769

ECOLOGICAL PSYCHOLOGY, 3(3), 253-261

An Acoustic Variable Specifying Time-to-Contact

Brian K. Shaw
Center for the Ecological Study of Perception and Action
University of Connecticut

Richard S. McGowan
Haskins Laboratories

M. T. Turvey
Center for the Ecological Study of Perception and Action
University of Connecticut and
Haskins Laboratories

The acoustic intensity structure generated by the rectilinear convergence of a compact sound source and a point of observation is analyzed. It is shown that an acoustic variable specifying the time-to-contact with the source is available at the point of observation, given certain assumptions about the source, the distance between the source and observation point, and the environment. This variable is $\tau = 2I/(dI/dt)$, where I is the time-averaged intensity. Possible extensions of the analysis are discussed.

The ecological approach to psychology emphasizes the critical role of perception in guiding action. Generally, however, ecologically oriented researchers have focused on the visual guidance of action, to the exclusion of other perceptual systems. The possibility that activity can be guided acoustically has received little attention, other than in species that make obvious use of echolocation. In this article we demonstrate that there is available acoustic information that

---

Requests for reprints should be sent to M. T. Turvey, Center for the Ecological Study of Perception and Action, Department of Psychology, U-20, University of Connecticut, 406 Babbidge Road, Storrs, CT 06269-1020.

could, in principle, be used to control action: information specific to the time-to-contact of the perceiver and a source of sound.

Time-to-contact information has played a central role in the study of the visual control of action. Lee (1976, 1980) showed that the inverse of the relative rate of dilation of an optical angle—a variable he labeled τ—specifies the time-to-contact of an observer with an object or surface, given a rectilinear approach and constant velocity. Evidence that the optical τ variable is in fact used by perceivers is reviewed in Lee (1980, 1990) and Tresilian (1990). Recently, the ability of humans, sighted and unsighted, to perceive time-to-contact auditorily has come under experimental investigation (Schiff & Oldak, 1990). Results suggest a rudimentary ability in sighted humans and a fairly sophisticated ability in unsighted humans. In this article we address the nature of the acoustic information that might underlie this ability; in short, we derive an acoustic τ.

BACKGROUND ON ECOLOGICAL ACOUSTICS

Most research on acoustic perception done to date that might be relevant to the guidance of action has addressed perceivers' ability to localize static sound sources in the environment. Interaural phase, time, and intensity differences in the sound reaching the two ears, as well as directionally dependent structuring of sound by the pinna, have been shown to be important variables in sound localization (see Handel, 1989, for a review of this research). Other relevant research has dealt with the phenomenon of echolocation, in animals like bats and dolphins and in humans (e.g., Lee, 1990, for a review of the work on bats; Handel, 1989, for the findings on humans). Rosenblum, Carello, and Pastore (1987) studied the localization of moving rather than static sound sources, an important step. They showed that interaural temporal differences, the Doppler effect, and intensity changes could all support accurate perception. Although all this work is important, it addresses limited questions. A more general attack on the acoustic guidance of action has yet to be made.

The challenge faced by researchers interested in taking an ecological approach to acoustic perception is to build an ecological acoustics, a description of sound at a level that speaks to the perceptual guidance of activity by organisms. Some issues that we see involved in this enterprise are the descriptive starting point and basic variables that will be needed to capture the higher order variables that might constitute information for perception. It seems clear that the minimal variable-set generally required will be (a) frequency or wavelength, (b) amplitude or intensity, and (c) time (Warren & Verbrugge, 1984). The question of the starting point for the description is harder. We can start with the frequency and intensity patterns at a single point of observation. But this leaves out the interaural factors that research has shown to be important. We can take two points of observation as our basis. But then the question arises as to whether sound-wave occlusions produced by the head and sound structurings produced by the pinna—which are important aspects of the information supporting localization—should be included in the description. Lee (1990) offered another possibility. He described a spatially extended acoustic array in a way analogous to Gibson's (1979/1986) description of the optic array—as a bundle of "acoustic cones" with their apex at a point of observation, each cone having different frequency and intensity characteristics depending on the area of the environment from which it emanated or was reflected. It is an open question whether this approach is the appropriate one. There are reasons to be cautious about the analogy between light and sound in this case. In acoustics, the sources tend to be multiple, intermittent, and important in and of themselves; the major role of light sources in an optic array analysis, on the other hand, is to allow for multiple reflection from the surface layout and, therefore, ambient light. Sound reflections are slower and damp out more quickly than light reflections. Perhaps most important, two ears do not allow for the registration of the fine-grained spatial structure of an array in the way that an eyes does.

Our decisions on these issues made for the purpose of our analysis are meant to be provisional only. We consider the intensity over time at a point of observation. Frequency is ignored, as are interaural differences. This proves adequate for the derivation of an acoustic τ. We have no doubt, however, that a general treatment of ecological acoustics will require more.

An ecological acoustics must of course be based in the physical acoustics that we have. For readers unfamiliar with acoustics, Dowling and Ffowcs-Williams (1983) provide a nice introduction. For the more sophisticated reader, Morse and Ingard (1986) is a standard in the field. In our analysis, we make explicit the assumptions and considerations involved in performing a mathematical-acoustical analysis of the situation of concern. Concepts from acoustics are brought in and defined where needed.

DEFINING THE SCOPE OF THE ANALYSIS

This analysis is based on the acoustic intensity field produced by a compact sound source. A source is *compact* when the size of the actual source of sound (the size of the vibrating object or group of objects producing sound pressure waves) is small compared to the wavelength of the sound. Recall that wavelength is given by dividing the speed of sound (approximately 350 m/s in air) by the frequency. A source length-scale that is 10% of the wavelength is permissible in many applications. Using this percentage, the maximum size of a compact source ranges from 175 to 0.175 cm for the audible frequency range 20 to 20,000 Hz. It is entirely possible for a source that is emitting sound with a broad frequency spectrum to be compact with respect to some of the wavelengths it is emitting but noncompact with respect to others. As an example, we can treat the human mouth itself as a sound source, paying no attention to the vocal tract. If the mouth's diameter is 5 cm, it is a compact source with respect to frequencies below 700 Hz and noncompact with respect to frequencies above 700 Hz. Note that, in fact, a portion of human vocalization does lie below 700 Hz (the great majority of vocalizations fall in the range of 300 to 3,500 Hz). This example suggests that many sound sources in natural environments may be compact at least with respect to some of their frequencies.

The property of compactness helps to simplify problems in mathematical acoustics. Sound sources are extended in space, meaning that the sound waves from spatially separated points on the source may be time delayed to some degree and therefore may have phase differences. If a source is compact the time delays and phase differences are negligible because the spaces between points on the source are small compared to the sound wavelength. This allows you to ignore the spatial extension of the source and to treat it as a point, a move known as the *point-source approximation*. This is not possible with noncompact sources, because the time delays and phase differences are not negligible; they can lead to complicated dependencies of intensity on orientation. Noncompact sources are not considered in this analysis.

Compact sound sources can be further classified by their number of poles. A *monopole* source is the simplest; it is realized in nature when material of a density different from that of the ambient air displaces the air in an unsteady way. Monopoles are described as spherically radiating point sources, because the contours of the intensity field form spheres around the source. The value of intensity is constant on a given sphere. Thus the intensity depends only on the distance from the source, not on the orientation or direction from it. Examples of monopole sources are horns, sirens, and the source of sound in a Bunsen-burner flame (when considering the lower frequency part of the spectrum). *Multipole* sound sources are collections of poles, each acting as a single source, vibrating at a particular relative phase. The poles must be close enough together for the multipole to qualify as compact. *Dipoles* are multipoles made up of two antiphase monopole sources. Examples of dipole sources are flying insects and wires whistling in the wind. Multipoles are nonspherically radiating sound sources, because the sound waves from the different poles cancel and augment in certain ways in different directions from the source, depending on their relative phase. Thus intensity depends on both distance and orientation (Dowling & Ffowcs-Williams, 1983; Morse & Ingard, 1986).

Further conditions on the analysis deal with the distance between the observer and the source. It is assumed that the observer is in the *geometric far field*, which is to say that the observer is at a distance from the source that is much larger than the length scale of the source itself. For the case of multipole sources, it is further assumed that the observer is in the *acoustic far field*, which is to say that he or she is several wavelengths of sound away from the source. For compact sources, for one to be in the acoustic far field means that one is also in the geometric far field.

Acoustic diffraction or scattering due to boundaries is not taken into account in this analysis; the expression for intensity is for the *free field*, which includes no redirected sound components. In cluttered environments, scattering could considerably complicate the simple inverse-square relationship between intensity and distance used here. However, the expression may still be a good approximation in situations in which some boundaries exist, because the secondary waves scattered from boundaries may be weak, or because the symmetry may be such that the expression is still valid. Also, as is discussed in an upcoming section, time-to-contact information may be more readily available and most useful for small times-to-contact, in the range of several seconds. For objects this close boundaries may be less of a problem.

The expression for intensity in our analysis is a time-averaged quantity, where the time averaging can be regarded as being over one cycle for periodic signals. This allows us to consider only those intensity changes with time that are due to the changing distance between the observer and source.

DERIVING τ[1]

We consider the acoustic intensity field structure produced by an observer moving rectilinearly with respect to a sound source or, equivalently, a sound source moving rectilinearly with respect to an observer. The movement is at constant speed. This situation is the same as that addressed by Lee's (1976, 1980) derivation of optical τ.

Assuming constant velocity, actual time-to-contact is given by the ratio of distance or position r to velocity, a ratio that has the units of time:

$$-\frac{\text{Position}}{\text{Velocity}} = -\frac{r}{\dot{r}} = \text{Time-to-Contact.} \tag{1}$$

Because distance is taken as positive, and its time derivative ($dr/dt = \dot{r}$) is negative (the distance is decreasing with time), we must put a negative sign on the ratio to get a positive time.

The key to identifying field properties specific to time-to-contact is to find ratios of field quantities that are equal to the time-to-contact ratio, dimensionally and numerically. For the optical-flow field, an appropriate ratio is that of an optical angle to its time derivative (Lee, 1976). For the acoustic case, we look to the intensity field, a scalar field, and to the inverse-square law.

The layout of time-averaged intensities around a monopole sound source is

---

[1] This derivation was first performed, to our knowledge, by Pittenger (1971) in an unpublished note. We performed the derivation independently of, and without the knowledge of, Pittenger's work.

given by the inverse-square law, which states that intensity is inversely proportional to the square of the distance from the source:

$$I = kr^{-2}. \tag{2}$$

The constant of proportionality k reflects, in part, the strength of the source. The relation for multipole sound sources is the same, by approximation, except that k is not a constant but a function of the direction from the source (recall that intensity varies with both distance and direction for multipoles). However, our analysis is confined to the rectilinear approach of a source to an observer or vice versa. In this situation the direction from the source to the observer is constant during the approach; therefore k is constant, and the expression for the inverse-square law (Equation 2) holds for both monopole and multipole sources.

The derivative of the intensity field with respect to r, its rate of change over changes in distance, is

$$\frac{dI}{dr} = -2kr^{-3}. \tag{3}$$

When we consider a moving observer or source, the distance changes over time. Therefore, r is a function of t (time), and

$$I = kr(t)^{-2}. \tag{4}$$

We take the time derivative of I,

$$\frac{dI}{dt} = \dot{I} = -2kr^{-3}\dot{r}. \tag{5}$$

The ratio of I to dI/dt is a time. It is, therefore, a candidate field property for τ. The question is, is this ratio equal to the time-to-contact ratio? The answer is that the ratio equals the time-to-contact ratio multiplied by ½:

$$\frac{I}{\dot{I}} = \frac{kr^{-2}}{-2kr^{-3}\dot{r}} = -\frac{r}{2\dot{r}}. \tag{6}$$

Therefore, assuming constant velocity

$$\tau = 2\frac{I}{\dot{I}} = \text{Time-to-Contact}. \tag{7}$$

DEPENDENCIES OF INTENSITY AND τ ON TIME

Figure 1 depicts intensity as a function of time during rectilinear approach to a sound source for various k values. The initial distance for all cases is 20 m. The three curves are for velocity $= -2$ m/s, and for $k = 10^{-7}$ W, 10^{-6} W, and 10^{-5}

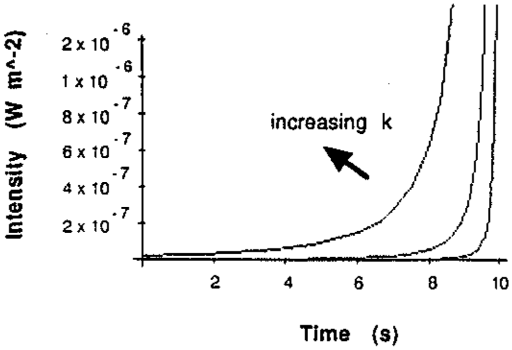

FIGURE 1 Intensity as a function of time for rectilinear approach with an initial distance of 20 m, velocity of -2 m/s, and three k values: 10^{-7} W, 10^{-6} W, and 10^{-5} W.

W. Contact is at time = 10 s; the time-to-contact, then, is 10 s minus the elapsed time. The graph demonstrates that changes in source strength, reflected in k, cause changes in the time evolution of intensity. Intensity is not a single-valued function of distance or time-to-contact. However, the ratio of intensity to its time derivative—equivalent to the ratio of a point on an intensity curve to its tangent—is a single-valued function of time-to-contact: It equals ½ the time-to-contact, as proven earlier. This is illustrated in Figure 2, which shows the relation of τ to time in the same approach situations as Figure 1. This graph shows the information in the time evolution of intensity that is invariant over changes in source strength and specific to time-to-contact.

Another point follows from Figure 1. Intensity is an inverse-square function of time (given constant velocity). It increases very slowly for larger times-to-contact, then begins to increase very quickly for smaller times-to-contact (with an infinite asymptote at time-to-contact = 0). The optical expansion produced by a rectilinear approach to an object or surface follows a similar trend over time. Given that changes of intensity over time are more pronounced in the smaller time-to-contact region, it is possible that perceivers would be better able to register the rate of change of intensity—and thus be better able to register acoustic τ—in that region. The same sort of argument can be made for the registration of optical τ. This could account for the observation that visually based judgments of time-to-contact are not highly accurate for times greater than 2 to 3 s (Schiff & Oldak, 1990). And in fact, most actions that require highly accurate contact information take place within the time scale of 1 s or less.

EXTENSIONS AND LIMITATIONS

Our analysis has considered only rectilinear convergence of the observer and source. It would be valuable to extend it to include a range of possible paths. This would introduce the question of whether or not there is information

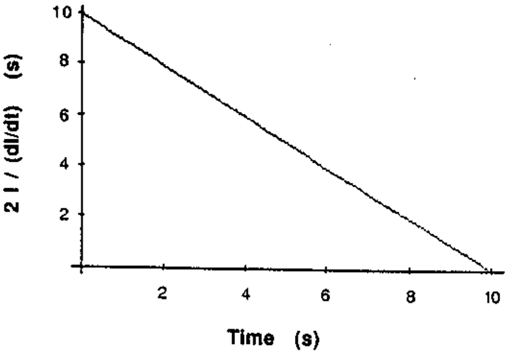

FIGURE 2 τ as a function of time for rectilinear approach with an initial distance of 20 m, velocity of -2 m/sec, and any k value.

specific to one's path—the direction of heading and so on. An observation along this line is that, for multipole sound sources, the intensity depends on both distance and direction and, therefore, movement that is not along a constant direction (rectilinear) could create significant intensity and τ variations. The implication is that τ variations could be used to direct behavior; for movement at a constant velocity, to minimize τ variation is to locomote on a path that is directed at the source.

There is a rich variety of sound sources in the environment. Here we have considered only single compact sources. Many sources, however, are not compact (for all or some of the frequencies they emit). The acoustic intensity fields generated by noncompact sources warrant investigation along similar lines. The intensity fields generated by more than one source warrant investigation as well.

The only acoustic variable that we have considered here is intensity. As stated in the introduction, a more general treatment will require frequency as well. What we have attempted is just a small start on an ecological acoustics. That there is information specifying time-to-contact in the simple intensity field bodes well for such an enterprise.

ACKNOWLEDGMENTS

This material is based on work supported under a National Science Foundation Graduate Fellowship awarded to Brian K. Shaw. Support was also provided by National Institutes of Health Grants HD–01994 and DC-00121 awarded to Haskins Laboratories.

We thank Claire Michaels, John Pittenger, Stephen Handel, Jim Todd, and Reinoud Bootsma for helpful reviews.

REFERENCES

Dowling, A. P., & Ffowcs-Williams, J. E. (1983). *Sound and sources of sound.* New York: Wiley.

Gibson, J. J. (1986). *The ecological approach to visual perception.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc. (Original work published 1979)

Handel, S. (1989). *Listening: An introduction to the perception of auditory events.* Cambridge, MA: MIT Press.

Lee, D. N. (1976). A theory of visual control of braking based on information about time to collision. *Perception, 5,* 437–459.

Lee, D. N. (1980). Visuo-motor coordination in space-time. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior* (pp. 281–295). New York: North-Holland.

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

Morse, P. M., & Ingard, K. U. (1986). *Theoretical acoustics.* Princeton, NJ: Princeton University Press.

Pittenger, J. B. (1971). *Time to collision information in sound intensity.* Unpublished manuscript.

Rosenblum, L. D., Carello, C., & Pastore, R. E. (1987). Relative effectiveness of three stimulus variables for locating a moving sound source. *Perception, 16,* 175–186.

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

Tresilian, J. R. (1990). Perceptual information for the timing of interceptive action. *Perception, 19,* 223–239.

Warren, W. H., & Verbrugge, R. R. (1984). Auditory perception of breaking and bouncing events: A case study in ecological acoustics. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 704–712.