Auditory motion parallax

William A. Yost[a,1]

When an object vibrates and produces an audible sound, many animals locate the position of the sound source based on only the sound. The horizontal (azimuth), vertical (elevation), and distance (range) of sound sources can often be determined, in some cases with high acuity (1). As most auditory systems have no spatial receptors and the physical properties of sound do not contain measures of extent (size and distance), the brain must process the sound in order for sound-source localization to occur. The study of sound-source localization has a long and rich history going back more than a century (2, 3). As a result, it might seem surprising that there can be new discoveries. But the PNAS article by Genzel et al. (4) has done just that.

In the late 1930s, Hans Wallach (5–7) published three papers in which he moved listeners and sound sources to test his main hypothesis that "Two sets of sensory data enter into the perceptual process of localization, (1) the changing binaural cues and (2) the data representing the changing position of the head" (7). Differences in the arrival time of sound at the two ears and interaural level differences were the two "binaural cues" assumed for determining sound-source location. Wallach (7) argued that vision, kinesthetic, and vestibular function provided head position cues. He had listeners rotate in a chair along the azimuth plane, judging the location of sounds presented from different azimuthal loudspeakers. Using 3D geometry and trigonometry, Wallach showed that certain rotational relationships between listeners and sound sources could disambiguate front–back reversal errors, produce azimuthal illusions, and allow listeners to judge sound-source elevation. This work influenced E. G. Boring, the famous psychologist at the time when Boring (2) summarized his history of sound-source localization with, "Wallach has made it quite clear that localization is not purely auditory, but the product of an integration of auditory, kinesthetic and, when the eyes are open, visual factors." For reasons that are not well established, there was almost no follow-up to evaluating Wallach's (7) hypothesis that sound-source localization is based on the integration of two cues (auditory-spatial and head-position cues). Several studies over the past decade have returned to Wallach's papers, delving more deeply into his hypothesizes (e.g., refs. 8–11). All of this work has involved azimuthal and elevated sound-source localization judgments when listeners and sound sources move. This work has indicated both the considerable strength of some of Wallach's ideas and some of the weaknesses. What is missing from this literature is a role of listener and sound motion in distance perception. The study by Genzel et al. (4) starts to fill in that gap in important ways.

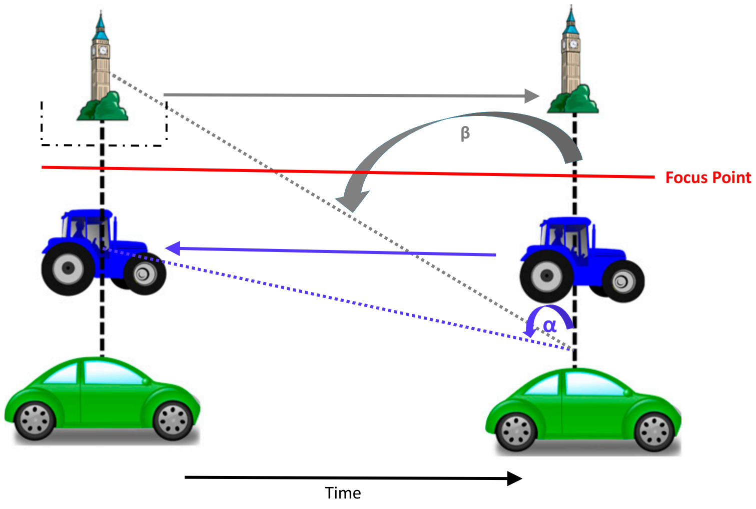

Fig. 1. Someone in the car on the left would view a tractor or a bell tower differently than when the car is on the right: that is, angle β is smaller than α, an example of parallax. If the car moved from left to right, the nearby tractor is likely to be perceived as moving faster past the car than the far-away bell tower, an example of motion parallax. With a focus point, near sources appear to move in a direction opposite that of the observer compared with far sources. And, near sources might occlude far sources when the observer is in one positon (*Left*) compared with another position (*Right*). If the sources only produced sound and could not be seen, then only motion parallax is likely to provide depth/distance information.

[a]Speech and Hearing Science, Arizona State University, Tempe, AZ 85287

Author contributions: W.A.Y. wrote the paper.

The author declares no conflict of interest.

See companion article on page 4264.

[1]Email: William.yost@asu.edu.

Published online April 5, 2018.

There are many examples of spatial information from one sensory system biasing the perception of spatial location in another system [e.g., the ventriloquism effect in which the perceived location of a sound is "captured" by the location of a visual image occurring at about the same time (12)]. Within a sense, one spatial location can bias the spatial perception of a different spatial location. For example, Teramoto et al. (13) showed that movement toward a sound source changes the perceived location of the sound source. Genzel et al.'s (4) study and the other recent research based on Wallach's (7) earlier work indicate that the actual ability to locate sound sources depends on an integration of different cues, not just one cue biasing perception based on another cue. Genzel et al. (4) provide strong evidence that auditory motion parallax, like the visual analog, allows one to judge relative sound-source distance.

Parallax is when the position or direction of an object appears to differ when viewed from two different locations. In Fig. 1, the position of the tractor or the bell tower would appear different if viewed from the car on the left compared with the car on the right (the angle β is smaller than the angle α). If the car moves from left to right, then motion parallax will cause the nearby tractor to appear to move faster past the car than the far-away bell tower (the change in the angle over time is faster for the tractor than for the bell tower). Thus, motion parallax leads to slow-moving objects being perceived as further away than fast-moving objects. In vision, if one focuses the retina between the tractor and bell tower as the car moves, the near-by tractor will move opposite to the car direction, but the bell tower will move in the same direction. Thus, direction of motion can indicate near or far objects. A near object (tractor) may occlude a far object (bell tower) in one (Fig. 1, *Left*), but not in another (Fig. 1, *Right*) condition. Occlusion might be a cue to judge object distance. In vision, parallax offers several possible cues for depth perception. If only the sound of the objects were used, then motion parallax, but only motion parallax, might provide a cue for distance perception (i.e., when listeners move, sound from near sound sources might appear to move faster across space than sound from far sound sources).

Genzel et al. (4) tested whether motion parallax could be a cue for relative sound source distance judgments. To do so, they needed a paradigm in which the only cue for judging relative distance was auditory motion parallax: that is, all other auditory distance cues were controlled for. While distance perception has not received a lot of attention in the literature (14), most research indicates that there are several possible cues for judging sound-source distance: (*i*) Expectation/experience, with a priori knowledge about the sound from a sound source, softer sounds will be further away than louder sounds. (*ii*) Sound from reflective surfaces (in a room) will reach listeners ears slightly after the direct sound arrives. The ratio of the direct-to-reflective sound level can indicate relative sound-source distance. (*iii*) When distances between sound source and listeners are large, the atmosphere reduces the level of sound before it reaches the ear in a frequency-dependent manner, so changes in a sound's spectrum can indicate relative distance. In most cases human listeners are not very accurate in making absolute distance judgments, generally underestimating the true distance of a sound source (14).

Genzel et al. (4) first show in a simple experiment that when listeners rotated their heads, they could differentiate a far sound source from a near source better than when they did not move their heads. The results were clear but variable, and cues other than the motion parallax caused by head rotation may have allowed for relative distance judgments. They then developed a clever "virtual" auditory motion parallax procedure in which the only possible cues for relative distance judgments were differences due to auditory motion parallax. Listeners judged the relative distance of one of two different pitches (i.e., was a high-pitch sound perceived closer or further than a low-pitch sound). Listeners either moved themselves (active movement) or they were moved in a chair (passive movement) past an array of virtual sound sources. All 12 listeners showed increasing percent correct distance judgments with increasing distance between the near and far sound source, with a smaller distance required for a threshold distance judgment in the active- compared with the passive-movement condition (although this difference was small). A control condition was run in which the sounds were moved virtually but the listeners remained stationary, to rule out the possibility that relative distance judgments were somehow associated with just a moving sound. Listeners' performance in this control experiment was near chance, strongly implicating motion parallax as being responsible for the relative distance judgments.

Genzel et al. provide strong evidence that auditory motion parallax, like the visual analog, allows one to judge relative sound-source distance.

While the experiments of Genzel et al. (4) are more simulations than direct investigations of auditory motion parallax, they were well done and clearly demonstrate that it is highly possible that humans can use motion parallax to judge relative sound-source distance. In addition, the studies are another in a series of recent research showing that sound-source localization requires, as Wallach (5) argued, the interaction of binaural cues and head position cues. The research also adds to the growing list of examples showing how listener and sound-source motion affect sound-source localization perception. A weakness of the Genzel et al. (4) study is the very small difference in performance between the active- and passive-motion conditions. The work of Wallach (5–7) and subsequent research suggest that there should be a difference in that there would appear to be fewer cues available concerning head motion in the passive condition than in the active condition, and as a consequence sound-source localization performance should be worse in the passive condition. The fact that there was not a clear difference means either that the role of head-motion cues is overstated (unlikely given the literature) or that the conditions used by Genzel et al. (4) did not sufficiently alter the amount of head-position information between the passive- and active-motion conditions. Given the small amount of movement and the possibility of learning, the latter seems a likely explanation. In any event, the role of different head-motion cues in auditory motion parallax needs more study.

While the results of Genzel et al. (4) indicate that listeners can use motion parallax to make relative distance judgments, what is not known is the extent to which auditory motion parallax is actually used to make sound-source distance judgments or to assist in navigation in ways that are similar to how parallax informs visual processing. In making comparisons to vision, it is crucial to recognize the many differences between sound and light, and between the auditory and visual systems. As Fig. 1 points out, there are many visual parallax cues for judging depth, but probably only one such cue using sound (i.e., auditory motion parallax). Thus, auditory motion parallax might be too weak to provide much of a benefit for judging relative sound-source distance in the actual world.

In real environments motion parallax is usually symmetrical as one navigates in something like a hallway. But motion parallax can be made asymmetrical in virtual environments, leading to errors in observer motion while navigating in a virtual world. The results of Genzel et al. (4) suggest that maybe alterations in auditory motion parallax in a virtual world could also influence listener motion. Thus, the work of Genzel et al. (4) suggests a wide range of interesting additional experiments that might reveal crucial information about sound-source localization perception in real and virtual spaces.

Acknowledgments

The author's research is supported by grants from the National Institute on Deafness and the Other Communication Disorders (NIH Grant 5R01DC015214) and by Oculus VR LLC.

1 Blauert J (1997) *Spatial Hearing: The Psychophysics of Human Sound Localization* (MIT Press, Cambridge, MA).
2 Boring EG (1942) *Sensation and Perception in the History of Experimental Psychology* (Appleton-Century-Croft, New York).
3 Yost WA (2017) History of sound source localization: 1850–1950. *Proceedings of Meetings on Acoustics* 30:1–5.
4 Genzel D, Schutte M, Brimijoin WO, MacNeilage PR, Wiegrebe L (2018) Psychophysical evidence for auditory motion parallax. *Proc Natl Acad Sci USA* 115:4264–4269.
5 Wallach H (1938) Ueber die wahrnehmung der schallrichtung. *Psychol Forsch* 22:238–266.
6 Wallach H (1939) On sound localization. *J Acous Soc Am* 10:270–274.
7 Wallach H (1940) Role of head movements and vestibular and visual cues in sound localization. *J Exp Psychol* 27:339–368.
8 Macpherson EA (2011) Head motion, spectral cues, and Wallach's 'Principle of least displacement' in sound localization. *Principals & Applications of Spatial. Hearing*, eds Suzuki Y, et al. (World Scientific, Hong Kong).
9 Brimijoin WO, Akeroyd MA (2014) The moving minimum audible angle is smaller during self motion than during source motion. *Front Neurosci* 8:273.
10 Yost WA, Zhong X, Najam A (2015) Judging sound rotation when listeners and sound rotate: Sound source localization is a multisensory process. *J Acoust Soc Am* 38:3293–3308.
11 Genzel D, Firzlaff U, Wiegrebe L, MacNeilage PR (2016) Dependence of auditory spatial updating on vestibular, proprioceptive, and efference copy signals. *J Neurophysiol* 116:765–775.
12 Alais D, Burr D (2004) The ventriloquist effect results from near-optimal bimodal integration. *Curr Biol* 14:257–262.
13 Teramoto W, Cui Z, Sakamoto S, Gyoba J (2014) Distortion of auditory space during visually induced self-motion in depth. *Front Psychol* 5:848.
14 Kolarik AJ, Moore BC, Zahorik P, Cirstea S, Pardhan S (2016) Auditory distance perception in humans: a review of cues, development, neuronal bases, and effects of sensory loss. *Atten Percept Psychophys* 78:373–395.