Auditory Distance Perception by Translating Observers

Jon M. Speigle and Jack M. Loomis

Department of Psychology
University of California
Santa Barbara, CA 93106

Abstract

We consider auditory distance perception of a moving observer and its relevance for the perception of stationary and moving sources. We begin with a review of some of the acoustic cues to source distance, focusing on the dynamic cues available under observer translation (motion parallax and acoustic tau). We report an experiment indicating the level of accuracy with which stationary and translating observers are able to localize stationary sources from 2 to 6 m away. Given the significant errors associated with these near distances, it would appear that the perceptual assessment of the 3-D trajectory of a real or virtual source, especially a distant one, is likely to be substantially in error. Even so, motion parallax and acoustic tau are informative about the relative motion between observer and source.

1.0 Introduction

Within the last several years, a number of virtual acoustic displays have come into existence [1], thus opening up a range of interesting possiblities for effectively displaying information to human operators in ways not possible before, not to mention their potential for furthering basic research on auditory space perception [2, 3]. Wenzel [1] reviews a number of these possible applications, some involving non-mobile users and others, mobile users. An example of the latter is providing pilots with an auditory representation of other aircraft in the vicinity, including those that are potential collision targets.

Another potential application is as part of the user interface of a navigation aid for the visually impaired [3, 4]. A test bed for such a system is now in the development stage by the second author and his colleagues [5] and will be used to assess the potential of such a system and to identify the major obstacles to its success. As one display option, we will be evaluating a virtual acoustic display. With such a display, the navigation system will lead the traveler along specified routes by means of virtual auditory beacons appearing perceptually in front of the traveler; furthermore, the system will indicate the positions of landmarks by having their labels, spoken by speech synthesizer, appear as virtual sounds at their correct locations within the auditory space of the traveler.

Critical to developing a sense of surrounding 3-D space and of moving targets within it is the auditory perception of distance. Most of the research on the auditory localization of emitted sound has been concerned with directional localization (e.g., [6-9]) and will not be discussed here. At a minimum, the auditory perception of distance involves externalization of the sound (perceiving the sound extracranially). This is an important issue for virtual acoustics and one that has received considerable attention over the years [1, 3, 10-16] . Assuming that the display is effective enough to promote externalization, the important question is whether sounds at different distances can be localized with any accuracy.

1.1 Static Distance Cues

A number of potential distance cues have been identified, some informative about absolute (egocentric) distance to a source and some informative only about the relative distances of two sources [17]. For an observer whose head is stationary, there are several potential monaural distance cues. The most important of these is sound level, which in an anechoic environment falls off by 6 dB with each doubling of distance [18, 19]. If source intensity is known to the listener, then familiar sound level can serve as an absolute distance cue, but if source intensity is unknown, sound level can serve only as a relative distance cue. Studies by Mershon and King [17] and others (cited in [18]) have demonstrated that the sound level of unfamiliar sounds does contribute to auditory perceived distance. Other static monaural cues are spectral content and the proportion of reverberant to direct sound, both of which have been shown to influence perceived distance [10, 17, 20-25].

0-8186-4910-0/93 $3.00 © 1993 IEEE

Going from monaural to binaural listening (with stationary head) adds a potential cue [26], but Molino [27] has shown it to be ineffective when a tone source was used. Head rotations provide still additional range information [28], but Simpson and Stanton [29] showed that head rotations do not improve the accuracy of distance judgments.

1.2 Dynamic Cues Resulting from Translation

Translation of the observer greatly increases the distance information available (subject to several assumptions). In analogy with what is known about optic flow and visual space perception (e.g., [30, 31]), at least two distinct types of "acoustic flow" information about distance can be identified [32]. One, absolute motion parallax, is the change in angular direction of a point source occasioned by the observer's translation. As can be seen in Figure 1, a displacement S of the observer causes the stationary source S_1 (assumed to be at ear-level) to undergo the depicted change in azimuth from α_1 to α_2. As is true of any method of triangulation, if the observer can assume the source to be stationary, can sense the change in azimuth of the source during the displacement, and knows the size of the displacement (or velocity), the observer can in principle compute the distance D of the source:

$$D = \frac{S \sin(\alpha_2)}{\sin(\alpha_2 - \alpha_1)}$$

What happens if the assumption of stationarity is relaxed? Figure 1 shows two other sources as well, S_2 moving concomitantly with the observer and S_3 moving concomitantly in the opposite direction. If these sources were to move such that their changes in azimuth were precisely the same as that for S_1, then the observer would have no means of distinguishing between them using azimuthal changes alone. (It is also the case that the variation in sound level is the same for the three cases, although absolute level differs.) This is to say that different combinations of distance and motion trajectory give rise to identical acoustic flow, a point well recognized in the analogous case involving optic flow [30, 33, 34]. Thus, perceiving the distance of a target from parallax alone is possible only under the assumption of stationarity; the more difficult problem of sensing the 3-D trajectory of a moving sound source is impossible without additional constraints.

Distance perception using motion parallax also requires that the observer be able to sense the changing azimuth of the target. Clearly if the observer's head were to rotate at the same time that it is translating, the changing azimuth of the source relative to the head would not by itself be informative about source distance. Again in analogy with the optic flow literature [35], motion parallax can be used to infer the distance of a stationary source only if the observer can partial out the changes in azimuth due to head rotation.

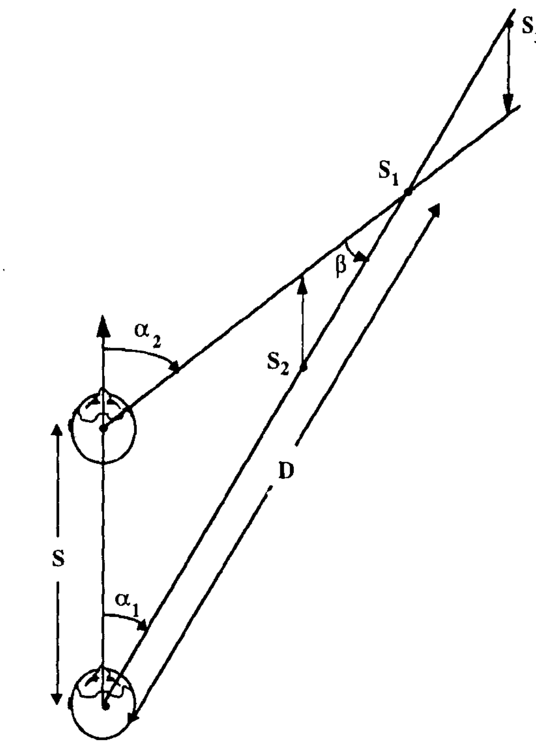

Figure 1. Stationary source S1 undergoes a change in azimuth (from α_1 to α_2) as the observer translates through a distance S. By assuming the source to be stationary, an observer can use the change in azimuth (motion parallax) to compute the source distance D. Note, however, that moving sources S_1 and S_2 produce the same motion parallax.

Finally, accurate perception of the distance of a stationary source, even when observers can partial out azimuthal changes due to rotation, requires that the observer know his/her own velocity (or displacements over time). Evidence of accurate perception of actively-controlled walking comes from studies on visually-directed action [36-39].

The other "acoustic flow" cue that is informative about source distance is "acoustic tau" [32, 40-43], an analog of the optic variable, tau, which is the ratio of the angular extent of a target divided by its angular rate of expansion [31]. When an observer and visible target are closing at a constant rate toward an eventual collision (observer and/or target is moving), optic tau specifies time-to-collision. Acoustic tau at a given moment is expressible as

$$\tau = \frac{2I}{dI/dt},$$

where I is sound intensity at the observer [41]. Like optic tau, acoustic tau specifies time-to-collision for a constant closure velocity between observer and source. Acoustic tau is informative about source distance only if the observer knows the closure velocity.

Ashmead and LeRoy [40] demonstrated that acoustic tau can be used by observers in perceiving the distance of stationary targets. In a control condition, observers heard noise bursts from sources ranging in distance from 5 to 19 m while standing with the head stationary at an origin. After extinction of the sound, observers attempted to walk blindfolded without additional auditory cues from the origin to the location where the source had been. This "open loop walking" response has previously been used as an effective means of assessing visually perceived distance [37-39].

In the movement conditions, the observer walked along a 1.5 m line that terminated at the origin of stationary listening. The sounds were presented while the observer walked blindfolded along this path, which was collinear with the source. After extinction of the sound, the observer executed the same open-loop response as in the stationary listening condition. Sound level was reduced in its reliability as a cue by randomly varying source intensity from trial to trial. Lacking absolute sound level as a reliable cue, the nine adult observers showed considerable improvement in the accuracy of their walking responses from the control to the tau condition. This result indicates that acoustic tau is used by observers as a cue to distance. Mean walking performance in the tau condition was quite accurate--mean signed error was less than 10% of source distance except for the nearest distances where overshooting of about 20% occurred.

The experiment we report is quite similar to that of Ashmead and LeRoy but differs in several important respects. We varied the initial source azimuth (to permit use of acoustic tau, motion parallax, or both in the translation conditions), source intensity was not varied from one trial to the next (making sound level a reliable distance cue), and source distance ranged from only 2 to 6 m. Some reverberation was present under these conditions.

2.0 Experiment

2.1 Method

The experiment was conducted outdoors in a grassy area approximately 20 m from the nearest building. Ambient noise was minimized by conducting the experiment during off hours. As depicted in the top panel of Figure 2, sources were located at distances of 2, 4 and 6 m from the "origin" and at azimuths of 0°, 30°, 60°, and 90°, relative to the initial facing direction of the observer.

The blindfolded observers were either stationary during stimulus presentation or allowed to walk 2 or 4 meters. In the movement conditions, observers were positioned approximately 3 or 5 meters from the origin and walked toward the origin, during which the stimulus was presented for either the final 2 or 4 m, respectively. The observer lightly grasped a guide rope while walking and stopped when the sound terminated. In the stationary condition, the observer was positioned at the origin, facing the 0° source axis. In all conditions, after the sound had been extinguished and the observer was standing still, the observer put on a pair of sound-attenuating ear protectors and then attempted to walk to the source. The distance and direction of the observer's terminal position relative to the origin were measured by an ultrasonic range finder and electronic compass, which were mounted on a tripod.

The stimulus used was a 20 Hz pulsetrain, a sound that is highly localizable in direction. The signal was amplified and played through a 9 cm speaker mounted at ear level. Because of the directionality of the speaker, it was necessary for an assistant to continually aim the speaker at the observer on the movement trials using a tripod and sighting mechanism. The ambient sound level averaged 37 dB (A weighting) and the stimulus plus ambient background measured 45 dB (A weighting) at a distance of 1 m.

Ten observers participated in the experiment for payment. Five were undergraduates, and five were graduate students. All reported having normal hearing and were naive as to the purpose of the experiment.

Observers were blindfolded before approaching the workspace and remained so throughout the experiment. The observer practiced walking along the rope until able to walk at a normal pace. Five practice trials were given: three moving and two stationary. Source pressure was the same as on the actual trials; however, the source was positioned randomly at locations different from those used in the experiment proper. No feedback was given as to how well the observer had localized the source.

In the experiment proper, each subject participated in 36 randomized trials, representing 12 source locations crossed with the three conditions. The experimenter was located at the sound generating equipment and the assistant, at the speaker tripod. The experimenter positioned the observer (still wearing the sound attenuators) at the starting position on the guide rope. After the assistant had positioned the source at one of the locations, the observer was signaled to remove the attenuators and informed as to whether the trial was stationary or moving. On stationary trials, the stimulus was presented for 3.0 sec. On movement trials, the observer started walking along the guide rope on "go." The stimulus turned on when the observer reached the 2 or 4 m marker and turned off when the observer passed through the origin. The observer stopped walking after the stimulus had been extinguished and put on the attenuators.

The observer paused for approximately 1 sec after donning the attenuators, giving the assistant time to move the speaker out of the way. The observer then walked to the perceived location of the source. The experimenter placed the measurement tripod at the origin and measured the distance and direction of the observer. At this point, the observer was led passively to the start point of the next trial. The observer received no feedback of any kind about his/her performance.

The distance that the observer had walked past the origin and the time elapsed between stimulus onset and the observer reaching the origin were also recorded. The mean overshoot was 76 cm in the 2m condition, and 62 cm in the 4m condition. Stimulus duration on the stationary trials was approximately 3 sec. The mean durations on the 2 and 4 m movement trials were 1.96 and 3.48 sec, respectively.

2.2 Results

The lower three panels of Figure 2 give the results for the stationary, 2 m, and 4 m conditions, respectively. The source locations are given by the solid symbols in correspondence with the upper panel. The centroids of the terminal locations of the walking responses are given by the unfilled symbols. The error bars represent the standard errors of the mean of the x and y coordinates of these terminal locations. Note that when the source was positioned along the 0° line, acoustic tau alone was available as a dynamic distance cue. For all other directions, both motion parallax and acoustic tau information were available.

For all three panels and azimuths, the ordering of the centroids in distance from the origin is correct with respect to that of the source locations. The figure indicates reasonable agreement between the target locations and the corresponding response centroids. Also apparent in the two movement conditions for non-zero azimuths is a clear indication of directional error; this surely is a consequence of overshooting the origin.

Figure 3 gives the mean walked distances to the three target distances, averaged over azimuths and observers. These walked distances were measured from the terminal

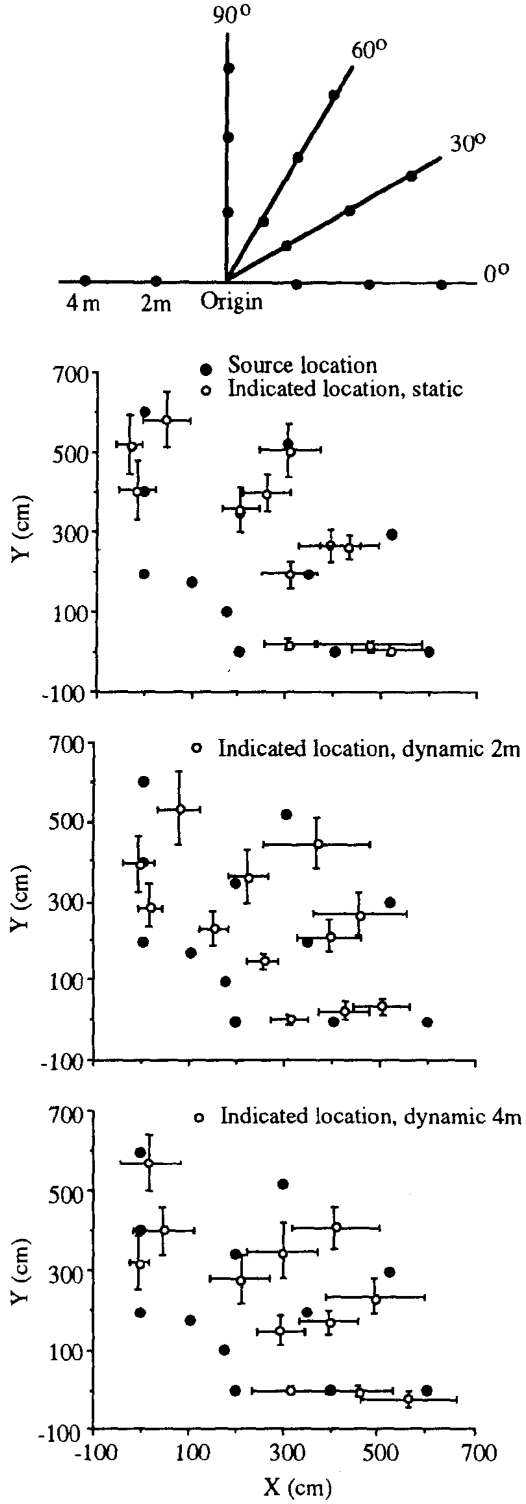

Figure 2. Experimental results. The top panel gives the configuration of source locations. The lower three panels give the mean indicated locations in the three conditions.

point of each response to the initial point (which coincided with the origin only in the stationary condition). The solid curves represent power functions fit to the data of the three conditions; the uppermost curve is for the stationary condition and the lowermost curve is for the 2 m walking condition.

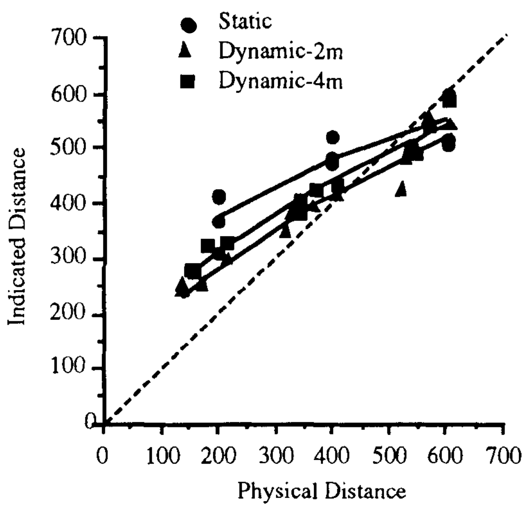

Figure 3. Mean indicated distance for the three conditions, averaged over source azimuth.

Figure 3 indicates a clear tendency to overshoot the near targets and to undershoot the far targets. A within-subjects ANOVA revealed a modest effect of condition, $F(2,18) = 3.98$, $p < .05$, which shows up in the figure mostly as a tendency to walk slightly farther in the stationary condition, especially at the near distances. The only other significant effect was that due to target distance, $F(3,27) = 36.59$, $p < .001$. The failure to find an interaction between condition and azimuth means that there is no clear differential effect of acoustic tau vs. motion parallax.

3.0 Discussion

Performance was only slightly better in the two walking conditions than in the stationary conditions signifying that the dynamic cues of motion parallax and acoustic tau supplement only slightly the static cues of sound level and reverberation that were present in the experimental setting. Ashmead and LeRoy [40] obtained a larger effect of observer motion (where only acoustic tau was available as a dynamic cue), but their observers had available a less reliable sound level cue in the two conditions.

The accuracy of auditory distance perception, as indicated by open-loop walking in the present experiment and in that by Ashmead and LeRoy [40], falls short of what is observed when observers perform the same open-loop walking response (without vision) after viewing targets binocularly; in the vision experiments, systematic error is on the order of tens of centimeters for targets ranging from 4 to 22 m [37-39], and the accuracy of the response to individual targets by individual observers is much better than it is in these auditory experiments. The presence of significant error at these short distances suggests that accurate perception of far source distances is unlikely to ensue with virtual displays simulating natural sounds, even when observers know that the virtual sounds represent stationary landmarks. However, pessimism is not completely warranted, for it is possible that virtual displays can augment natural distance cues with coding by pitch or familar speech sounds (whispering, shouting, etc; [44]) or that observers will be able to learn to make effective use of the dynamic cues under the assumption that the virtual sources are stationary.

Using virtual acoustic displays to portray moving targets to a moving observer becomes problematic if distance to the virtual source is not correctly perceived. With respect to visual space perception, Gogel [34, 45] has shown that misperceiving the distance of a visual target or misperceiving one's own motion has predictable effects on the perceived 3-D trajectory. Both types of error can make stationary targets appear to move, can make moving targets appear stationary, and can induce misperception of the 3-D trajectories of moving targets. As Mershon and Hutson [46] and Cox [47] have demonstrated, Gogel's analysis applies to audition as well; thus, misperceptions of the trajectories of stationary and moving sound sources are to be expected on the basis of errors in auditory distance perception and errors in perceived self-motion. For example, Figure 4 shows that if a stationary source is perceived to be closer than it actually is and if the observer correctly senses the changing azimuth of the source and perceives his/her self-motion correctly, then the apparent source will appear to move along with the observer; similarly, if the source appears more distant than it is, it will appear to move in a direction opposite to the observer. (Some of the apparent motions depicted in Figure 4 correspond to the physical motions of the sources in Figure 1 that produce the same binaural stimulation as the stationary source.)

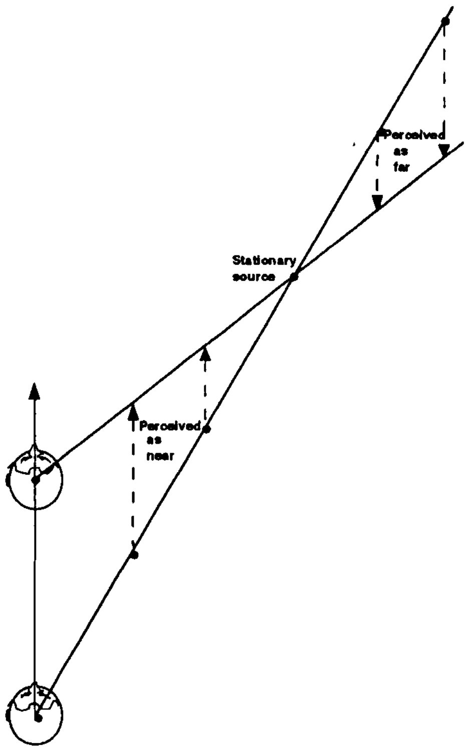

Figure 4. A translating observer misperceiving the distance of a stationary source will perceive the source as moving [34]. If the distance is underperceived, the perceived motion will be in the same direction as the observer motion. If the distance is overperceived, the perceived motion will be in the opposite direction.

Even though the above analysis suggests a somewhat pessimistic outlook for correctly perceiving the absolute 3-D trajectory of a moving source, a moving observer still may be able to accurately assess aspects of the relative motion between self and source. For the purposes of collision avoidance, for example, the motion parallax of the source is sufficient to indicate whether collision will occur—the necessary and sufficient condition for collision between the observer and a target, when both are moving with unaccelerated motion, is constancy of the angular direction of the source, a result known from the optic flow literature [48]. Similarly, acoustic tau is informative about time to collision even when the relative distance and relative velocity are unknown. Thus, some tasks that at first glance seem to require that the observer perceive source distance, in fact do not, making them candidates for the effective application of virtual displays.

Acknowledgments

This research was supported by Grant R01 EY09740 from the National Eye Institute. The authors thank Donald Mershon and Daniel Ashmead for very helpful discussion as well as for providing their research articles in pre-publication form.

References

1. Wenzel, E.M., "Localization in virtual acoustic displays", *Presence: Teleoperators and Virtual Environments*, **1**, 1992, 80-107.

2. Durlach, N., "Auditory localization in teleoperator and virtual environment systems", *Perception*, **20**, 1991, 543-554.

3. Loomis, J.M., Hebert, C., and Cicinelli, J.G., "Active localization of virtual sounds", *Journal of the Acoustical Society of America*, **88**, 1990, 1757-1764.

4. Golledge, R.G., Loomis, J. M., Klatzky, R. L., Flury A. and Yang, X., "Designing a personal guidance system to aid navigation without sight: progress on the GIS component", *International Journal of Geographic Information Systems*, **5**, 1991, 373-395.

5. Loomis, J.M., Golledge, R.G., and Klatzky, R.L., "Personal guidance system for the visually impaired using GPS, GIS, and VR technologies", in *Conference on Virtual Reality and Persons with Disabilities*. Millbrae, CA, June 17-18: 1993.

6. Blauert, J., *Spatial Hearing*. Cambridge, MA: MIT Press, 1983.

7. Durlach, N.I. and Colburn, H.S., "Binaural phenomena", in *Handbook of Perception*, E.C. Carterette and M.P. Friedman (Eds.), New York: Academic Press, 1978, 365-466.

8. Middlebrooks, J.C. and Green, D.M., "Sound localization by human listeners", *Annual Review of Psychology*, **42**, 1991, 135-159.

9. Mills, A.W., "Auditory localization", in *Foundations of Modern Auditory Theory, Vol. 2*, J.V. Tobias (Eds.), New York: Academic Press, 1972, 303-348.

10. Durlach, N.I., Rigopoulos, A., Pang, X.D., Woods, W.S., Kulkarni, A., Colburn, H.S., and Wenzel, E.M., "On the externalization of auditory images", *Presence: Teleoperators and Virtual Environments*, **1**, 1992, 251-257.

11. Koenig, W., "Subjective effects in binaural hearing", *Journal of the Acoustical Society of America*, **22**, 1950, 61-62.

12. Levy, E.T. and Butler, R.A., "Stimulus factors which influence the perceived externalization of sound presented through headphones", *Journal of Auditory Research*, **18**, 1978, 41-50.

13. Loomis, J.M., "Distal attribution and presence", *Presence: Teleoperators and Virtual Environments*, **1**, 1992, 113-119.

14. Plenge, G., "On the differences between localization and lateralization", *Journal of the Acoustical Society of America*, **56**, 1974, 944-951.

15. Toole, F.E., "In-head localization of acoustic images", *Journal of the Acoustical Society of America*, **48**, 1970, 943-949.

16. Wallach, H., "The role of head movements and vestibular and visual cues in sound localization", *Journal of Experimental Psychology*, **27**, 1940, 339-368.

17. Mershon, D.H. and King, L.E., "Intensity and reverberation as factors in the auditory perception of egocentric distance", *Perception & Psychophysics*, **18**, 1975, 409-415.

18. Coleman, P.D., "An analysis of cues to auditory depth perception in free space", *Psychological Bulletin*, **60**, 1963, 302-315.

19. White, F.A., *Our acoustic environment*. New York: Wiley & Sons, 1975.

20. Butler, R.A., Levy, E.T., and Neff, W.D., "Apparent distance of sounds recorded in echoic and anechoic chambers", *Journal of Experimental Psychology: Human Perception and Performance*, **6**, 1980, 745-750.

21. Coleman, P.D., "Dual role of frequency spectrum in determination of auditory distance", *Journal of the Acoustical Society of America*, **44**, 1968, 631-632.

22. Little, A.D., Mershon, D.H., and Cox, P.H., "Spectral content as a cue to perceived auditory distance", *Perception*, **21**, 1991, 405-416.

23. Mershon, D.H., Ballenger, W. L., Little, A. D., McMurtry, P. L., and Buchanan, J. L., "Effects of room reflectance and background noise on perceived auditory distance", *Perception*, **18**, 1989, 403-416.

24. Mershon, D.H. and Bowers, J.N., "Absolute and relative cues for the auditory perception of egocentric distance", *Perception*, **8**, 1979, 311-322.

25. Sheeline, C.W., *An investigation of the effects of direct and reverberant signal interaction on auditory distance perception*, Doctoral Thesis, Stanford University, 1983.

26. Hirsch, H.R., "Perception of the range of a sound source of unknown strength", *Journal of the Acoustical Society of America*, **43**, 1968, 373-374.

27. Molino, J., "Perceiving the range of a sound source when the direction is known", *Journal of the Acoustical Society of America*, **53**, 1973, 1301-1305.

28. Lambert, R.M., "Dynamic theory of sound-source localization", *Journal of the Acoustical Society of America*, **56**, 1974, 165-170.

29. Simpson, W.E. and Stanton, L.D., "Head movement does not facilitate perception of the distance of a source of sound", *American Journal of Psychology*, **86**, 1973, 151-159.

30. Nakayama, K. and Loomis, J.M., "Optical velocity patterns, velocity-sensitive neurons, and space perception: A hypothesis", *Perception*, **3**, 1974, 63-80.

31. Lee, D.N., "The optic flow field: the foundation of vision", *Philosophical Transactions of the Royal Society of London: B*, **290**, 1980, 169-179.

32. Lee, D.N., "Getting around with light or sound", in *Perception and control of self-motion*, R. Warren and A.H. Wertheim (Eds.), Hillsdale, NJ: Lawrence Erlbaum Associates, Inc., 1990, 487-505.

33. Eriksson, E.S., "Distance perception and the ambiguity of visual stimulation", *Perception & Psychophysics*, **13**, 1973, 379-381.

34. Gogel, W.C., "A theory of phenomenal geometry and its applications", *Perception and Psychophysics*, **48**, 1990, 105-123.

35. Longuet-Higgins, H.C. and Prazdny, K., "The interpretation of a moving retinal image", *Proceedings of the Royal Society of London B*, **208**, 1980, 385-397.

36. Fukusima, S.S., Loomis, J.M., and Da Silva, J.A., "Accurate distance perception accessed by two triangulation methods", Paper presented at the annual meeting of the Psychonomic Society, San Francisco, CA, 1991.

37. Loomis, J.M., Da Silva, J.A., Fugita, N., and Fukusima, S.S., "Visual space perception and visually directed action", *Journal of Experimental Psychology: Human Perception and Performance*, **18**, 1992, 906-921.

38. Rieser, J.J., Ashmead, D.H., Talor, C.R., and Youngquist, G.A., "Visual perception and the guidance of locomotion without vision to previously seen targets", *Perception*, **19**, 1990, 675-689.

39. Thomson, J., "Is continuous visual monitoring necessary in visually guided locomotion?", *Journal of Experimental Psychology: Human Perception and Performance*, **9**, 1983, 427-443.

40. Ashmead, D.H. and LeRoy, D., "Auditory distance perception: The contribution of listener motion", submitted, 1993.

41. Shaw, B.K., McGowan, R.S., and Turvey, M.T., "An acoustic variable specifying time-to-contact", *Ecological Psychology*, **3**, 1991, 253-261.

42. Guski, R., "Acoustic tau: An easy analogue to visual tau?", *Ecological Psychology*, **4**, 1992, 189-197.

43. Schiff, W. and Oldak, R., "Accuracy of judging time to arrival: Effects of modality, trajectory, and gender", *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 1990, 303-316.

44. Mershon, D.H. and Philbeck, J.W., "Auditory perceived distance of familiar speech sounds", In preparation, 1994.

45. Gogel, W.C., "Analysis of the perception of motion concomitant with a lateral motion of the head", *Perception & Psychophysics*, **32**, 1982, 241-250.

46. Mershon, D.H. and Hutson, W.E., "Toward the indirect measurement of perceived auditory distance", *Bulletin of the Psychonomic Society*, **29**, 1991, 109-112.

47. Cox, P.H., Jr., *Applying phenomenal geometry to the measurement of perceived auditory distance*, Master's Thesis, Department of Psychology, North Carolina State University, 1992.

48. Michaels, R.M. and Cozan, L.W., "Perceptual and field factors causing lateral displacement", *Public Roads*, **32**, 1965, 233-240.