and fixation point indicates no effect of trajectory length or target duration on RM displacement.

Auditory localization was found to be shifted toward the fixation point for both moving as well as stationary targets [Fig. 2(b)]. RM resulted in larger localization errors for the moving targets compared to the localization error for stationary targets,

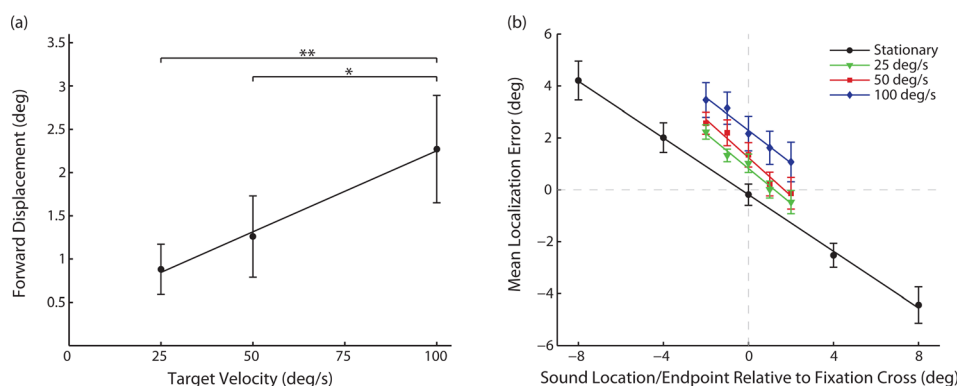


Fig. 2. (Color online) (a) Displacement in the perception of the point where the target stopped moving (means \pm SE) in relation to target velocity, illustrated with a linear fit ($R^2 > 0.99$). Asterisks indicate significant differences in post hoc pairwise comparisons. (b) Localization error (means \pm SE) in relation to the difference between target location (stationary sounds) or target endpoint (moving sounds) and the fixation cross. Positive localization errors represent mislocalization to the right of target locations or endpoints, negative localization errors represent mislocalization to the left of target locations or endpoints. To allow for a comparison between target motion directions, localization error for moving targets is defined as corresponding to a rightwards direction of motion. All linear fits reach $R^2 \geq 0.97$.

with larger localization errors for faster velocities [see also Fig. 2(a)]. The error in localization showed low variance for stationary targets [mean standard error (SE) = 0.58°] as well as for moving targets if the eye fixation point was taken into account (mean SE: 25°/s, 0.33° ; 50°/s, 0.49° ; 100°/s, 0.67°), indicating that the perception was highly reproducible between trials and sessions. For stationary targets presented at the fixation point, the mean localization error was 0.19° ($\pm 0.41^\circ$ SE), and the mean standard deviation across participants was 1.64° ($\pm 0.34^\circ$ SE).

4. Discussion

The effectiveness of the VAS method is demonstrated by a close relationship of the localization errors for stationary targets presented at the fixation points to previously measured free-field localization errors of the auditory system in the frontal region (e.g., Ref. 14). Also, the localization errors for stationary targets presented beyond the fixation points showed a low variance. These results confirm the consistent set of veridical target localization cues through VAS. For moving targets, localization was displaced in the direction of target motion for all three velocities tested, which is strong evidence for a RM effect.

Our results provide strong evidence that both stationary and moving targets were mislocalized toward the gray crosses that served as fixation points [Fig. 2(b)]. The gradients of the slopes for the linear fits are very similar for stationary and moving targets. Also, the distance between the fits for the moving targets resembles the effect of target velocity on the size of RM [see Fig. 2(a)]. Therefore, for moving targets, the effect of target offset position in relation to fixation point appears to be additive with the RM effect. When this effect of target offset position is accounted for [Fig. 2(b)], the variance in localization for the three target velocities shown in Fig. 2(a) is considerably reduced. The most likely explanation is that localization was biased toward eye position—although we did not measure eye position, subjects were instructed to hold fixation throughout the trial. Several auditory studies have demonstrated a localization bias toward eye position for stationary targets (e.g., Ref. 15). Except for Mateeff and Hohnsbein,⁵ previous studies on auditory RM did not provide participants with fixation targets so the effects of eye position remained uncontrolled. Measuring eye position in future studies on auditory RM would be useful in clarifying this effect.

Our results show a linear increase in auditory RM with target velocity, replicating results from multiple studies on visual RM.¹ The effect of target velocity on

visual RM and the range of RM size are similar to our auditory results. Most of these experiments used rotational stimuli rather than translational stimuli, which, to some extent, complicates comparisons with this study. One visual RM study¹⁰ used methodologies comparable to the present study and found visual RM to increase from approximately 0.8° to 2.2° for target velocities increasing from $5.8^\circ/\text{s}$ to $34.8^\circ/\text{s}$. In comparison to previous auditory studies on the effect of target velocity on RM, the range of RM size in our study is lower and previous results would not be predicted by the trend shown in Fig. 2(a). Unlike all recent studies, which employed action paradigms, we used a perceptual response measurement paradigm to avoid motor response biases. The lower range of RM size in our study is in line with the finding that action paradigms result in larger visual RM size than perceptual paradigms.¹

Results from stimuli moving toward the midline in the study by Schmiedchen *et al.*⁹ confirm the effect of target velocity found in this study. However, our results do not support the finding that stimuli moving toward the periphery resulted in an opposite effect of target velocity on RM size. The positive relationship between target velocity and RM size found in this study also contradicts the findings by Getzmann *et al.*,⁶ who showed that a slow $8^\circ/\text{s}$ continuous target resulted in larger RM than a faster $16^\circ/\text{s}$ continuous target. The authors hypothesized that the faster moving target may have been insufficient to engage the mechanism underlying RM because the shorter duration provided less spatial information. A subsequent study by Getzmann and Lewald⁸ found that a $16^\circ/\text{s}$ target with identical spectrum but much longer duration resulted in a nearly identical small RM size. This would indicate that target duration *per se* is not underlying the effect of target velocity on RM size in the study by Getzmann *et al.*⁶ Indeed, our study did not show an effect of target duration. However, spatial information may play a role in the mechanism underlying RM, which is limited for example by the spatial acuity of the auditory system in the region of target endpoints, or by the target spectrum. Compared to the study by Schmiedchen *et al.*,⁹ localization acuity was higher in regions where targets moving toward the periphery ended. Also, the current study is the first to provide the full range of localization cues by using broadband noise stimuli. Besides the differences in response measurement paradigm described above, these factors influencing spatial information may be underlying the difference in findings compared to the studies by Getzmann *et al.*⁶ and Schmiedchen *et al.*⁹ The similarities between RM in the present study and visual RM suggests that our stimuli were adequate in engaging the RM mechanism at all velocities examined. Future studies are needed to clarify the effect of spatial information on auditory RM.

This study indicates that the auditory system may possess a predictive mechanism for auditory motion. In addition, the similarity of our results with visual RM suggests that a similar or common mechanism, likely to be high-level (e.g., Refs. 2 and 8), could underlie both effects. If this were the case, then much could be learned about auditory motion processing from the significant body of knowledge regarding visual RM. Better characterization of auditory RM and its role in motion processing awaits further research.

Acknowledgments

We thank Heather Kelly for technical and logistical support. The study was partly funded by the Deutsche Forschungsgemeinschaft (SFB/TRR 31) and the Australian Research Council.

References and links

- ¹T. L. Hubbard, "Representational momentum and related displacements in spatial memory: A review of the findings," *Psychon. Bull. Rev.* **12**, 822–851 (2005).
- ²T. L. Hubbard, "Approaches to representational momentum: Theories and models," in *Space and Time in Perception and Action*, edited by R. Nijhawan and B. Khurana (Cambridge University Press, Cambridge, UK, 2010), pp. 338–365.
- ³D. R. Perrott and A. D. Musicant, "Minimum auditory movement angle: Binaural localization of moving sound sources," *J. Acoust. Soc. Am.* **62**, 1463–1466 (1977).

- ⁴J. Freyd and R. Finke, "Representational momentum," *J. Exp. Psychol. Learn. Mem. Cogn.* **10**, 126–132 (1984).
- ⁵S. Mateeff and J. Hohnsbein, "Dynamic auditory localization: Perceived position of a moving sound source," *Acta Physiol. Pharmacol. Bulg.* **14**, 32–38 (1988).
- ⁶S. Getzmann, J. Lewald, and R. Guski, "Representational momentum in spatial hearing," *Perception* **33**, 591–599 (2004).
- ⁷S. Getzmann and J. Lewald, "Localization of moving sound," *Percept. Psychophys.* **69**, 1022–1034 (2007).
- ⁸S. Getzmann and J. Lewald, "Constancy of target velocity as a critical factor in the emergence of auditory and visual representational momentum," *Exp. Brain Res.* **193**, 437–443 (2009).
- ⁹K. Schmiedchen, C. Freigang, R. Rübsamen, and N. Richter, "A comparison of visual and auditory representational momentum in spatial tasks," *Atten. Percept. Psychophys.* **75**, 1507–1519 (2013).
- ¹⁰T. L. Hubbard and J. J. Bharucha, "Judged displacement in apparent vertical and horizontal motion," *Percept. Psychophys.* **44**, 211–221 (1988).
- ¹¹S. Carlile, *Virtual Auditory Space: Generation and Applications* (RG Landes, New York, 1996).
- ¹²S. Carlile and T. Blackman, "Relearning auditory spectral cues for locations inside and outside the visual field," *J. Assoc. Res. Otolaryngol.* **15**, 249–263 (2014).
- ¹³S. Carlile and V. Best, "Discrimination of sound source velocity in human listeners," *J. Acoust. Soc. Am.* **111**, 1026–1035 (2002).
- ¹⁴S. Carlile, P. Leong, and S. Hyams, "The nature and distribution of errors in sound localization by human listeners," *Hear. Res.* **114**, 179–196 (1997).
- ¹⁵Q. N. Cui, B. Razavi, W. E. O. Neill, and G. D. Paige, "Perception of auditory, visual, and egocentric spatial alignment adapts differently to changes in eye position," *J. Neurophysiol.* **103**, 1020–1035 (2010).