

1

Introduction to Sound Source Localization

RICHARD R. FAY AND ARTHUR N. POPPER

The sense of hearing has evolved and been maintained so that organisms can make use of sound in their environment not only for communication, but also to glean information about the general acoustic milieu that enhances general survival (e.g., Fay and Popper 2000). Information in the environment enables animals to learn about sources that are in many different directions, and particularly signals that are outside of the detection range of other senses. Still, while sound is inherently important for overall survival, its value would be very limited if the receiving organism did not know the position of the source. In effect, to make maximum use of a sound from a predator, the receiver must know not only that the predator is present, but also where the predator is in order to escape most effectively.

As a consequence, one of the fundamental and most important features of sound source detection is the ability of an animal to estimate source location as a first step in behaving appropriately in response to the sound. The need for sound source localization thus has become a fundamental feature of hearing in most hearing organisms, and one could argue that it is inconceivable that a sense of hearing could have evolved at all without the ability to locate and segregate sources.

At the same time, the strategies for computation of sound location must be different in different species, and must depend on the nature of the information arriving from the ears to the brain, and this in turn must depend on interactions between the physics of sound and the characteristics of the receivers.

This volume treats sound source localization from this comparative and evolutionary perspective. The auditory receivers of vertebrates and invertebrates vary widely (e.g., Hoy et al. 1998; Manley et al. 2004), from the tympanic, pressure receiver ears of insects, birds, and mammals, to the otolithic ears of fishes that are inherently directional, to the pressure gradient receiver ears of amphibians (and some reptiles and birds) that function like inherently directional pressure receivers. In spite of these differences in the ways that different ears respond to sound, there apparently are only a few acoustic cues for source location: interaural time differences (ITD), interaural intensity differences (IID), the shape of the sound spectrum reaching the ears, and the axis of acoustic

particle motion. The latter cue can be used in conjunction with inherent receiver directionality to create neural representations that are functionally equivalent to those provided by the interaural acoustic cues (IID and ITD) (Fay, Chapter 3). Several chapters of this volume make clear the diversity of mechanisms that have developed to exploit these acoustic cues in sound source localization.

Our understanding of hearing mechanisms in general has been guided by our extensive knowledge of human hearing (at a behavioral or functional level), and as such a considerable amount of work on localization in general has focused on the cues and mechanisms that operate in the human auditory system. Thus, we understand the importance of the interaural cues in determining azimuth and, more recently, the role of spectral shape in determining elevation. Colburn and Kulkarni (Chapter 8) authoritatively summarize the sorts of models that have developed in an attempt to account for sound source localization as it occurs in human hearing. The fundamental question that is asked by Colburn and Kulkarni is, "How does the localizer interpret the received signals to determine the location of the sound source?" The authors evaluate several theoretical schemes that locate the sound source while at the same time extracting information about the acoustic environment, making use of a priori and multimodal information about the original acoustic signal and environment, and estimating other properties of the signal. Models are roughly categorized as those that find the maxima in a process called "steering the beam" (e.g., cross-correlation models) and those that find the minima in the array output by "steering the null" (e.g., equalization and cancellation model). Modeling is generally most successful when there is only one source of sound, no reverberation, no conflicting cues, and no unusual spectra (for elevation judgments), and when the model is restricted to judgments of lateral position.

Trahiotis, Bernstein, Stern, and Buell (Chapter 7) evaluate binaural hearing, broadly defined, from the point of view of interaural correlation as originally suggested by Lloyd A. Jeffress (1948). After defining the indices of interaural correlation and their application to binaural perception, the concept of the three-dimensional cross-correlation function is introduced and it is shown how it is possible to understand binaural perception in terms of various types of pattern processing operations on this function. The authors then present neurophysiological evidence for cross-correlation mechanisms in binaural hearing, a 50-year-old model (Jeffress 1948) that has maintained its value both as a useful guide to the quantitative understanding of binaural hearing in humans and as a neurobiological explanation for the computation of sensory maps in the auditory system (see Kubke and Carr, Chapter 6). Trahiotis et al. use insights that arise from functional/behavioral investigations on binaural hearing mechanisms to support the principles of "sloppy workmanship" and the "principle of diversity" (Huggins and Licklider 1951) applied to the mechanistic implementations of binaural processing. The principle of sloppy workmanship refers to the danger of postulating a neural structure that is precisely arranged in detail; it is important to recognize that the postulated mechanism need function only in a statistical sense. The principle of diversity states that there are many ways to skin a cat

and that the nervous system may use all of them. The implication for theory, here, is that conceptions that appear to be alternatives may supplement one another.

These principles are illustrated by Kubke and Carr (Chapter 6), who examine the development of auditory centers responsible for sound localization in birds and mammals. In both taxa, the developmental processes that shape the basic plan of the auditory circuit are complemented by plastic modifications that fine tune the neuronal connections to adapt to the experience of each individual animal. The resulting neuronally computed auditory space map associates particular binaural cues with specific sound source locations. But since binaural information will be determined by head size and shape of each individual, the auditory system must be able to adapt the basic connectivity plan to each animal. Thus, accurate associations between binaural cues and space assignments can develop only after the requirements of each individual are determined. The process therefore requires experience-dependent plasticity. In general, the neural circuits responsible for sound source localization can be recalibrated throughout life.

Comparative psychoacoustics is the link between the neural mechanisms responsible for localization and models of sound source localization applied to human hearing. Brown and May (Chapter 5) review the literature on behavioral studies of sound source localization in mammals with a focus on the cues operating for judgments in azimuth, elevation, and distance (proximity). It is clear from comparative work that the mechanisms for coding these three dimensions of directional hearing are entirely different, and may have had quite distinct evolutionary histories. The interaural cues of intensity and time are the primary ones for determination of azimuth for most mammalian species, including humans. The cues for elevation are less well understood, but seem to be related to the spectrum of the received sound, as filtered by the head-related transfer function (HRTF). It is remarkable, really, that the cues for estimating azimuth (ILD and ITD), and those for estimating elevation (the HRTF) are so fundamentally different. Processing of the HRTF depends on the tonotopic axis of the cochlea and the acuity of frequency analysis (the cue is essentially mapped onto the cochlea), while processing ITDs is independent of the acuity of frequency analysis. One could argue that processing ILDs is at least partially a matter of frequency analysis because the cue could exist in one frequency band and not in another, and its existence depends on the precision of frequency analysis. But judgments of elevation are a direct consequence of processing the spectrum, while the ILD cue must be processed in the level domain to be useful. The cues for distance or proximity are subtle and there is no direct connection between a given cue and the distance it specifies. Thus, the mechanisms for estimating azimuth, elevation, and distance are different, and it is possible that each may have had a different evolutionary history and that acuity along these dimensions may vary independently from species to species.

It is remarkable that vertebrates have so much in common when it comes to sound source localization. Even sound localization by fishes, with their otolithic

ears and nonhomologous brainstem nuclei, can be viewed as a variation on a theme shared by many species (see Fay Chapter 3). Sound source localization among fishes is incompletely understood, but what we do understand seems familiar. First, localization in azimuth appears to be matter of binaural processing. Interaural acoustic cues are small or nonexistent owing to the high speed of sound underwater. In addition, fishes have close-set ears and (in some cases) indirect stimulation of the two ears via an unpaired swim bladder. Yet, the peripheral auditory system of fishes seems to reconstitute the functionally equivalent interaural cues through ears that are inherently directional. For example, while ILDs probably do not exist for fishes, the neurally coded output of the ears represents interaural response differences by virtue of the ears inherent directionality. Since response latency is a function of response magnitude in some primary afferents, there is also the equivalent of ITDs in the response of the ears. Just as in terrestrial vertebrates, the coding for azimuth and elevation seems to be independent of one another. In fishes, the ear is not tonotopically organized, but rather is organized directly with respect to sound source elevation (hair cell orientation). Tetrapod ears, having tonotopic organization, estimate elevation through an analysis of the sound spectrum as filtered by the HRTF. Fish ears analyze elevation directly through an across-fiber peripheral code that reflects hair cell orientation.

The inherent directionality of the ears of fishes puts them into the category of pressure-gradient receivers. It has become clear that pressure gradient or pressure-difference receivers are more widespread among animals than previously thought. Pressure-difference receivers are inherently directional receivers that can be of two basic types. In tetrapods the most familiar type consists of a tympanic ear for which sound may find pathways to both sides, creating a pressure difference across the membrane. Depending on the phase and amplitude of sound reaching both sides of the membrane, interaural differences in time and magnitude can become quite large. These interaural differences do not amplify the interaural differences that normally would accompany close set ears, but rather determine interaural differences essentially arbitrarily. These ears are thought to be inherently directional simply because their response varies as an arbitrary function of azimuth or elevation. The less familiar type of pressure-gradient receiver consists of an array of particle velocity receivers, such as hair cells with particular directional orientations (in fishes), or insect antennae, filiform hairs, or terminal cerci that respond directly and in a directional manner to acoustic particle motion (Robert, Chapter 2). These are properly referred to as pressure-gradient receivers because acoustic particle motion occurs to extent that pressure gradients exist. These receptors require a different sort of central processing and computation than is required for pure pressure receivers. Rather than, for example, computing azimuth based on the differences in the time and intensity of sound reaching the two ears, the axis of acoustic particle motion is probably estimated as the "best" axis from the population of active fibers varying in most sensitive axis.

These pressure difference receivers occur widely among species, and are prob-

ably primitive among vertebrates and invertebrates alike. Thus, as shown by Robert (Chapter 2), mosquitos, caterpillars, and crickets are known to have non-tympanal auditory receptors, and many insects have pressure-gradient tympanal ears. The ears of fishes are all pressure-gradient receivers, responding directly to acoustic particle motion (see Fay, Chapter 3). Those fishes that have pure pressure receivers (only confirmed in one fish taxa, the Ostariophysi) are a distinct minority. Among anuran amphibians, pressure gradient ears are ubiquitous and are thought to represent the primitive condition. Reptilian and avian ears are characterized by interaural canals, and although experiments must be done to confirm pressure difference hearing in each species, it is possible that most species detect pressure gradients, at least at low frequencies (see Christensen-Dalsgaard, Chapter 4). Thus, mammals are the only vertebrate group to have lost sensitivity to pressure gradients and to this mode of directional hearing. The pure pressure receiver characteristic of the mammalian ears seems to have given up inherent directionality and interaction with the respiratory system for sensitivity to high frequencies and all the advantages that come with high-frequency hearing.

References

- Fay RR, Popper AN (2000) Evolution of hearing in vertebrates: The inner ears and processing. *Hear Res* 149:1–10.
- Hoy R, Popper AN, Fay RR (eds) (1998) *Comparative Hearing: Insects*. New York: Springer-Verlag.
- Huggins WH, Licklider JCR (1951) Place mechanisms of auditory frequency analysis. *J Acoust Soc Am* 23:290–299.
- Jeffress, LA (1948) A place mechanism of sound localization. *J Comp Physiol* 41:35–39.
- Manley GA, Popper AN, Fay RR (eds) (2004) *Evolution of the Vertebrate Auditory System*. New York: Springer.