

# 5

## Comparative Mammalian Sound Localization

CHARLES H. BROWN AND BRADFORD J. MAY

### 1. Introduction

In natural environments, the approach of a competitor, a predator, a relative, a mate, or one's prey may be conveyed by subtle fluctuations within the acoustic environment. In many instances it is likely that the early detection of an intruder is conveyed not by a sound that is unusual or uncommon because of its amplitude or frequency composition, but rather by a sound that is distinctive chiefly because it occurred at an "inappropriate" location within the acoustic landscape. Here, one's ability to survive depends not on unusual sound detection capabilities, but rather on a sound localization system that permits a listener to effortlessly, yet vigilantly, track the relative positions of the sources of sounds that signal safety or danger. Moreover, the absence of a "safe" sound may be as significant to many birds and mammals as is the presence of an "unsafe" one; for an intruder's approach may be disclosed by either the production of unexpected sounds, or by the abrupt cessation of "expected" sounds that were previously sustained or ongoing in some regular pattern. Movements made unstealthily will disrupt the chorus of cicadas or the sounds of birds, or other animals, and a ripple of silence may spread across the landscape signaling that something (or someone) is nearby. The subtlest acoustic changes may be biologically the most telling. Clumsy predators are apt to go hungry, and an evolutionary premium has been placed upon the ability of most animals to quickly discern the position of a sound that does not belong (or the position of an unexpected cessation of those sounds that do belong). In the struggle for survival, the determination of the origin of a signal may be assigned a level of importance that equals or exceeds that of being able to recognize the sound, or being able to identify the perpetrator of the disturbance. It is in this biological context that the mechanisms underlying sound localization evolved, and through the course of the succession of animals on earth the sound localization abilities of many species have come to exhibit remarkable acuity and species specificity.

The position of the source is a cardinal perceptual attribute of sound. Under normal conditions, for human listeners, the source of a stimulus is instantly and

effortlessly assigned a position with reference to the orientation of the listener. The localization of sound is seemingly reflexive, the perception of direction is “instantaneous,” and localization does not appear to be derived by some kind of deductive cognitive process. That is, under most conditions listeners do not actively think about having to triangulate the possible origin of the sound given what they heard at their two ears. Just as a sound is perceived as having some magnitude, pitch (or noisiness), loudness, and duration, it also is perceived as having a distance dimension (it is near or far), an elevation dimension (above or below), and an azimuth dimension (left or right of the observer). Only when listeners wear earphones do sounds routinely lack a coherent or natural spatial image, and under these conditions the normal filtering characteristics of the external ear and ear canal have been bypassed, and the normal correlation between the timing and amplitude of the signals at the two ears has been violated.

Batteau et al. (1965) noted that many sounds presented through earphones are reported to have an origin somewhere inside the listener’s head. They showed that sounds presented through earphones would be perceived as having a normal external position and could be accurately located in space if the signals fed to the left ear and right ears originate from microphones positioned approximately 17.5 cm apart (a normal head width) and if the microphones were fitted with replicas of human pinnae. The apparent origin of the signal is “external” to the listener when sounds are presented this way, and if the position of a sound source delivered to the microphone array is moved to the left or to the right, the perceived location of the sound source moves accordingly. If the artificial pinnae are removed from the microphones, or if the normal free-field-to-eardrum transfer functions are artificially manipulated, localization accuracy suffers (Batteau et al. 1965; Wightman and Kistler 1989a,b, 1992; Middlebrooks 1992, 1999).

It is known that the position of sound is a core dimension of auditory perception in adult human subjects, and there is good reason to believe that the same is true for human infants, and for most vertebrates. That is to say, the position of a startling sound appears to “command” most vertebrates to orient toward its site of origin. Although auditory experience may modify and adjust localization during development (Knudsen 1983, Knudsen et al. 1984), reflexive orientation to sound position is evident at or near birth in a wide variety of subjects including laughing gulls (Beer 1969, 1970), Peking ducklings (Gottlieb 1965), infant cats (Clements and Kelly 1978a), rats (Potash and Kelly 1980), guinea pigs (Clements and Kelly 1978b), and humans (Muir and Field 1979; Wetheimer 1961). The data suggest that most vertebrates, including both altricial and precocial species, are able to reflexively locate the origin of sound nearly as soon as the ear canal opens and they are able to hear.

In many organisms sound localization mechanisms may initiate and actively guide saccadic eye movements to the site of potentially important events. Animals with binocular frontal visual systems, such as most primates, have limited peripheral or hemispheric vision, and these species may be particularly dependent on a high-acuity directional hearing system to rapidly direct the eyes to the location of a disturbance (Harrison and Irving 1966). Furthermore, the more

restricted the width of the horizontal binocular visual field in various mammals, the greater the acuity of their sound localization abilities (Heffner and Heffner 1985, 1992). This enhanced acuity may be critical for accurately aiming the eyes.

The perception of many events is bimodal. Speech perception, for example, is influenced by both visual information regarding tongue and lip configuration, and by the corresponding acoustic signal. When these two modalities of information are out of synchrony, or artificially separated in space, the result is disturbing to both adult and infant human subjects (Aronson and Rosenbloom 1971; Mendelson and Haith 1976). The preservation of the normal congruence between visual and auditory space is important for the development of sound location discriminations in animals (Beecher and Harrison 1971). Animals appear to be prepared to learn to direct visually guided responses toward objects positioned at the origin of sounds, and correspondingly unprepared to learn to direct responses toward objects that have been repositioned so that the contiguity of visual and auditory space has been violated (Beecher and Harrison 1971, Harrison et al. 1971; Harrison 1990, 1992). For an organism to be able to react appropriately to events occurring at different locations in space it is necessary that the visual and acoustic perceptual maps be aligned and in register. Visual deprivation early in development alters sound localization in rats (Spigelman and Bryden 1967), cats (Rauschecker and Harris 1983; Rauschecker and Knierpert 1994), ferrets *Mustela putorius* (King et al. 1988; King and Parsons 1999), and barn owls (*Tyto alba*) (Knudsen and Knudsen 1985, 1990; Knudsen and Brainard 1991; Knudsen et al. 1991). Theoretically it is possible that spatial maps are organized by visual experience. Thus, visual deprivation would lead to impairments in sound localization. Alternatively, it is possible that visual deprivation produces compensatory sharpening of directional hearing. Some recent findings are consistent with both of these contrasting perspectives, and these data are described in Section 4.5.

Although sound localization mechanisms evolved because of their significance to survival in the natural world (Masterton et al. 1969), sound localization abilities have nearly always been studied in synthetic, quiet, echo-free environments (or even with earphones), and testing has often been conducted with tones, clicks or band-limited bursts of noise. The intent of this tradition has been to assess the absolute limits of precision of directional hearing, though at the expense of exploring how well sound localization abilities function under more normal conditions. The sections that follow describe the physical cues for sound localization available to terrestrial vertebrates and the behavioral methodologies commonly used to assess the sound localization capabilities of animals, and then go on to survey the sound localization abilities of selected mammals.

## 2. Localization Cues

The origin of sound in space is referenced relative to the orientation of the listener, and sound position is accordingly expressed in terms of its azimuth, elevation, and distance from the listener.

## *2.1 Sound Source Azimuth: The Horizontal Coordinate of Sound Location*

Sound localization is dependent upon the comparison of the sound waves incident at each ear in most terrestrial vertebrates. These interaural (or binaural) differences are the result of two factors: (1) the difference in distance ( $\Delta d$ ) the sound wave must travel to reach the tympanic membrane of the two ears and (2) differences in the transfer function of the signal incident at each ear. In the case of pure tones, or very narrow bandwidth signals, differences in the transfer function at each ear are reduced to interaural level (or amplitude) differences of the waveform incident at each ear. We defer discussion of spectral cues and head-related transfer functions (HRTFs) to Section 5, and we begin the examination of directional hearing with simple sine waves. The first factor, the difference in propagation distance for the near and far ear, results in differences in the time-of-arrival and in differences in the phase of the signal at each ear. When a sound is presented from a position off to one side of a listener (not on the midline, or at  $0^\circ$  azimuth), corresponding points in the sound wave will necessarily be received by the “near” ear (the ear on the side of the head which is toward the source of the sound) before it reaches the “far” ear. The velocity of sound in air is nominally 343 m/s; given this velocity, for each additional cm the sound wave must travel to reach the far ear, the wave will arrive  $29 \mu\text{s}$  later than it will at the near ear. Hence, interaural differences in the time-of-arrival of corresponding points in the sound wave may serve as one of the principal cues for directional hearing.

For the case in which the stimulus is a simple sustained cyclic wave, such as a pure tone, differences in the arrival time of the near- and far-ear waveforms will result in interaural differences in the phase of the wave as long as the arrival time difference is not equal to the period (or integral multiples of the period) of the signal. For example, the additional time required for the sound wave to reach the far ear may be a fraction of the period of the wave, such as one fourth of the period. In this case the corresponding interaural difference in signal phase would be one fourth of  $360^\circ$ , or  $90^\circ$ . Increments or decrements in the arrival time of near- and far-ear waveforms would result in corresponding increases or decreases in the difference in interaural phase. If the position of the source of the sound were moved so that the arrival time difference is increased from a fraction of the period to exactly match the period of the signal, the near- and far-ear waves would again be in phase. The sound wave incident at the far ear would be precisely one cycle behind the corresponding wave at the near ear. For the special cases in which the arrival time difference between the near- and far-ear waves happens to coincide with two times, three times, or other integral multiples of the period of the signal, the near- and far-ear waves will again be in register, and there will be no interaural differences in phase. Although these phase ambiguities apply to tonal stimuli, they are probably unimportant for most complex stimuli for which other cues for sound localization are available. Furthermore, even with tones, it is unlikely that the source of a sound will be located such that the arrival time differences will equal the period (or multiples of the

period) of the signal; consequently, interaural differences in arrival time are usually reflected by interaural phase differences, and these differences may serve as a cue for localizing the azimuth of the source.

Interaural level differences (ILDs) are an additional cue for the perception of the azimuth of sound position. ILDs may occur when the origin of the sound is off to one side, and it is a consequence of the shape of the torso, head, pinna, and external ear canal, as well as the properties of sound diffraction, reflection, and refraction with these structures. The magnitude of sound diffraction, reflection and refraction is dependent on the relative dimensions of the wavelength of the sound wave and the size and shape of the reflective structures. In general, ILDs are most important for signals composed of wavelengths that are less than the diameter of the listener's head. Shorter wavelengths (e.g., higher-frequency signals) produce more prominent ILDs, and the characteristics of these differences are highly dependent on the specific geometry of the listener's head and pinna.

## 2.2 Geometrical Considerations

As a first approximation, an acoustically opaque sphere may model the head with the ears diametrically opposed (Rayleigh 1876, 1945). With this idealization, the shape and dimensions of the pinna are ignored. Furthermore, the idealized model assumes that the head is immobile (unable to scan the sound field), and that a point sound source is positioned greater than 1 m from the listener. Under these conditions, a plane may approximate the wave front. Interaural distance differences ( $\Delta d$ ) will occur for all sound locations other than those that lie on the median plane. As depicted in Figure 5.1, for a sound source to the right of a listener at azimuth  $X$ , the additional distance that the sound must travel to reach the far ear (left ear) is given by the sum of the linear distance  $r(\sin X)$  and the curvilinear distance  $r(X)$ . That is, the difference ( $\Delta d$ ) in the sound path-length for the two ears is given by Eq. (5.1),

$$\Delta d = r(X + \sin X) \quad (5.1)$$

where  $\Delta d$  is the distance difference in cm,  $r$  is the radius of the listener's head in cm, and the sound source azimuth angle  $X$  is measured in radians.

The path length difference to the two ears is acoustically realized by the interaural difference in time-of-arrival of corresponding points in the waveforms incident at the two ears. Time-of-arrival differences ( $\Delta t$ ) are calculated by dividing the distance difference by the velocity of sound. Given a sound velocity in air of 343 m/s, then the relationship between  $\Delta t$  and azimuth is provided by Eq. (5.2),

$$\Delta t = r(X + \sin X) / 3.43 \times 10^4 \quad (5.2)$$

where  $\Delta t$  is the temporal difference in  $\mu\text{s}$ ,  $r$  is the radius of the observer's head in cm, and the azimuth angle  $X$  is measured in radians.

Three factors merit emphasis. First,  $\Delta t$  approaches a maximum value as az-

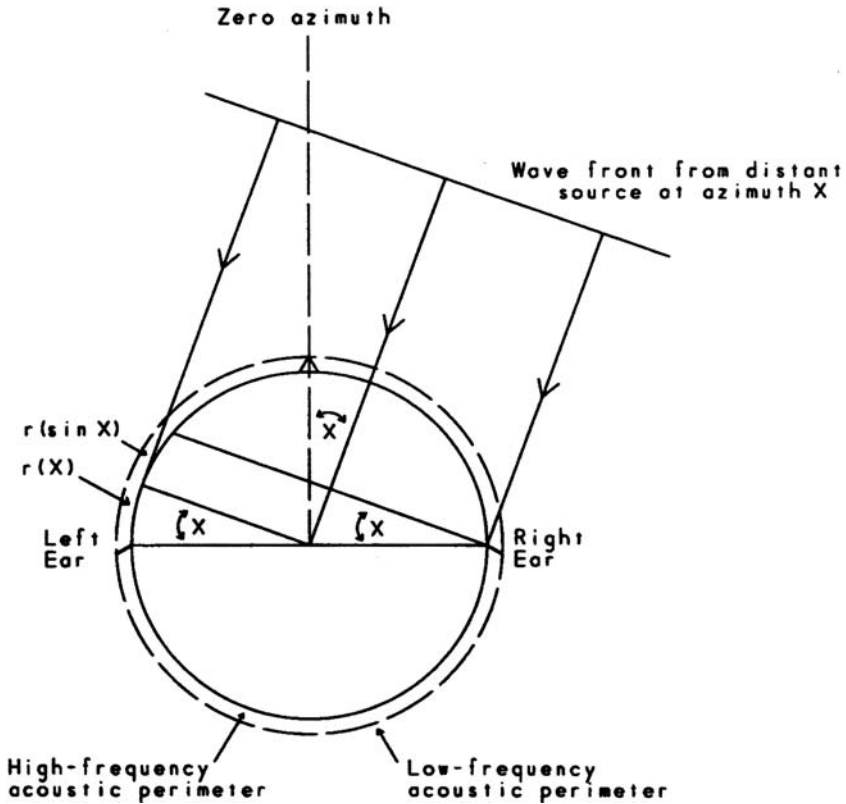


FIGURE 5.1. Geometrical considerations for interaural differences in ongoing time or phase differences as a function of frequency. The signal is presented at azimuth  $X$  from a distant location from the listener. The interaural differences in signal phase obtained with low-frequency tones are produced by a larger effective acoustic radius of the head compared to that obtained by a high-frequency signals. High- and low-frequency signals are scaled relative to head size. A high-frequency signal is one in which the wavelength of the sound is equal to or less than two times the radius of the listener's head, while a low-frequency tone is one in which the wavelength is eight times the radius of the head or greater. In the intermediate frequency zone (defined by the interval  $2r < X < 8r$ ) the effective radius of the head gradually changes from the two boundary conditions illustrated here. See Kuhn (1977) for a detailed treatment of these phenomena.

imuth  $X$  approaches  $\pi/2$  radians or  $3\pi/2$  radians ( $90^\circ$  or  $270^\circ$ ). Assuming  $r = 8.75$  cm, the usual value assigned for humans, then maximum  $\Delta t = 656 \mu\text{s}$ . That is, at  $\pi/2$  radians ( $90^\circ$ ) the sound wave arrives at the far ear  $656 \mu\text{s}$  after it arrives at the near ear. Second, for any given azimuth  $X$ ,  $\Delta t$  varies directly with  $r$ . As a result, listeners with large heads will experience a greater interaural time-of-arrival difference than will subjects with small heads. Consequently, if

the neural processing resolution of time-of-arrival differences is approximately equal across mammalian species, then species with large heads will be able to perceive finer changes in azimuth using this cue than will species with small heads. Third, interaural time-of-arrival differences do not define a specific locus in three-dimensional space. That is, sources that differ in elevation may still have the same interaural path-length difference. Furthermore, the hemifield behind a listener is a mirror image of that in front, and locations above, behind, or below a listener may have the same interaural time-of-arrival difference as those in the front hemifield. If an organism's trunk and limbs did not serve as acoustic obstacles influencing interaural time-of-arrival differences, equal differences in time-of-arrival would be given by all points falling on a surface of a cone centered on the aural axis. In most instances, however, the limbs and torso are acoustic obstacles that impact significantly on path-length differences (and hence, interaural time-of-arrival differences) for sound sources positioned above, behind, or below a listener (Kuhn 1979). As a consequence of this complication, while equal time-of-arrival differences may be empirically measured for displacements of the sound source in the horizontal and vertical planes, equal time-of-arrival differences are unlikely to correspond to the surface of a cone.

### *2.3 Sound Frequency and Effective Radius*

Time-of-arrival cues are produced for both the leading edge of the wavefront, and for ongoing time or phase differences in the waveforms. Interaural differences in time-of-arrival cues are influenced by frequency, head size, and azimuth. Kuhn (1977, 1987) has shown that the effective acoustic radius of the head is larger than the skull perimeter when low-frequency stimuli are presented, but equal to the skull perimeter when high-frequency sounds are presented. In general, when the wavelength of the stimulus is less than or equal to the diameter of the skull (a high-frequency signal) the effective acoustic radius approximates that of the perimeter of the skull. When the wavelength of the stimulus is greater than or equal to four times the diameter of the skull (a low-frequency signal) the effective acoustic radius expands to a larger value with a magnitude that is probably governed by the degree of prognathism, the protrusion of the jaw and nose, and by the size of the pinna. In humans the effective acoustic radius of the head for low-frequency signals is about 150% of that for high-frequency signals (Kuhn 1977), and the transition in the effective radius occurs at about 2000 Hz. In animals with pronounced prognathism of the nose and jaw, and hypertrophied pinna even greater expansion of the effective acoustic radius of the head for low-frequency signals would be expected. This phenomenon is not intuitively apparent, and it is attributed to a frequency dependency in the pattern of standing waves created around acoustic barriers (Kuhn 1977).

The functional acoustic radius for the leading edge of a sound wave, however, is equal to the skull perimeter, and it is not influenced by the relative frequency of the signal (Kuhn 1977). Thus, as depicted in Figure 5.1, low-frequency

signals have an enhanced difference in interaural phase because the effective acoustic radius of the head is expanded for the fine structure of these signals.

The cues that are available for time-domain processing are influenced by the rise and fall times of the signal and the complexity of the frequency spectrum and envelope (or amplitude modulation) of the waveform. Signals, which seem to begin and end imperceptibly, have a slow rise and fall time (gradual onset and offset) and lack a crisp leading edge. Consequently, time-domain localization would likely be restricted to the comparison of interaural differences in the phase of the fine structure of the signal, or to the comparison of interaural differences in the amplitude contour, or envelope, of the waveform (Henning 1974; McFadden and Pasanen 1976). In the case of an unmodulated, slow onset and offset pure tone, time-domain processing would necessarily be restricted to an analysis of the interaural phase differences of the fine structure of the tone. However, in spectrally and temporally complex signals, the envelope will be modulated, and the envelope of these modulations will be incident at the near and far ear at correspondingly different times-of-arrival. Human subjects are able to localize signals by processing interaural differences in signal envelopes (Henning 1974; McFadden and Pasanen 1976), and these envelope cues influence sound localization in other mammals as well (Brown et al. 1980). Thus, time-domain processing of localization cues may be analyzed by processing interaural differences of the cycle-by-cycle fine structure of the signal, or by the processing of interaural differences in the time-of-arrival of the more global modulations of the envelope of complex signals (Middlebrooks and Green 1990).

## 2.4 Azimuth Ambiguity

Interaural differences in signal phase may provide ambiguous information regarding the position of the source. By way of example, assume that the radius of the head is 8.75 cm, and that the maximum time difference is 656 ms for the fine structure of the signals in question. In this example, as a simplification, ignore the fact the effective acoustic radius may change for signals of different frequency. This interaural time difference (ITD) ( $\Delta t = 656 \mu\text{s}$ ) would result in interaural phase differences of  $90^\circ$ ,  $180^\circ$ , and  $360^\circ$  for pure tones of 380 Hz, 760 Hz, and 1520 Hz, respectively. This example illustrates two points. First, the relationship between interaural phase difference and spatial location is frequency dependent. A phase difference of  $30^\circ$  indicates one position for a tone of one frequency, but a different position for a tone of another frequency. Second, more than one location may produce the same difference in interaural phase when the period of the waveform is equal to or less than twice the maximum interaural difference in time-of-arrival. In this simplified example for human listeners, such location ambiguities will occur for frequencies with periods less than or equal to  $1312 \mu\text{s}$ . Here, a 760-Hz stimulus will produce an  $180^\circ$  difference in interaural phase when the stimulus is presented either at azimuth  $\pi/2$  radians or  $3\pi/2$  radians ( $90^\circ$  to the right or left). Hence, interaural phase information alone will not discriminate between these two locations. Similarly,



for all frequencies greater than 760 Hz, the interaural difference in signal phase produced for a source at any given azimuth will be perfectly matched by at least one other azimuth. The possibility of ambiguity in azimuth for interaural phase differences of mid-range and high-frequency signals suggests that phase information should be utilized in sound localization only for low-frequency signals. Furthermore, the smaller the head size, the higher the frequency limit for unambiguous localization via interaural phase. A small rodent with a maximum  $\Delta t$  of only 100  $\mu\text{s}$  will not experience ambiguous azimuths for phase differences of signals below 5000 Hz in frequency.

The perception of interaural differences in the phase of the fine structure is restricted to relatively low-frequency signals. Both physiological and behavioral observations indicate that the mammalian auditory system is unable to resolve interaural differences in signal phase for frequencies above some critical value. The critical value may differ for various species, and it is usually observed in the region of 1 kHz to 5 kHz (Klumpp and Eady 1956; Kiang et al. 1965; Rose et al. 1967; Anderson 1973; Brown et al. 1978a; Johnson 1980).

The evolutionary significance of interaural temporal processing for directional hearing is seen in physiological specializations dedicated to this function. Auditory specializations for measuring ITDs are first observed in the brainstem, where powerful endbulb synapses securely couple auditory nerve inputs to cochlear nucleus bushy cells. These calyceal endings faithfully transmit the timing of auditory inputs to binaural neurons in the medial and lateral superior olive. The ITD sensitivity of neurons in the olive and inferior colliculus have been measured with binaural beat stimuli that establish a dynamic time delay for tones or amplitude modulated noise in the two ears (Yin and Kuwada 1983; Batra et al. 1997; Ramachandran and May 2002). ITD-sensitive neurons appear to encode a specific time delay by their maximum (medial superior olive—MSO) or minimum discharge rates (lateral superior olive—LSO). A cross-species comparison of the distribution of best neural delays suggests an overrepresentation of ITD cues that fall within the biological constraints imposed by the effective radius of the head.

Although the geometrical model presented in Figure 5.1 may describe interaural differences in time-of-arrival rather accurately, the same is not true for interaural differences in signal level, or ear differences in HRTFs. In the case of ILDs, subtle variations in the shape of the skull and pinnae have a pronounced impact on the magnitude of the observed differences in interaural level. Using a Shilling artificial human head, Harris (1972) conducted measurements of ILDs with either no pinna, or large pinna or small pinna chosen to sample human pinna variations. These studies were conducted with a microphone diaphragm placed in the position of the tympanic membrane, and differences in the sound pressure level incident at each eardrum were measured as a function of the azimuth of the sound source, the frequency of the signal, and the size of the pinnae (large pinna, small pinna or no pinna). Harris's measurements, presented in Figure 5.2, show that at low frequencies (e.g., 500 Hz) ILDs were very small, while at high frequencies (e.g., 8 kHz) they were prominent. The results also

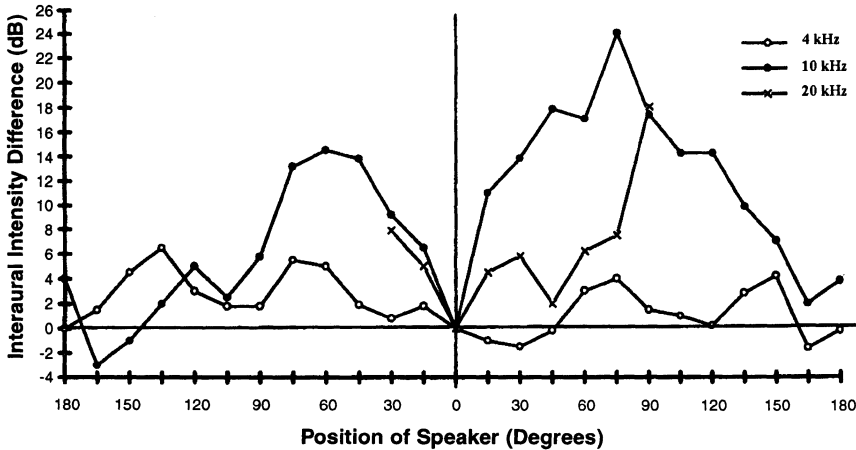


FIGURE 5.2. Sound shadows (ILDs) at representative frequencies produced by rotating a speaker around an artificial head fitted with large, small, or no pinnae. Points further from the center of the circle indicate that the signal level was more intense at the ear on that side of midline. (From Harris 1972. Reprinted with permission.)

indicate that this is an area in which mathematical models do not substitute for empirical measurements. For example, it is surprising that at some azimuth and frequency combinations, ILDs were greater for the no-pinnae condition than they were for either the large- or small-pinnae conditions.

Harrison and Downey (1970) used probe microphones placed by the tympanic membrane to measure ILDs in humans, rats, bats and squirrel monkeys. Their data showed that interaural level differences tended to increase with frequency, and they encountered very large ILDs with nonhuman subjects. Figure 5.3 displays ILDs for an individual squirrel monkey (*Saimiri sciureus*). In general, as signal frequency was increased, ILDs also increased, and at certain azimuth and frequency combinations ILDs could exceed 20 dB. However, because the magnitude of ILDs was influenced by small variations in the morphology of the head and pinnae, ILDs did not vary monotonically with changes in azimuth. It is possible that with tonal, or narrow-bandwidth signals two or more azimuths may give rise to the same overall ILDs, and sound position may then be ambiguous. Broad-bandwidth, high-frequency signals may be accurately localized via the ILD mechanism, however. Brown et al. (1978a) have argued that at each azimuth, the left and right ears will have a spectral transfer function, and the difference between the near- and far-ear functions will give rise to a binaural difference spectrum (Fig. 5.4). The properties of the binaural difference spectrum may be unique to each location, and if the stimulus were broadband, then accurate sound localization would be realized. It is possible that the binaural difference spectrum can also be used to derive sound source elevation as well (Rice et al. 1992), and this will be described in Section 5.

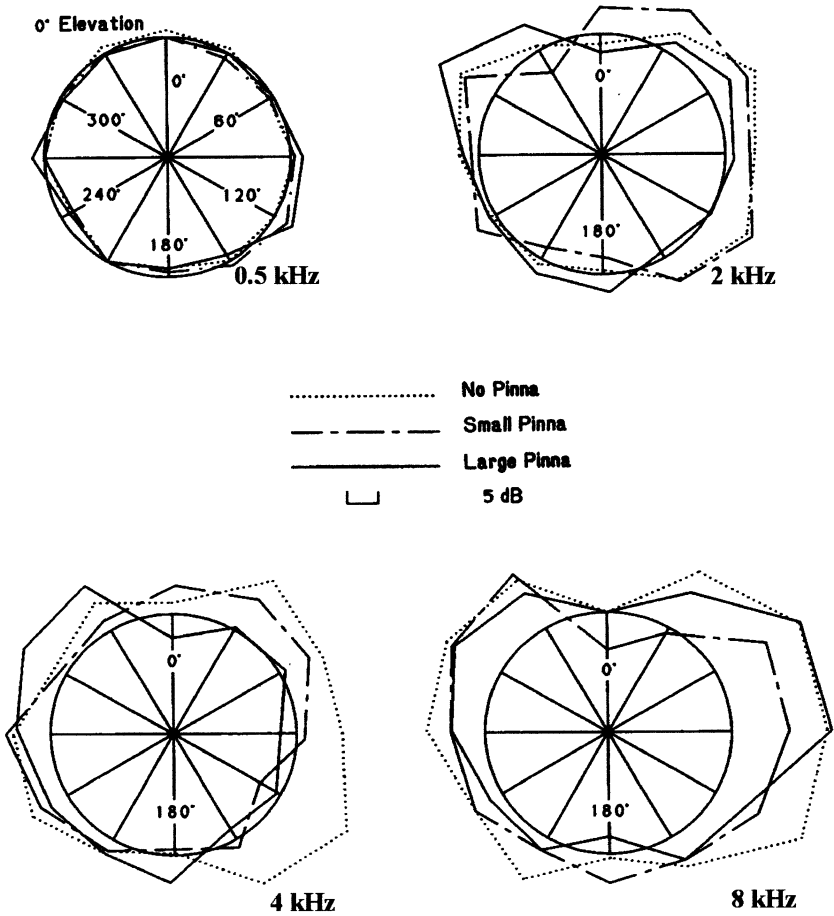


FIGURE 5.3. Interaural level differences measured in squirrel monkeys (*Saimiri sciurens*) as a function of the position of speaker azimuth ( $0^\circ$  to  $180^\circ$  on either side of zero azimuth) at three tone frequencies. (From Harrison and Downey 1970. Reprinted with permission.)

In summary, the localization of sound azimuth may be dependent on the perception of interaural differences in time-of-arrival, signal level, and spectral differences. At signal frequencies above the limit for which interaural phase differences become ambiguous, ILDs or spectral differences may become a viable cue. Thus, sound localization may be governed by a multicomponent perceptual system. It is possible that some mammals may be more dependent on one mechanism, while other mammals are more dependent on the other. Species differences in the relative development of several brainstem nuclei are consistent with this possibility. Furthermore, it is likely that head size differences and

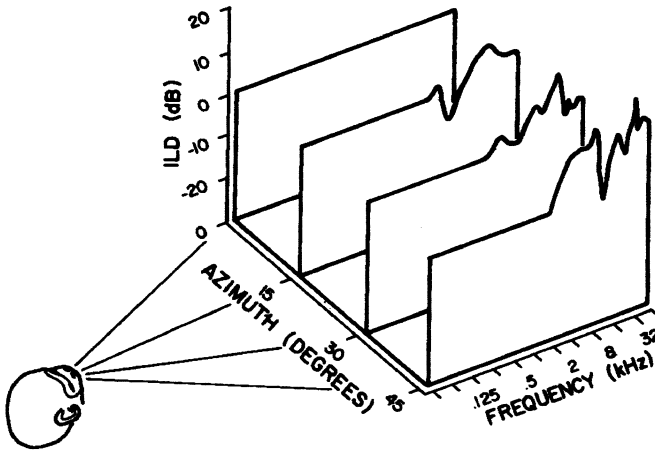


FIGURE 5.4. Hypothetical interaural sound pressure level differences as a function of the azimuth of the source and signal frequency. Negative level differences in the difference spectrum are generated when the signal level incident at the far ear exceeds that recorded at the near ear. (From Brown et al. 1978a. Reprinted with permission.)

pinna size morphology may amplify the significance of one mechanism relative to that for the other.

### 2.5 The Vertical Coordinate of Sound Location

The perception of source height, or elevation, may be particularly significant for the arboreal primates, marine mammals, and other nonterrestrial organisms. At mid-canopy levels in the rain forest and in marine habitats, biologically significant sounds may arise from positions above and below the listener, as well as from the right and left. If listeners were unable to move their pinnae, and if the right and left ears were acoustically symmetrical, then vertical localization would not involve binaural processing, unless, of course, listeners strategically cocked their heads (Menzel 1980). The relative level of different frequency components of a complex sound change as the location of the source is moved in elevation, and these variations in level are the property of the listener's HRTFs. Although the general properties of these transfer functions are similar for both ears, they are not bilaterally identical, and left- and right-ear asymmetries in the elevation-dependent sound transformation functions are important for vertical sound localization in cats (Musicant et al. 1990; Rice et al. 1992), humans (Shaw 1974a,b; Middlebrooks et al. 1989; Middlebrooks 1992, 1999), and barn owls *Tyto alba* (Payne 1962; Norberg 1977; Knudsen et al. 1979).

In human listeners (and probably in most terrestrial mammals), the perception of vertical position is largely dependent on the fact that the transformation func-

tion of the external ear is elevation dependent. Using high-frequency, broadband signals, the apparent spectral content of the sound changes with elevation of the stimulus (Butler 1969; Gardner 1973; Hebrank and Wright 1974; Kuhn 1979). Because the asymmetries and convolutions of the pinna and external ear canal must be relatively large compared to the wavelength of the signal for the expression of elevation dependent differences in the external ear transformation function, this cue would require relatively high-frequency, broadband signals, and high-frequency hearing (Shaw 1974a,b; Kuhn 1979; Wightman and Kistler 1989a,b). However, lower-frequency signals may reflect off the ground and the organism's torso in an elevation-dependent manner (Kuhn 1979, Brown et al. 1982), and it is possible that some degree of vertical localization is possible with low-frequency signals. Nearly all mammals have excellent high-frequency hearing, and this general trait in many cases may be at least as critical for vertical localization as it is for horizontal localization.

While humans have short, essentially fixed pinna, most terrestrial mammals have extended, mobile pinna; and asymmetries in pinna shape or orientation (Searle et al. 1975) may enhance vertical localization just as ear canal asymmetries enhance the perception of acoustic elevation in the barn owl, *Tyto alba* (Payne 1962; Norberg 1977; Knudsen et al. 1979). Although marine mammals have either no pinna or small pinnae, accurate vertical localization may still be possible (Renaud and Popper 1975). Much work remains to be conducted regarding marine mammal localization; it is unknown how sound is propagated around the head, torso, and ear canal of various marine mammals, and it is unknown if left-right asymmetries exist in the transformation functions for sounds presented at different elevations.

## 2.6 *The Distance Coordinate of Sound Location*

The perception of acoustic proximity (distance or depth) is very poorly understood, yet its analog has been well studied in the visual system. In visual perception, both binocular and monocular cues may provide information regarding the relative proximity of visual targets. The chief binocular cue is binocular disparity; a near object is seen from two slightly different angles by the two eyes. When the observer views a near object, more distant objects in the background will necessarily fall on different areas of the left and right retinas. A second binocular cue is convergence, the inward turn of the eyes required to maintain stereoscopic vision. This cue becomes more pronounced as the visual target is positioned progressively closer to the subject (Kaufman 1979). Because there is very little change in either of these binocular cues for two targets positioned at 10 m or 20 m, for example, it is likely that relative distance judgments for distal targets are more dependent on monocular cues than on binocular ones.

The monocular cues for distance perception (for a review see Kaufman 1979) include size constancy (retinal image size varies with changes in object distance); interposition (near objects are in front of, or partially obscure, more distant objects); linear perspective (parallel lines appear to converge at the ho-

rizon); textural perspective (the density of items per unit of retinal area increases with distance); aerial perspective (distant objects appear to lose their color saturation and appear to be tinged with blue); relative brightness (objects at greater distances from a light source have less luminance than do objects positioned closer to the source); and relative motion parallax (the apparent location of distant objects is shifted less by a change in the position of the viewer than are the perceived locations of closer objects).

In light of the richness of our appreciation of the cues underlying visual depth and distance perception, it is surprising that so little is known about either the putative perceptual cues underlying the perception of acoustic distance, or the abilities of various species to perceive differences in the proximity of acoustic events. Nevertheless, there is good reason to believe that the perception of acoustic proximity has undergone intense selection for many species. Payne (1962) showed that, in a totally darkened room, barn owls (*Tyto alba*) were able to accurately fly from an elevated perch to broadcasts of recordings of the rustling noise produced by the movements of a mouse. Because the barn owl flies headfirst, yet captures prey feet-first, it must be able to accurately estimate the acoustic azimuth, elevation and distance to be able to position its body for the strike. If the owl were unable to perceive acoustic proximity, it would risk breaking its descent either too soon or too late. Playback experiments have shown that great tits, *Parus major* (McGregor and Krebs 1984; McGregor et al. 1983), howler monkeys *Alouatta palliata* (Whitehead 1987), and gray-cheeked mangabeys *Lophocebus albigena* (Waser 1977) use acoustic cues to gauge distance and judge the possibility of incursions into one's territory by a rival individual or group. Distance perception is also important for prey capture and object avoidance in bats (Denzinger and Schnitzler 1998; Masters and Raver 2000). It is likely that the perception of acoustic distance is important for many species.

Although binocular vision is important for distance and depth perception, there is little available evidence to suggest that binaural hearing is either important, or unimportant, for the perception of acoustic proximity. It is likely that many of the monocular and binocular cues for distance perception have a rough analog in the acoustic domain.

### 2.6.1 Monaural Cues for Auditory Distance Perception

The prime candidates for monaural distance perception include:

1. Amplitude, sound level or auditory image constancy (the amplitude of the signal varies with distance usually in accordance with the inverse-square law (Gamble 1909; Coleman 1963). Hence, the raucous calls of the hornbill, *Bycanistes subcylindricus*, grow softer as the bird flies to a more distant part of the forest.
2. Frequency spectrum at near distances (von Békésy 1938). At distances less than 4 feet, the low-frequency components of complex signals are relatively more prominent than are mid-frequency and high-frequency components, and

as the source of the signal is moved progressively closer to the listener low-frequency components become even more prominent.

3. Frequency spectrum at far distances (Hornbostel 1923; Coleman 1963). The molecular absorption coefficient for sound in air depends on humidity, temperature, and frequency. At a temperature of 20°C, and a relative humidity of 50%, the absorption coefficient of a 10-kHz tone is about 20-fold greater than that for a 1-kHz tone (Nyborg and Mintzer 1955). Hence, high frequencies are attenuated more rapidly than are low frequencies, and at successively greater transmission distances, the frequency spectrum of complex signals shows a progressive loss of the high-frequency components (Waser and Brown 1986). This cue resembles aerial perspective in the visual domain. That is, just as more distant views are characterized by the loss of longer wavelength hues, more distant sounds are characterized by the loss of high-frequency components.
4. Reverberation. The temporal patterning of signals becomes “smeared” as the delayed reflected waves overlay the direct wave (Mershon and Bowers 1979). Hence, the ratio of direct to reflected waves can provide distance information. This phenomenon is more likely to provide usable information in forested habitats than it is in open habitats.
5. Temporal distortion. Changes in wind velocity, wind direction, or convection current flow result in changes in the duration and pitch of signals transmitted through a nonstationary medium. Signals broadcast from greater distances are probably more susceptible to disturbance by this phenomenon, but this has not been studied in detail (Brown and Gomez 1992).
6. Movement parallax. The relative location of distant sources is shifted less by a change in location of a listener than are the perceived locations of closer sources. This cue is a direct analog to relative motion parallax in the visual domain. It is probable that this cue requires rather large displacements in space for it to play a role in distance judgments for head cocking and other rotational movements of the head and neck may be insufficient to aid distance judgments in some situations (Simpson and Stanton 1973).

### 2.6.2 Binaural Cues for Auditory Distance Perception

Binaural cues for the perception of acoustic distance include:

1. Binaural intensity ratio. When the source of a signal is at a position other than 0° azimuth the signal may be greater in amplitude at the near ear relative to the amplitude of the signal at the far ear. This difference in sound amplitude, the binaural intensity ratio, varies as a function of head size, azimuth, signal frequency, and transmission distance (Hartley and Fry 1921; Firestone 1930).
2. Binaural differences in signal phase. In addition to the binaural intensity ratio, empirical measurements have shown that binaural differences in signal phase vary as a function of transmission distance as well as head size, azi-

- moth, and signal frequency (Hartley and Fry 1921). Thus, it is possible that binaural differences in signal phase may help cue transmission distance.
3. Acoustic field width. At the front row of the auditorium the orchestra may occupy a whole hemifield, while at the rear of an auditorium, the orchestra occupies a more restricted portion of the acoustic field. Hence, the perceived distance to an acoustic source that is not a point source varies inversely with the perceived width of the acoustic field. Although this putative cue is likely binaural in origin, it resembles the monocular cue of textural perspective in the visual domain.
  4. Scattered sound direction and field width. In forested habitats, sound becomes scattered by tree trunks. The greater the transmission distance, the greater the magnitude of the field width of the scattered sound, and the perceived width of the field of this scatter may influence distance judgments.

There are very little data to indicate the relative potency of the various putative monaural and binaural cues for judgments of distance, and much research remains to be done in this area. The utility of these cues for the perception of acoustic proximity depends on how reliably they change with distance. The initial cue listed above, auditory image constancy, is simply a change in signal amplitude, while all the other cues enumerated here are associated with a change in sound quality, sound distortion, or a change in sound characteristics at each ear. The only cue, which has received full examination, is auditory image constancy (e.g., amplitude constancy); however, studies of sound transmission in natural habitats have shown that amplitude may fluctuate 20 dB or more in short intervals of time (Wiley and Richards 1978; Waser and Brown 1986). Fluctuations of this magnitude may lead to errors in judgment of three or four doublings of acoustic distance (a sound presented at 25 m under unfavorable conditions may be received at a lower amplitude than the same sound broadcast at 100 m presented under more favorable conditions). Hence, sound amplitude per se is generally regarded as a poor index of transmission distance.

In all habitats, the natural environment degrades sounds, and these more complicated habitat-induced changes in sound quality may more reliably cue acoustic proximity. Brown and Waser (1988) have shown that exemplars of representative vocal classes are degraded differently by the acoustics of natural habitats. Changes in sound quality have been measured with respect to the frequency composition of the call and with respect to the temporal patterning of the signal (Brown and Waser 1988, Brown and Gomez 1992).

Figure 5.5 shows sound spectrograms of the blue monkey (*Cercopithecus mitis*) grunt utterance at the source (panel A), and three recordings of the same call after having been broadcast 100 m in savanna (panels B–D). While the signal displayed in panel B retains the overall structure of the source (panel A), the signal shown in panel C is missing the low-frequency portion of the call (the band of energy at about 500 Hz), and the signal displayed in panel D is missing the two higher-frequency components of the call (the bands of energy at about 1500 Hz, and 3000 Hz). These recordings were conducted in succes-



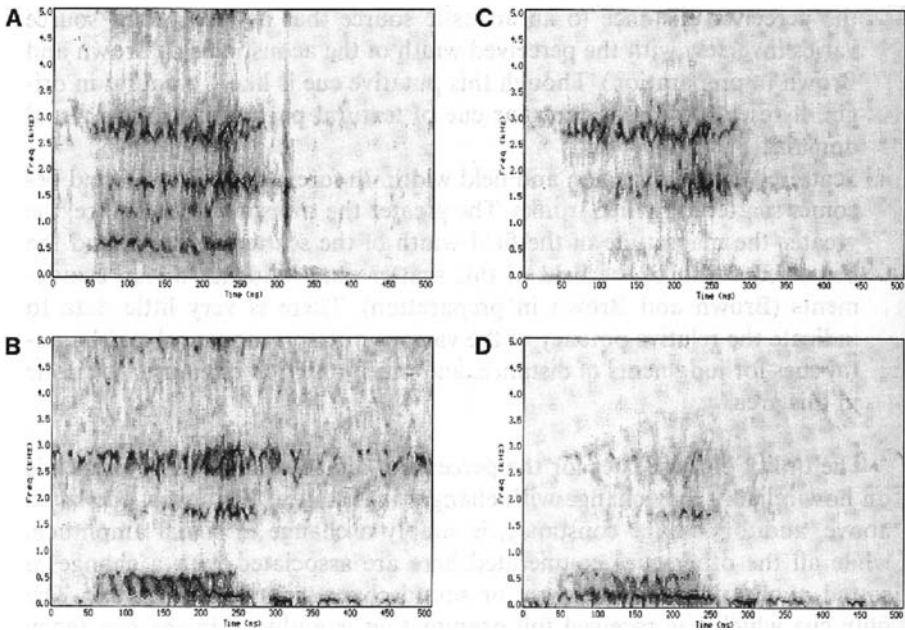


FIGURE 5.5. Sound spectrograms of a grunt call given by a blue monkey (*Cercopithecus mitis*). (A) The call at the source. The signal is composed of three energy bands centered at approximately 500 Hz, 1500 Hz, and 2500 Hz. (B–D) Spectrograms of the call recorded at a transmission distance of 100 m in the savanna habitat. (B) The recording was noisy but all elements of the call were present. (C) The 500-Hz-frequency band was absent. (D) The 1500-Hz and 2500-Hz-frequency bands were strongly attenuated. (From Brown and Gomez 1992. Reprinted with permission.)

sion under rather stable climatic conditions within a 2-hour interval at the same site (for a review of the factors in natural habitats that lead to different patterns of distortion see Brown and Gomez 1992). These recordings dramatize the fact that the structure of signals may be altered by the acoustics of the habitat. Environmentally induced degradation of acoustic signals occurs in probably all natural settings, and some types of distortion may be useful for estimating the distance to the source.

It is possible to adopt signal-processing techniques to quantitatively measure the magnitude of habitat-induced distortion of vocalizations (Brown and Waser 1988; Brown and Gomez 1992). The data show that some vocalizations (e.g., the blue monkey's boom) are relatively unchanged by the acoustics of the habitat, while other calls (e.g., the blue monkey's chirp or pyow) are more susceptible to degradation. The overall pattern of these degradation scores indicates that different utterances are degraded in different ways by environmental acoustics. Just as some vocalizations are ventriloquial while others are easily localized in azimuth (Brown 1982a), the present observations suggests that some

vocalizations may be good for revealing acoustic proximity, while other utterances may obscure the relative proximity of the vocalizer. Presumably, the presence or absence of “distance information” in various calls is relevant to the social function of different vocalizations. Many forest monkeys emit calls that appear to mark the position of the vocalizer. These calls may be involved in regulating the spacing, distribution, and movements of individuals out of visual contact.

Recent work has shown that soundscapes, the background sounds in natural habitats, contain highly nonrandom structures (Nelken et al. 1999), and selection may have favored the evolution of sound processing strategies that exploit the coherence in soundscapes to render signals more separable from the background, and hence more audible, and potentially more locatable. Thus, relative to their audibility in white noise, signals are more audible in masking noises, which exhibit comodulation (Hall et al. 1984; Moore 1999) common to the spectrotemporal fluctuations of natural soundscapes. It is possible that variations in signal structures that influence sound localization in both azimuth and distance are related to release from masking phenomena (Hall et al. 1984; Moore 1999), and the factors that influence signal detection in natural environments may be related to directional hearing. This is a promising area for future research.

### 3. Sound Localization Methodology

Many animals will orient toward, and approach, the origin of some sounds. The accuracy of approach has been used to study sound localization in the gray-cheeked mangabey monkey (*Lophocebus albigena*) (Waser 1977), tree frogs (*Hyla cinera* and *H. gratiosa*) (Feng et al. 1976), cats (Casseday and Neff 1973), and many other species. In some instances, food or some other reward has been used to maintain this behavior. In such approach procedures, the accuracy of localization is dependent on the ability of the auditory system to process a change in sensation associated with a change in the position of the source, and in the ability of the motor systems of the animal to accurately guide the subject towards the perceived location of the acoustic target. Species differences in the acuity of localization, measured by the approach procedure, may be due to differences in the precision of the perceptual system, or alternatively these apparent acuity differences may be due to variations in the accuracy of motor systems.

Orientation paradigms have also been developed to measure the acuity of localization. With these methods a head turn or body turn is used to indicate the perception of sound direction (Knudsen and Konishi 1978; Knudsen et al. 1979; Whittington et al. 1981; Brown 1982a; Perrot et al. 1987; Makous and Middlebrooks 1990; Huang and May 1996a,b; May and Huang 1996). Figure 5.6 illustrates this procedure using results from a food-reinforced orientation task. The acoustic stimuli were brief bursts of broadband noise that were presented from one of eight randomly selected locations in an anechoic room.

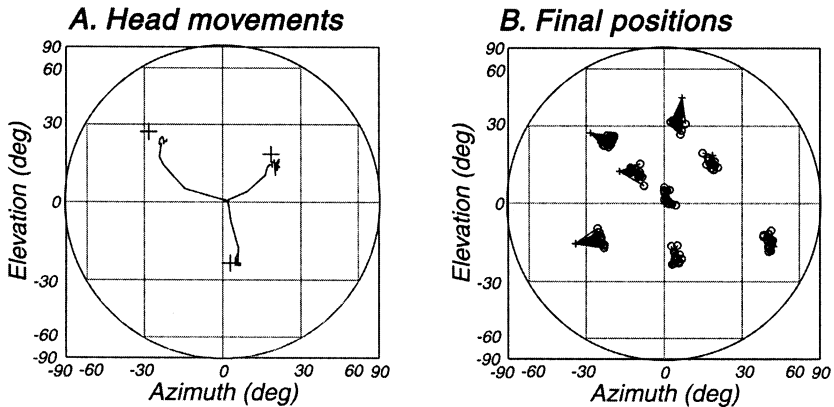


FIGURE 5.6. Sound orientation accuracy of a representative cat. The subject was required to direct its head to the source of a broad bandwidth noise burst. The sound source changed randomly across eight possible speaker locations (*plus symbols*). (A) The path of head movements from a fixation point ( $0^\circ$  azimuth,  $0^\circ$  elevation) to a final stable position for tests with three different target locations. (B) The final head orientation for all tests in one session. *Lines* connect each response to the actual speaker location. (Adapted from Huang and May, 1996a.)

Figure 5.6A tracks the cat's head movements on three successive trials from a standardized fixation point ( $0^\circ$  azimuth,  $0^\circ$  elevation) to the location of the speakers (*plus symbols*). In each case, the subject rapidly acquires the target location then holds the response for several seconds as it waits for a visual cue that signals the end of the trial. Figure 5.6B shows the cat's final stable head orientation for all trials in the testing session (*open symbols*).

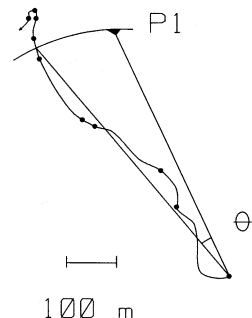
The results shown in Figure 5.6 indicate that spectrally rich sounds evoke orientation responses that are accurate in the vertical as well as the horizontal plane (May and Huang 1996). Both orientation and approach procedures are categorized as egocentric methods (Brown and May 1990). Here localization is made not with reference to an external acoustic marker, but rather with reference to the subject's physical orientation in space. With egocentric procedures, differences in sound localization acuity may be due to limitations in the accuracy of the perceptual system, or to limitations of the motor system.

Behavioral tasks in which listeners have been trained to operate response levers to indicate the detection of a change in sound location have been used with both human (Mills 1958) and animal subjects (Brown et al. 1978a). These ear-centered, or otocentric, procedures are designed to assess the acuity of the perceptual system, and they do not require the participation of the spatial/motor system (Brown and May 1990). Hence, with these procedures listeners report the detection of a change in sound location, but they do not indicate where the sound originated relative to their own orientation.

Given the procedural variations possible between these different methodologies it is important to note that independent measurements of sound localization acuity in normal animals appear to be remarkably consistent and robust. There is a high degree of agreement in the results using both egocentric and otocentric methods within (Heffner and Heffner 1988d), and between laboratories (Brown and May 1990). Using the approach procedure under field conditions in the natural habitat, Waser (1977) showed that one of the Cercopithecoidea monkeys, the gray-cheeked mangabey (*Lophocebus albigena*), was able to localize the whoopgobble, a complex long-distance vocalization (Brown 1989), with an average error of only  $6^\circ$  (Fig. 5.7). Under laboratory conditions, using otocentric methods with two other species of the Cercopithecoidea monkeys (*Macaca nemestrina* and *M. mulatta*), the localization of complex vocal signals ranged from  $3^\circ$  to  $15^\circ$  depending on the specific acoustical characteristics of the utterance (Brown et al. 1978b, 1979). The mean localization error of macaque monkey broad bandwidth or frequency-modulated calls, those that are most comparable to the mangabey's gobble, is about  $3^\circ$ . It is remarkable that a field phonotaxis study conducted in the monkey's native habitat in Uganda and a laboratory investigation yield results that are so similar.

When comparable stimuli are used, the congruence in the data produced by different laboratories employing different methods is even more striking. Figure 5.8 shows ITD thresholds measured using earphones (Houben and Gourevitch 1979), and those calculated from free-field measurements of the acuity of localization (Brown et al. 1978a) as a function of tone frequency. Data for human subjects (Klumpp and Eady 1956) are compared with macaque monkey data. These data show that the physical characteristics of the signal have a strong impact on the accuracy of sound localization. This is true for both simple synthetic signals, such as tones, and complex natural signals, such as vocalizations.

FIGURE 5.7. Sound initiated approach in gray-cheeked mangabeys (*Lophocebus albigena*) evoked by the playback of a whoopgobble vocalization. The playback was conducted in the Kibale forest in Uganda with native populations of mangabeys. P1 is the location of the broadcast loud speaker. At the time of broadcast the focal subject was located at the apex of angle  $\theta$ . Mangabeys are arboreal, and only rarely descend to the ground. The track that the monkey takes is then partially governed by the location of the branches of adjacent trees in the forest. Owing to the thickness of rain forest vegetation field assistants cannot follow the path of the monkey directly, but are able to observe the movement of the monkey to identified trees in the forest (denoted by periods in the figure). The angle  $\theta$  is the discrepancy between the mean direction of approach and the playback site. (From Waser 1977. Reprinted with permission.)



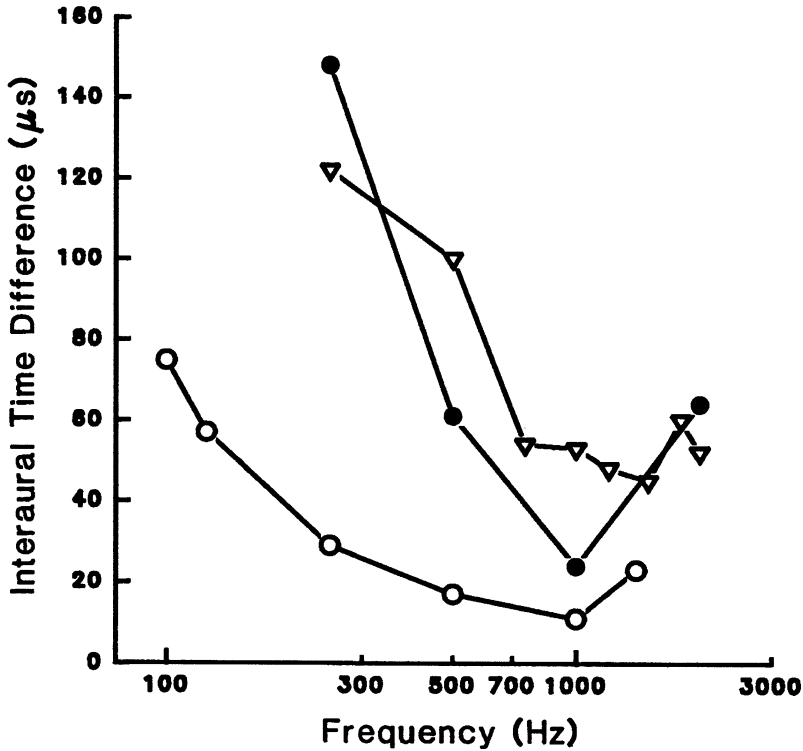


FIGURE 5.8. ITD thresholds. Thresholds, measured in microseconds, are displayed as a function of signal frequency. Macaque monkeys (*Macaca mulatta* and *M. nemestrina*) thresholds (black squares) are transposed from free-field localization measurements (Brown et al. 1978a); interaural time difference thresholds measured with earphones from monkeys (open triangles: Houben and Gourevitch 1979); and humans (open circles: Klumpp and Eady 1956). (From Brown et al. 1978a. Reprinted with permission.)

Furthermore, the data show that measurements of a species' acuity for sound localization are robust, and relatively independent of method. These observations indicate that it is possible to measure with high precision the acuity of sound localization that is representative of the abilities of the species, but that the data derived are dependent on the physical characteristics of the test signal.

## 4. The Acuity of Sound Localization

### 4.1 The Perception of Acoustic Azimuth

The just detectable change in the position of the sound source, the minimum audible angle (MAA), has generally been regarded as the most precise index of

the acuity of localization. Figure 5.8 presents individual psychometric sound localization functions for three macaque monkeys. The test signal was a macaque coo vocalization. The psychometric functions were derived from monkeys who had been trained to hold a contact-sensitive key when sounds were pulsed repetitively from a source at  $0^\circ$  azimuth, directly in front of the monkey, and release contact with the key when the sound was pulsed from a source at any other azimuth. The monkey's rate of guessing (its catch-trial rate) was very low, less than 8% (this rate is displayed over the  $0^\circ$  point in Fig. 5.9). The monkey's ability to detect a change in the azimuth of the sound source increased with the magnitude of change in source location reaching about 100% correct by  $30^\circ$ . These psychometric functions conform to the class ogive shape (Cain and Marks 1971), and the 50% correct detection point (the MAA) is measured in degrees and calculated from the psychometric functions.

Investigators have tended to measure the acuity of directional hearing with biologically significant stimuli, such as vocalizations (Feng et al. 1976; Waser 1977; Brown et al. 1978a, 1979; Rheinlaender et al. 1979), or more commonly with synthetic signals that are either simple, such as pure tones (Casseday and Neff 1973; Terhune 1974; Brown et al. 1978a; Heffner and Heffner 1982), or spectrally more complex, such as clicks or noise bursts (Ravizza and Masterton 1972; Brown et al. 1980; Heffner and Heffner 1982, 1983, 1987, 1988a,b). Biologically significant signals have tended to be used with phonotactic proce-

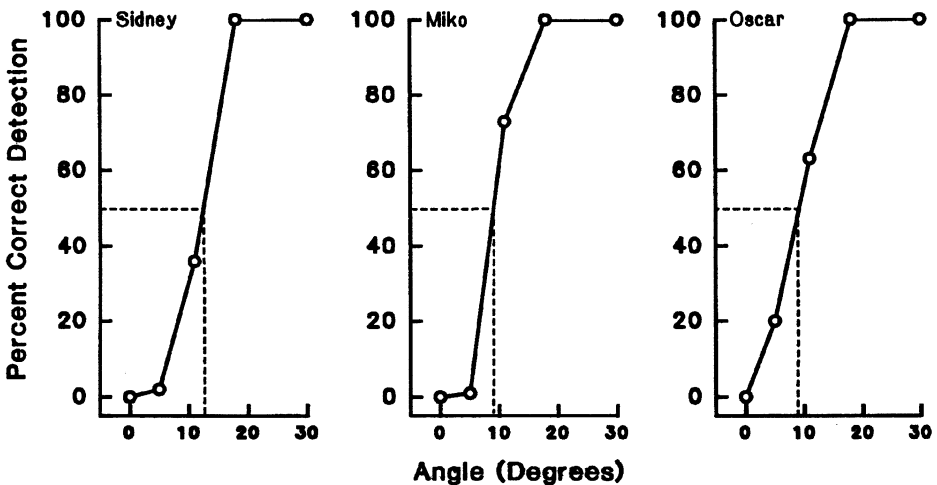


FIGURE 5.9. Psychometric functions for the localization of a macaque coo call. Functions are shown for three individual monkeys (Sidney, Miko, and Oscar). The monkey's rate of guessing (catch-trial rate) is displayed over the zero-degree point, and the monkey's percentage of correct detection for the trials presented at the four comparison locations increased with angle. The calculation of the MAA is shown by the *dashed line*. (From Brown et al. 1979. Reprinted with permission.)

dures or in studies in which the relative locatability of various natural signals was the topic of interest, while synthetic signals have tended to be used in studies that have focused on assessing the limits of the perceptual system.

## 4.2 *Biologically Significant Signals*

The different physical characteristics of various classes of complex natural stimuli, such as vocalizations, may influence the acuity of localization. In macaque monkeys, the effective bandwidth (or magnitude of frequency modulation) of the dominant frequency band of the call has a strong effect on sound localization (Fig. 5.10). Increments in the effective bandwidth of the signal enhance the accuracy of localization. MAAs for macaque coo calls span approximately a fivefold range, from about 3° to 15°. Macaque monkeys also produce a wide variety of noisy barks, grunts, and growls, and these harsh sounding, atonal, broad-bandwidth calls are all accurately localized as well (Brown et al. 1979). Complex natural signals that exhibit a broad effective bandwidth (produced either by frequency modulating a relatively tonal sound, or by generating an atonal, broad-bandwidth sound) are probably localized at the limits of resolution of the organism's perceptual system (Brown 1982b; May et al. 1986). The mate attracting calls, rallying calls, and position marking calls given by a wide variety of mammals typically exhibit a broad effective bandwidth, that likely promotes sound localization at the listener's limit of resolution.

## 4.3 *Pure Tones*

Comparative data for the localization of pure tones as a function of frequency are shown in Figure 5.11. While the human data suggest that stimulus frequency has a relatively modest effect on the localization of tones (Mills 1958), it tends to have a pronounced effect on the accuracy of localization by nonhuman mammals. At the best frequency, macaque monkeys (*Macaca mulatta* and *M. nemestrina*) (Brown et al. 1978a), harbor seals (*Phoca vitulina*) (Terhune 1974), and elephants (*Elephas maximus*) (Heffner and Heffner 1982) exhibit a resolution of about 4°, while human listeners are able to resolve about 1° (Mills 1958). At their worst frequency, human subjects are still able to resolve angles of about 3°, while most of the other mammals tested may require angles of 20° or more. Thus, human subjects tend to be more accurate at localizing the source of sounds across the frequency spectrum than are most other mammals. Testing with pure tones has almost exclusively been conducted with signals that are gated on and off slowly, and that are not modulated in amplitude so that the envelopes of the waveforms do not provide information that may influence localization. Under these conditions, human listeners localize low-frequency tones with a mechanism sensitive to interaural time differences, while the localization of high-frequency tones is governed by a mechanism sensitive to ILDs (Mills 1960). The same frequency effects have been shown to hold for monkeys (*M. mulatta* and *M.*

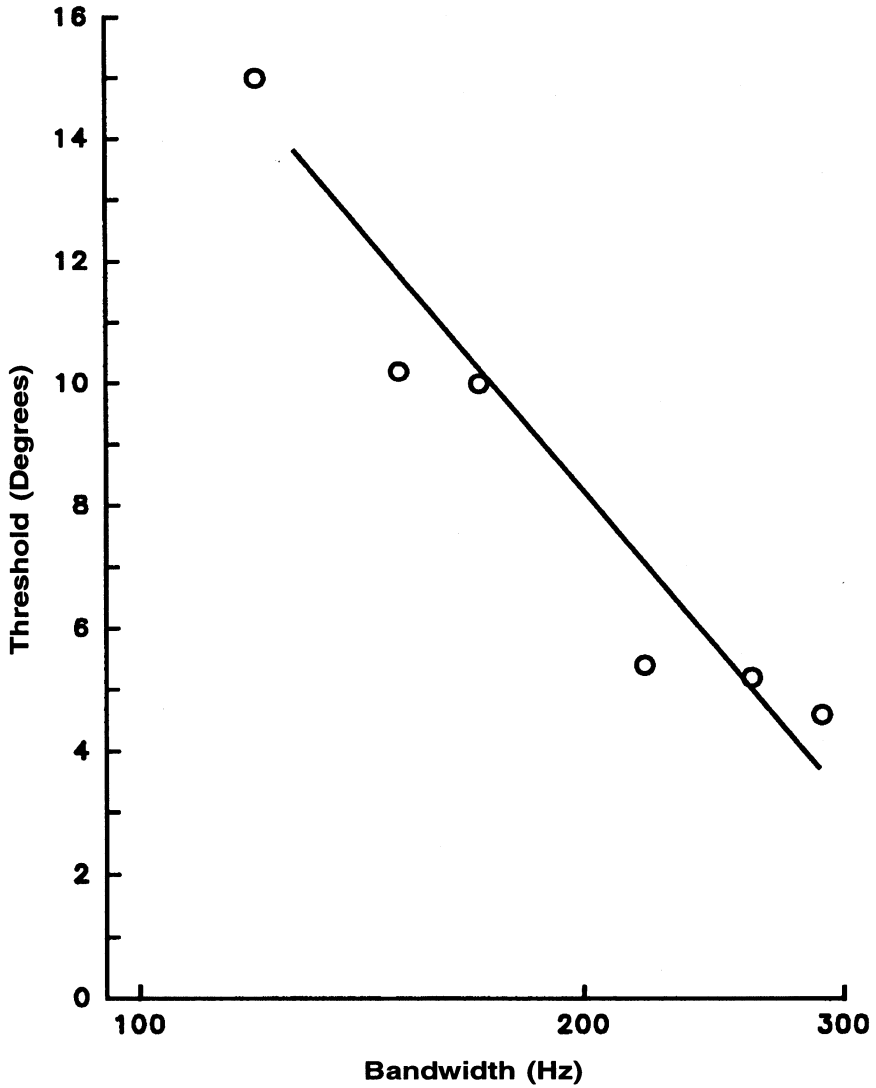


FIGURE 5.10. Macaque monkey (*Macaca mulatta* and *M. nemestrina*) localization thresholds for six coo vocalizations displayed as a function of the effective bandwidth (frequency modulated bandwidth) of the dominant band of the call. The correlation between threshold and call bandwidth was  $-0.59$ . (From Brown et al. 1978b. Reprinted with permission.)



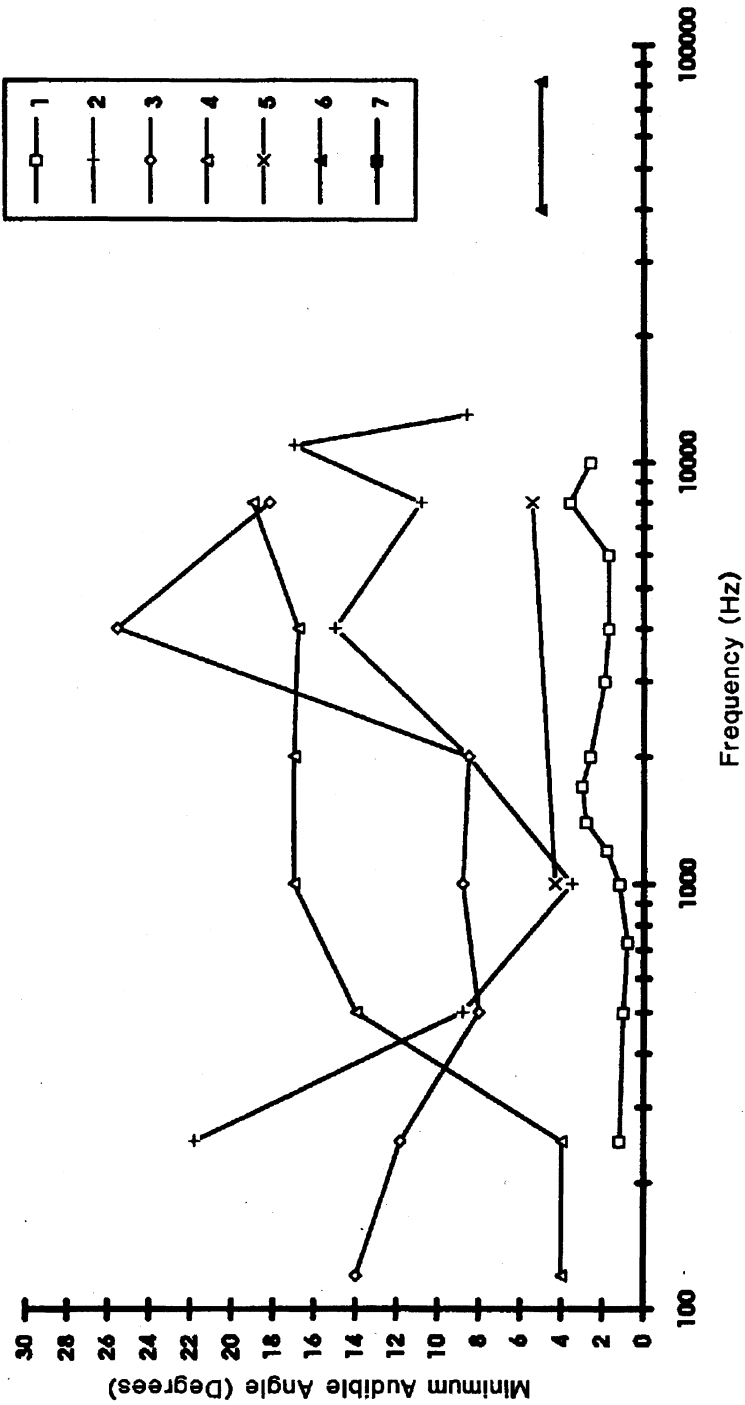


FIGURE 5.11. Sound localization acuity (MAAs) of tones as a function of signal frequency for representative mammals. 1, Human (*Homo sapiens*) (Mills 1958); 2, rhesus and pig-tailed monkeys (*Macaca mulatta* and *M. nemestrina*) (Brown et al. 1978a); 3, cat (*Felis catus*) (Casseday and Neff 1973); 4, elephant (*Elephas maximus*) (Heffner and Heffner, 1982); 5, harbor seal (*Phoca vitulina*) (Terhune 1974); 6, little brown bat (*Myotis oxygnathos*) (Ayrapet'yants and Konstantinov 1974); 7, greater horseshoe bat (*Rhinolophos ferrumequinum*) (Ayrapet'yants and Konstantinov 1974). (Fay 1988. Adapted with permission.)

*nemestrina*) (Brown et al. 1978a; Houben and Gourevitch 1979), and are presumed to apply for most other mammals as well.

The literature on the comparative localization of tones suggests that both mechanism for localization by human subjects are equally accurate, while in most other mammals one mechanism may be less accurate, and perhaps less significant, than the other. In this context, an extensive physiological and anatomical literature (Heffner and Masterton 1990) has shown that high-frequency localization primarily involves brainstem structures in the medial nucleus of the trapezoid body (MNTB) and LSO, while low-frequency localization primarily involves structures in the MSO. The relative development of these nuclei varies across mammals; in some species the MNTB–LSO system is undeveloped or nearly absent, while in other species the MSO system is undeveloped or nearly absent. In general, as the physical size of the mammal increases, the greater the development of the MSO system, and a concomitant reduction in the MNTB–LSO system is observed (Heffner and Masterton 1990). Thus, variations in the development of auditory structures in the ascending pathway may underlie species differences in their ability to fully utilize interaural time-of-arrival difference cues, or ILD cues. These variations may account for the observed species differences in the pure tone localization data (Fig. 5.11). However, while human subjects localize high-frequency tones well, their MNTB–LSO system is only marginally developed. Hence, although it appears that much is understood regarding the anatomical and physiological mechanisms subserving sound localization, significant puzzles still remain.

#### 4.4 Complex Stimuli

Comparative data for the localization of complex stimuli (e.g., vocalizations, clicks or noise bursts) are displayed in Figure 5.12. Here the MAA is plotted in reference to head size. As noted in Section 2.2, all other things being equal, both ITDs and ILDs should increase with head size. Thus, large mammals should exhibit greater sound localization acuity simply because the physical magnitude of these interaural cues increase with head size. This trend is generally observed (Fig. 5.12). However, the correlation between threshold and head size is only  $-0.32$ . Hence, some mammals are either significantly less sensitive, or more sensitive, to sound direction than would be expected by the size of their heads. Species located below the diagonal regression line shown in Figure 5.12 have better localization acuity than would be expected by their head size, while those positioned above the regression line have less acute directional hearing than would be expected. Thus, regardless of the magnitude of the physical cues available for localization, some species are particularly good localizers, while others are not.

How can these differences in the relative acuity for directional hearing be explained? Four species [gopher (Go), blind mole rat (Bm), pallid bat (Pb), dolphin (Do)] are particularly discrepant from the others tested. The gopher and blind mole rat are fossorial species, spending most of their time under-

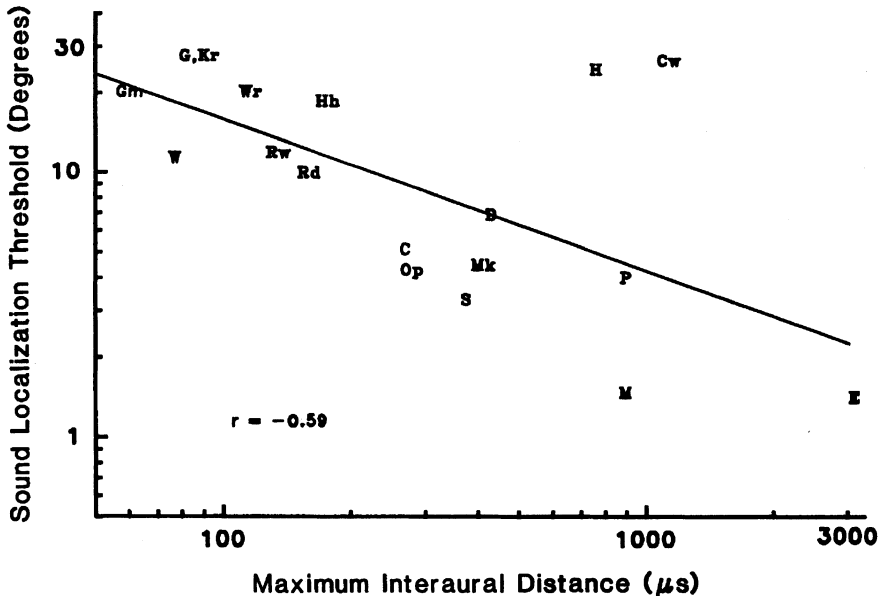


FIGURE 5.12. Sound localization threshold as a function of head size in 18 representative mammals. Acuity is displayed for a broad bandwidth sound, noise, or a click stimulus in the horizontal plane. Gm, grasshopper mouse (*Onychomys leucogaster*) (Heffner and Heffner 1988a); W, least weasel (*Mustela nivalis*) (Heffner and Heffner 1987); G, gerbil (*Meriones unguiculatus*) (Heffner and Heffner 1988c); Kr, kangaroo rat (*Dipodomys merriami*) (Heffner and Masterton 1980); Rv, wild Norway rat (*Rattus norvegicus*) (Heffner and Heffner 1985); Rd, domestic Norway rat and Wistar albino rat (*R. norvegicus*) (Kelly 1980); Wr, wood rat (*Neotoma floridiana*) (Heffner and Heffner, 1988a); Hb, hedgehog (*Paraechinus hypomelas*) (Chambers 1971); C, cat (*Felis catus*) (Heffner and Heffner 1988d); Op, opossum (*Didelphis virginiana*) (Ravizza and Masterton 1972); S, harbor seal (*Phoca vitulina*) (Terhune 1974); Mk, rhesus and pig-tailed macaque monkeys (*Macaca mulatta*) and (*M. nemestrina*) (Brown et al. 1980); D, dog (*Canis canis*) (H.E. Heffner, unpublished); H, horse (*Equus caballus*) (Heffner and Heffner 1984); M, human (*Homo sapiens*) (Heffner and Heffner 1988c); P, domestic pig (*Sus scrofa*) (Heffner and Heffner 1989); Cw, cattle (*Bos taurus*) (Heffner and Heffner 1992); E, elephant (*Elephas maximus*) (Heffner and Heffner 1982).

ground. These two species, along with the naked mole rat (Nm), have degenerate hearing characterized by poor sensitivity and poor high-frequency hearing (Heffner et al. 1987), and their sound localization acuity is also impaired. Thus, radiation into a niche in which hearing in general, and sound localization in particular, are less important biologically may result in a comparative reduction of these sensory capacities. The other two highly atypical species, the dolphin (*Tursiops truncatus*) and pallid bat (*Antrozous pallidus*), are echolocators, and

selection for some forms of echolocation may also heighten sound localization acuity. If these four species are removed from the correlation, the association between head size and localization acuity increases to  $-0.57$ . The corresponding correlation has improved substantially, but much of the variance in the association between these two variables has not been accounted for. It has been argued that the relationship between vision and sound localization may be an important factor in explaining some of this variance.

In a classic paper, Harrison and Irving (1966) argued that accurate sound localization abilities are particularly important for redirecting the site of gaze for species with high-acuity tunnel vision. That is, the horizontal width of the field of high-acuity vision tends to be much narrower in animals with high-acuity binocular visual systems (such as primates) compared to animals with nonoverlapping hemispheric visual systems (such as rabbits). In most mammals, ganglion cell density varies across the vertical and horizontal coordinates of the retina, and regions of high ganglion cell density are associated with high acuity vision. Heffner and Heffner (1988c) have defined the region of best vision as that portion of the retina in which the ganglion cell density is at least 75% of the maximum observed for that species. Using this approach they have shown that mammals with comparatively narrow fields of best vision have better localization acuity compared to those with broader fields of best vision. The relationship between sound localization and best visual field width breaks down for burrowing fossorial mammals which have radiated into the subterranean habitat, and which in turn exhibit a comparative blunting of both the visual and acoustic senses. Apparently, just as acute vision is superfluous in the absence of light, acute hearing and sound localization is of little value in the absence of a free field. That is, subterranean tunnels may channel sound similarly to a waveguide, and the resulting absence of a free field may change the direction of selection for acoustic processing. In general, these observations support the notion that for many species of mammals, one key function of directional hearing systems is to acoustically guide the orientation of the visual system.

#### *4.5 Plasticity and Sound Localization Acuity*

Anecdotal reports have long suggested that some blind humans appear to develop unusually keen auditory abilities. Specifically, the perception of acoustic space in some blind individuals has appeared to significantly exceed the abilities of subjects with normal sight. These reports raise the possibility that perceptual compensation may result when visual processing centers have been reassigned to acoustic processing following the onset of blindness. Thus, it is possible that the loss of use of one sensory modality may lead to a reorganization of the cortex to favor the processing of the remaining viable sensory modalities. Recent physiological studies have obtained results consistent with the idea that early blindness may result in cross-modal reorganization of the cortex, and this reorganization may produce compensatory effects for sound localization (Kujala

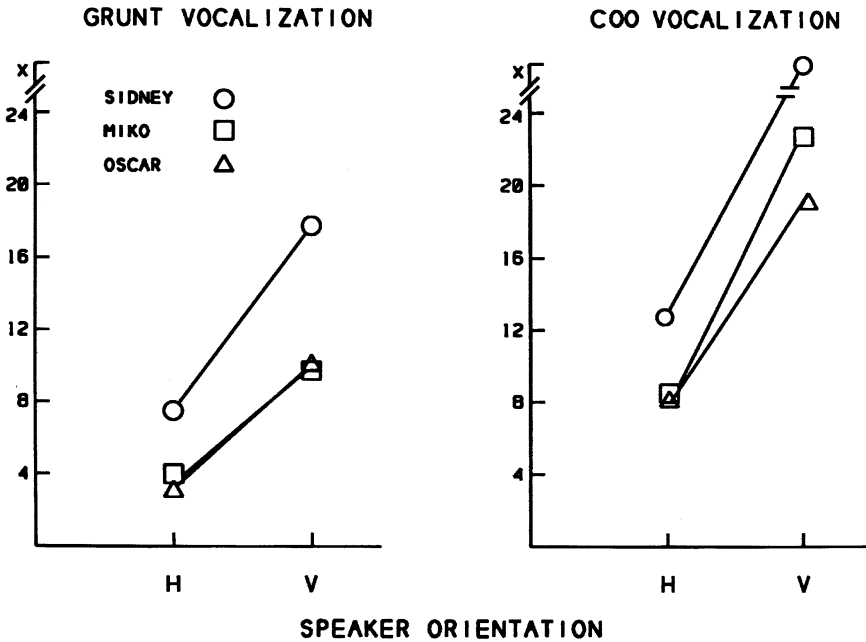


FIGURE 5.13. Horizontal (H) and vertical (V) minimum audible angles for a macaque grunt vocalization, and a macaque coo vocalization for three macaque monkeys (Sidney, Miko, and Oscar). An X indicates that the performance level of the subject never exceeded chance. (From Brown et al. 1982. Reprinted with permission.)

et al. 1992, 1995, 2000). Behavioral data consistent with this hypothesis have been reported for both cats and ferrets. When testing was conducted at a reference azimuth of 0°, MAAs were 16° and 15°, respectively, for ferrets with normal vision compared to those deprived of binocular vision at infancy (King and Parsons 1999). However, when testing was conducted at a reference azimuth of 45°, MAAs for the normal and visually deprived groups were 34° and 16° respectively (King and Parsons 1999). Thus, visual deprivation resulted in an improvement in the acuity of spatial hearing for stimuli located at lateral azimuths, but not at midline locations. King and Parsons (1999) also compared minimal audible angles for subjects blinded at adulthood, and they obtained a similar pattern of results. A complementary pattern of findings has also been reported for cats (Rauschecker and Kniepert 1994). Visually deprived cats showed enhanced sound localization abilities for signals broadcast from lateral and rear positions relative to normally sighted controls. This trend was strong for cats deprived of vision in infancy, and only approached (but did not achieve) statistical significance in adult deprived cats. Recent studies with humans have found that visually impaired, but not totally blind, subjects localize sounds with less accuracy than sighted controls (Lessard et al. 1998). However, 50% of the

subjects who were totally blind were superior to sighted controls in a monaural localization task (Lessard et al. 1998). Further evidence suggests, that like cats and ferrets, blindness in humans may have a more pronounced effect for the localization of sounds presented from peripheral locations relative to localization near the midline (Roder et al. 1999). In concert, these findings support the concept of compensatory plasticity; however, the etiology and severity of blindness, as well as its age at onset, may influence its significance for directional hearing.

Although the literature on plasticity has implicated changes in cortical structures, it is also possible that plasticity is expressed by changes at subcortical sites. For example, physiological studies have shown that metabotropic receptors in the dorsal cochlear nucleus (DCN) are capable of modulating synaptic transmission in a manner that resembles neural plasticity in the cerebellum (Molitor and Manis 1997; Devor 2000). Furthermore, the DCN is implicated in spectral processing (Spirou and Young 1991; Nelken and Young 1994), and the resolution of variations in spectral content may be particularly important for the resolution of front/back confusions and the localization of lateral azimuths, spatial regions particularly susceptible to the effects of blindness.

#### 4.5.1 The Perception of Acoustic Elevation

The literature is much more limited concerning the accuracy of perception of acoustic elevation. In arboreal living species, or in marine mammals, the determination of acoustic elevation may be as significant as the determination of azimuth. Vertical and horizontal minimum audible angles for primate vocalizations are shown for macaque monkeys (*M. mulatta* and *M. nemestrina*) in Figure 5.13. The test vocalizations were a macaque coo call, and a macaque grunt call. The grunt, which is broader in bandwidth, was localized more accurately than the coo. The median vertical localization thresholds were approximately 9° and 20°, respectively. For these same signals, the acuity of vertical localization was approximately two to three times less accurate than was localization in the horizontal plane. High-frequency hearing and high-frequency broadband stimuli are important for accurate vertical localization. If the signal contains sufficient high-frequency information, macaque monkeys may detect vertical displacements of only 3 to 4° (Fig. 5.14). This observation corresponds with the expectations based on the cues for perception of elevation discussed in Section 2.5. However, as shown in Figure 5.13, it is likely that the perception of sound azimuth is more accurate than is the perception of elevation for most signals.

Table 5.1 presents the acuity of vertical localization for representative mammals for the best signals tested. With a vertical acuity of 23°, the chinchilla (*Chinchilla laniger*) (Heffner et al. 1995) was the least acute mammal tested, while the bottlenose dolphin (*Tursiops truncatus*) (Renaud and Popper 1975) at 2° was the most precise vertical localizer. However, the literature is too sparse to permit much exploration of the role of pinna shape or size, visual field size,

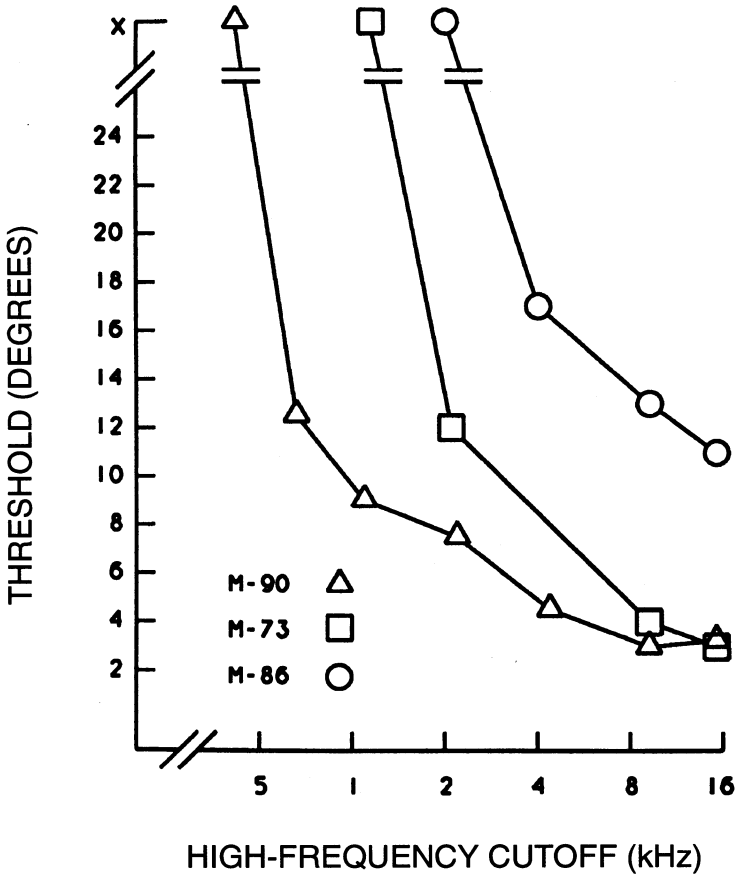


FIGURE 5.14. Vertical minimum audible angles for band-limited noise as a function of the high-frequency cutoff of the noise band for three macaque monkeys (M-90, M-73, and M-86). The low-frequency limit of the noise was 125 Hz. An X indicates that the performance level of the subject did not exceed chance. (From Brown et al. 1982. Reprinted with permission.)

or brainstem anatomical correlates with vertical acuity. Hopefully, investigators in the future will direct further attention to the problem of the perception of elevation.

### 5. Spectral Cues for Sound Localization

Behavioral assessments of the perception of sound source elevation by human listeners have contributed greatly to our current understanding of the role of spectral information in directional hearing. Although the basic principles of

TABLE 5.1. Vertical localization acuity in representative mammals.

Group	Species	Acuity	Source
Rodentia	Chinchilla	23°	Heffner et al. (1995)
Marsupialia	Opossum	13°	Ravizza and Masterton (1972)
Carnivora	Cat	4°	Martin and Webster (1987)
Primate	Rhesus/pig-tailed monkey	3°	Brown et al. (1982)
	Human	3°	Wettschurek (1973)
Cetacea	Dolphin	2°	Renaud and Popper (1975)

The data summarized in this table are rounded to the nearest integer, and are for the best signal tested. In some instances the test signal was a pure tone; in most cases, however, the best test signal was a band of noise, a click, or a species-specific vocalization.

these processes have been known for decades (Hebrank and Wright 1974; Butler and Belendiuk 1977; Watkins 1978), the maturation of digital signal processing techniques has resulted in significant recent advancements for psychoacoustic and physiological research in this area of the hearing sciences. Now, the salient directional features of human HRTFs are known in sufficient detail to allow the simulation of realistic auditory environments with earphones and other closed-field acoustic systems (Wightman and Kistler 1989b; Carlile and Pralong 1994; Pralong 1996; Kulkarni and Colburn 1998). In the future, these so-called virtual sound fields are likely to become a routine dimension of audio devices ranging from home entertainment centers to assistive aids for the hearing impaired.

In the laboratory, HRTF-based sounds provide an important functional context for exploring how spatial information is derived from the spectrotemporal properties of complex acoustic stimuli at processing levels ranging from the auditory nerve to cortex (Young et al. 1992; Imig et al. 1997; May and Huang 1997; Delgutte et al. 1999; Xu et al. 1999), just as ITD and ILD testing procedures have led to a better understanding of the binaural auditory system (Moore 1991). Much of our current knowledge regarding the auditory processing of spectral cues for sound localization has been gained from electrophysiological studies of the domestic cat.

Functional interpretations of the neural response patterns linked to directional hearing have been made possible by a long history of psychoacoustical studies in cats. The natural sound localization abilities of the cat have been described over a variety of stimulus conditions (Casseday and Neff 1973; Martin and Webster 1987; Heffner and Heffner 1988d; Populin and Yin 1998a), and the information processing roles of the major ascending auditory pathways have been confirmed by evaluating the behavioral deficits that follow surgical lesioning procedures (Moore et al. 1974; Casseday and Neff 1975; Neff and Casseday 1977; May 2000). This work suggests that the biological necessity for accurate sound localization has exerted a profound influence on the information processing pathways of the auditory system. Anatomical specializations for processing ILD and ITD cues are obvious in the striking binaural innervation patterns of the superior olive. Selectivity for the spectral features of complex sounds is



created by the frequency-dependent convergence of inhibitory inputs within the auditory brainstem (Spirou et al. 1993, Imig et al. 2000). These neural networks are more difficult to distinguish anatomically but no less important in the auditory behaviors of the cat (Sutherland et al. 1998, May 2000).

### 5.1 The HRTF of the Cat

The filtering properties of the cat's head and pinna are known in detail and provide biologically relevant stimulus parameters for evaluating the neural and perceptual basis of directional hearing (Musicant et al. 1990; Rice et al. 1992). Representative HRTFs of the cat are shown in Figure 5.15 using measurements from the study of Rice et al. Each function is for a different source location and describes the gain of sound energy that propagates to the eardrum relative to the free-field amplitude spectrum of the stimulus. The data were recorded

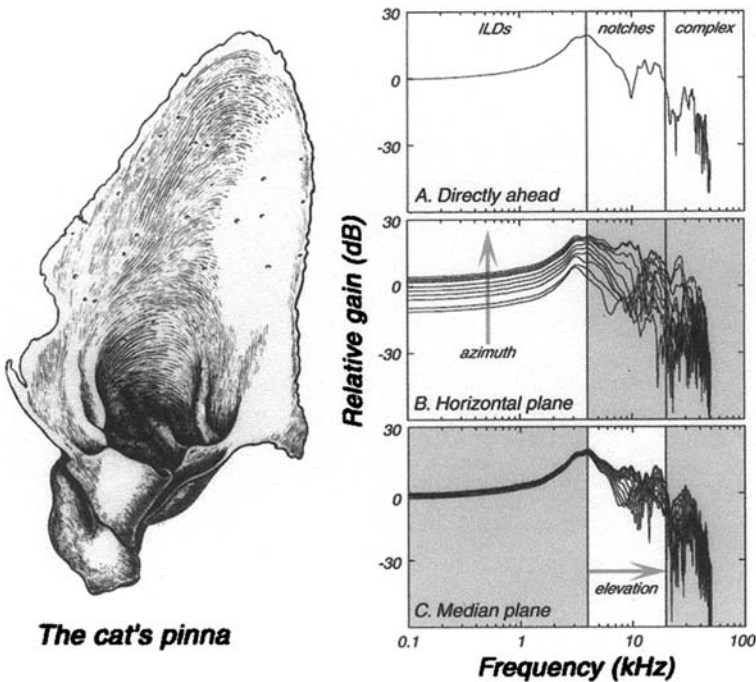


FIGURE 5.15. HRTFs of the cat. (A) Filtering effects of the pinna fall into three frequency domains. This example was measured with the sound directly in front of the subject ( $0^\circ$  AZ,  $0^\circ$  EL). (B) Low frequencies convey ILDs as sounds move in the horizontal plane around the head. (C) Mid-frequencies exhibit a prominent notch that varies in frequency with changes in sound source elevation. High-frequency spectral cues are complex and show a less orderly relationship to the direction of a sound source. (From Rice et al. 1992. Reprinted with permission.)

by surgically implanting a probe microphone in the ear canal of an anesthetized cat. The transfer function in Figure 5.15A was measured with the sound source directly in front of the subject's head ( $0^\circ$  AZ,  $0^\circ$  EL). As proposed by Rice and colleagues, three major directional properties of the HRTF are evident in this example. At low frequencies ( $<5$  kHz), the function displays a broad amplification that rises to an energy peak around 4 to 6 kHz. At mid frequencies (5 to 20 kHz), the HRTF exhibits a single prominent energy minimum, or spectral notch. At high frequencies ( $>20$  kHz), a complex pattern of peaks and notches is observed as the overall gain of the transfer function falls to low sound pressure levels.

Figure 5.15B summarizes how the HRTF changes with the azimuth of a sound source in the horizontal plane. The transfer functions that are superimposed in this example were obtained by placing the sound source at 11 locations in the frontal sound field ( $\pm 75^\circ$  in increments of  $15^\circ$ ). In comparison to the HRTF in Figure 5.15A, the more lateralized HRTFs display a low-frequency gain that is either attenuated by the "sound shadow" of the head when the source is in the far field opposite the microphone, or amplified by the pinna when the source is in the near field. This directionally dependent change in gain is the acoustic basis of the ILD (Shaw 1974a,b; Kuhn 1987; Martin and Webster 1989). By contrast, Figure 5.15C shows how the low-frequency filtering properties of the HRTF are virtually unaffected as the sound source passes through 17 locations in the median plane ( $-30^\circ$  to  $90^\circ$  in increments of  $7.5^\circ$ ). These findings suggest that in the cat binaural processes related to the ILD provide a poor representation of elevation.

The spectral notch at the mid-frequencies of the HRTF changes in frequency as the sound source moves in the horizontal or median plane. This potential localization cue supplements ILD information at lateralized spatial locations in Figure 5.15B, but appears to have singular importance for signaling elevation changes in Figure 5.15C. Neurophysiological studies of the central auditory system have identified neural populations in the DCN and inferior colliculus (IC) that are selectively responsive to HRTF-based spectral notches (Young et al. 1992; Imig et al. 2000). The highly nonlinear spectral integration properties of these neurons may represent a pathway for encoding sound localization cues in the spectral domain (Nelken et al. 1997; Spirou et al. 1999), just as the binaural pathways of the MSO and LSO are specialized for processing interaural time and level differences.

The directional cues conveyed by the high-frequency filtering properties of the HRTF are complex and highly variable. Current analyses have not revealed global directional relationships in the spectral characteristics of high-frequency peaks and notches, but elevated frontal locations do show more high-pass filtering effects. An interesting perceptual phenomenon that may arise from this property of the HRTF is the observation that high-frequency tones or noise bands are often heard as elevated sources regardless of their actual location (Pratt 1930; Butler and Belendiuk 1977; Blauert 1997). These systematic errors in narrow-band localization can be explained by matching the proximal stimulus spectrum

in the subject's ear to the directional properties of the HRTF (Middlebrooks 1992; van Schaik et al. 1999). For example, a narrow band of noise with a center frequency of 12 kHz will be attributed to locations where an individual's HRTF selectively passes those frequency components. Confusions of this nature are most apparent in vertical localization because the perception of horizontal location is enhanced by binaural directional information, like the ILD cues described in Figure 5.15B.

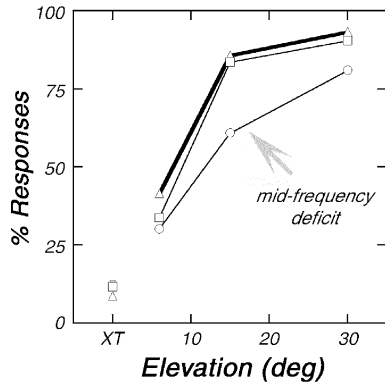
Three-dimensional virtual sound locations can be reproduced in a closed-field by adding false HRTF-filtering effects to the source spectrum (Wightman and Kistler 1989b; Pralong 1996; Langendijk and Bronkhorst 2000). Given this interesting perceptual effect, how can the auditory system derive HRTF effects without knowing the characteristics of the original source spectrum? This signal processing problem is avoided under normal listening conditions because the spectral shapes of natural sounds tend to have locally constant slopes that are capable of revealing the sharp peaks and notches of the HRTF (Zakarauskas and Cynader 1993). It is also true that the listener gets simultaneous "looks" at the spectrum from the different directional perspectives of the two ears. If the sound is sustained, the HRTF will change with movements of the source or the listener's head to reveal the underlying spectrum (Wightman and Kistler 1999). Animals with mobile pinna, like cats, can also translate the HRTF by moving the ears independently of the head (Young et al. 1996; Populin and Yin 1998b), a behavior that adds the dimension of proprioceptive feedback to spectral processing networks in the auditory brainstem (Kanold and Young 2001). Nevertheless, optimal localization of sound source elevation is observed for familiar sounds (McGregor et al. 1985; Blauert 1997) and the filtering effects of the listener's own ears (Wenzel et al. 1993; Hofman et al. 1998; Middlebrooks 1999).

## *5.2 Spectral Cues for the Discrimination of Changes in Sound Source Direction*

The most common procedure for characterizing directional hearing in nonhuman animals is the MAA task in which the subject indicates the detection of a change in location by responding on a lever (May et al. 1995; Huang and May 1996b) or suppressing an ongoing behavior to avoid electrical shocks (Martin and Webster 1987; Heffner and Heffner 1988d). These methods have the advantage of relatively short training periods and produce psychometric data that allow easy quantification of directional acuity (Mills 1958).

Figure 5.16 shows average psychometric functions that were obtained by testing three cats with the MAA task (Huang and May 1996b). The cats were required to hold down on a response lever when bursts of noise were presented from a reference speaker ( $0^\circ$  AZ,  $0^\circ$  EL), and to release the lever when the sound source shifted to another speaker. The comparison speakers were arranged in the median plane (Fig. 5.16A) or the horizontal plane (Fig. 5.16B). The percentage of correct lever releases is plotted in relation to the magnitude of the

### A. Median plane



### B. Horizontal plane

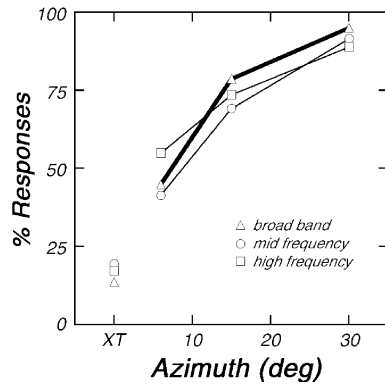


FIGURE 5.16. Effects of stimulus frequency on directional acuity in the median plane (A) and horizontal plane (B). Psychometric functions show the percentage of correct responses for directional changes relative to a reference speaker at 0° AZ, 0° EL. The percentage of incorrect responses for catch trials (XTs) are plotted as individual symbols to the left of the psychometric functions. The individual functions reflect the average responses of three cats to broadband (> 5 kHz), mid frequency (5 to 20 kHz), or high-pass noise (> 20 kHz). (Adapted from Huang and May 1996b.)

directional change between the reference and comparison speaker. Responses to catch trials (*XTs*) are indicated to the left of the psychometric functions. Catch trials were conducted just like MAA tests, but no speaker change was presented. The subject's responses to catch trials are presumed to reflect the probability of false-positive responses that result from guessing. The MAA is defined as the change in location (elevation or azimuth) that corresponds with the signal detection criterion of  $d' = 1$  (based on the probabilities of correct responses to MAA trials and error responses to catch trials).

The psychometric functions in Figure 5.16 reveal the effects of frequency on directional acuity. Three frequency ranges were selected to evaluate the perceptual significance of the HRTF directional filtering effects that are summarized in Figure 5.15. Broadband noise contained spectral cues above 5 kHz. Mid-frequency noise was limited to the systematic spectral notch region from 5 to 20 kHz. High-frequency noise conveyed only the complex spectral cues above 20 kHz.

The behavioral results in Figure 5.16 indicate that best directional acuity was achieved with broadband noise, and this agrees with the results of Martin and Webster (1987). Since this stimulus condition included both mid-frequency and high-frequency spectral cues, the relative contribution of the individual spectral domains can be assessed by comparing response accuracy under broadband and the band-limited testing conditions. In general, no difference was observed in directional acuity between broadband and high-frequency noise. The subjects failed to detect changes in location more often when tests were conducted with mid-frequency noise. This deficit was most evident for sound sources in the median plane.

Computational models based on the auditory nerve encoding of spectral cues for sound localization offer an interpretation of the behavioral results in Figure 5.16 (May and Huang 1997). These models demonstrate a sensitive representation of directional change among neurons that respond best to the high-frequency components of HRTF-shaped noise. Even small changes in sound source location are capable of producing large and pervasive changes in the high-frequency HRTF, as shown in Figure 5.15. These spectral variations are reflected in the discharge rates of auditory neurons. The neural response is not inherently directional, but it is a sufficient cue for accurate performance of the MAA task where the subject is only required to respond to acoustic differences that are correlated with directional changes. Singular mid-frequency notches provide less effective information because they are more localized in frequency and smaller in magnitude than the multiple high-frequency notches. This mid-frequency deficit is less apparent for sound sources in the horizontal plane because spectral cues are augmented with binaural directional information.

### *5.3 Spectral Cues for the Perception of an Absolute Directional Identity*

An alternative behavioral method for measuring sound localization accuracy involves training the subject to point toward the direction of a sound or approach the source. This procedure is necessary for studies in which the perceived location of the sound is an important parameter of the experiment. For example, the investigator may be interested in the systematic errors that are induced by modifying the source spectrum of the localization stimulus. After such manipulations, the subject might grossly misinterpret the actual location of the modified stimulus but still respond correctly to a change from one speaker location to another in the MAA task.

Sound-directed orientation behaviors of the head (Thompson and Masterton 1978) or eyes (Populin and Yin 1998a) have been used to characterize the perception of directional identity in cats. Unlike an approach procedure, which is constrained by source locations arrayed along the floor of the testing arena (Casseday and Neff 1973), head-orientation tasks can measure the simultaneous localization of stimulus azimuth and elevation. Orientation is a natural reflexive behavior that can be used as a short-term response metric for unexpected sounds

in naïve subjects (Sutherland et al. 1998), or it can be shaped into a food-reinforced operant paradigm that is capable of sustaining long-term psychophysical analyses of the acoustic cues for directional hearing (May and Huang 1996).

Performance in a food-reinforced orientation task is summarized by the behavioral results shown in Figure 5.17 (Huang and May 1996a). This cat earned food rewards by accurately orienting its head toward randomly selected sound sources in an anechoic chamber. An electromagnetic sensor that was worn during the testing session tracked head movements. Trials in the orientation task were conducted with discrete presentations of brief noise bursts (40 ms), so the subject could not influence localization accuracy by moving its head or ears

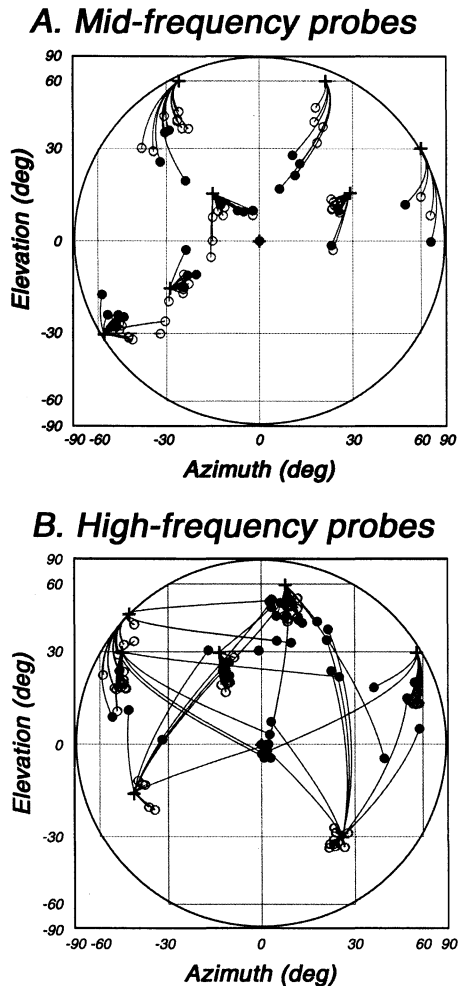


FIGURE 5.17. Effects of stimulus frequency on sound orientation behavior. Broadband stimuli were presented on 85% of the tests in these sessions to assess baseline accuracy (*open symbols*). The remaining probe trials evaluated the reliability of directional cues that were conveyed by restricted frequency regions of the head-related transfer function (*filled symbols*). (A) Orientation responses for 5 to 20 kHz band-pass noise versus broadband noise. (B) Orientation responses for 20 kHz high-pass noise versus broadband noise. Results for each condition are based on performance during one session. Additional plotting conventions are described in Figure 5.6. (Adapted from Huang and May 1966b.)

after stimulus onset. Most of the trials were conducted with bursts of broadband noise containing the full complement of HRTF-based localization cues. Orientation responses to these spectrally rich stimuli (open symbols in gray) were in good agreement with actual sound source locations (plus symbols).

The effect of frequency on the perception of sound source direction was evaluated by probing the cat's orientation behavior with bursts of mid-frequency noise (Fig. 5.17A) and high-frequency noise (Fig. 5.17B). Probe stimuli were intermingled at random intervals among the more frequent broadband noise bursts. Only slight changes in orientation accuracy were noted when the probes maintained mid-frequency cues. The reliability of mid-frequency spectral information in the perception of sound source elevation is attributed to auditory processing of directional notches in the cat's HRTF (Fig. 5.15C) (May and Huang 1997). High-frequency probes resulted in head movements that were inaccurate and highly variable, particularly with respect to the vertical coordinate of the source. These results suggest that cats do not utilize the complex filtering effects of the high-frequency HRTF for directional hearing even though the spectral cues in this high-frequency region should provide an excellent source of information for the perception of source direction (Fig. 5.16).

#### *5.4 Spectral Processing Pathways in the Central Auditory System*

Just as there are pathways in the central nervous system to enhance binaural directional hearing, behavioral and electrophysiological studies are beginning to reveal neural specializations for the auditory processing of spectral cues for sound localization. In the DCN, ascending inputs from the auditory nerve combine with a complex local inhibitory circuitry and descending projections from throughout the brain to create a notch-sensitive projection neuron that is also capable of integrating information about the orientation of the moveable pinna (Young et al. 1992; Imig et al. 2000). The target neurons for these projections in the central nucleus of the inferior colliculus (ICC) show spatially selective receptive fields that are sensitive to HRTF-filtering effects (Ramachandran et al. 1999; Davis 2002; Davis et al. 2003).

The functional significance of the putative spectral processing pathway has been explored with behavioral procedures by evaluating the auditory deficits that follow surgical lesions of the dorsal acoustic strial fibers that link the DCN to ICC (Sutherland et al. 1998; May 2000). As shown in Figure 5.18, these fibers exit the DCN and combine with the intermediate acoustic strial fibers from the ventral cochlear nucleus (VCN). The lesion was made by transecting the striae at the most dorsal limit of the nucleus. Previous studies have shown that the surgical procedure has little effect on hearing sensitivity because this more generalized auditory information ascends from the VCN to the binaural brainstem nuclei and inferior colliculus by way of the trapezoid body (Masterton et al. 1994).

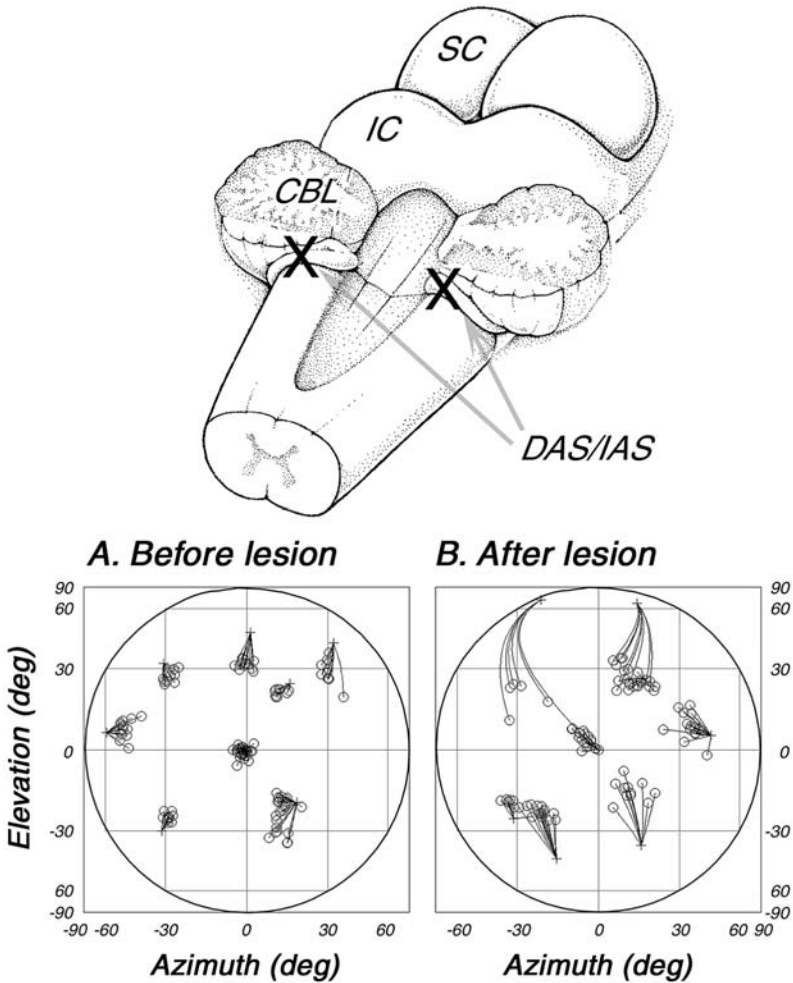


FIGURE 5.18. Effects of dorsal cochlear nucleus lesions on sound orientation behavior. The anatomical drawing illustrates the placement of surgical lesions (X). Orientation responses observed during one session before the lesion (A) and during another session with the same cat after the output pathways of the DCN were transected (B). Plotting conventions are described in Figure 5.6. CBL, cerebellum; DAS/IAS, dorsal and intermediate acoustic striae; IC, inferior colliculus; SC, superior colliculus. (Adapted from May 2000.)

The orientation plots in Figure 5.18 compare the head pointing behaviors of one cat before and after a bilateral lesion of DCN projections (May 2000). These tests were conducted with bandpass noise to restrict the domain of HRTF-based spectral information to the mid-frequency notches that exist at 5 to 20 kHz. As predicted by the results of the probe testing procedure in Figure 5.17A,

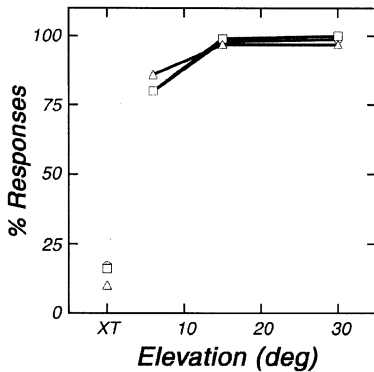


the cat exhibited excellent orientation accuracy prior to the lesion (Fig. 5.18A). Large errors were noted after the lesion disrupted the spectral processing pathways of the DCN and ICC (Fig. 5.18B).

An analysis of the patterns of errors in the lesioned cat indicates that the localization deficits were statistically significant only in terms of response elevation. Regardless of the actual location of the sound source, the subject's orientation responses were seldom directed at elevations beyond  $\pm 30^\circ$ . These systematic underestimations could exceed  $60^\circ$  for extreme source locations. It is likely that the subject maintained accurate azimuthal localization after the DCN lesion by relying on nonspectral directional information; for example, ILD and ITD cues that were processed in the intact binaural pathways of the auditory brainstem nuclei.

Cats with DCN lesions also have been studied with MAA procedures to confirm the specificity of the orientation deficits in Figure 18 (May 2000). These experiments are summarized by the psychometric functions in Figure 5.19.

### A. Median plane



### B. Horizontal plane

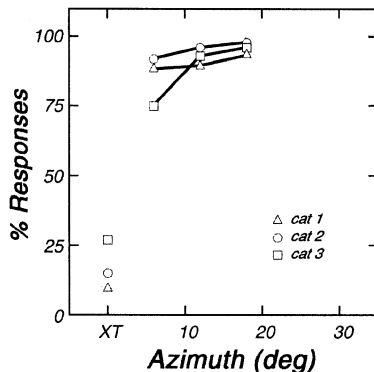


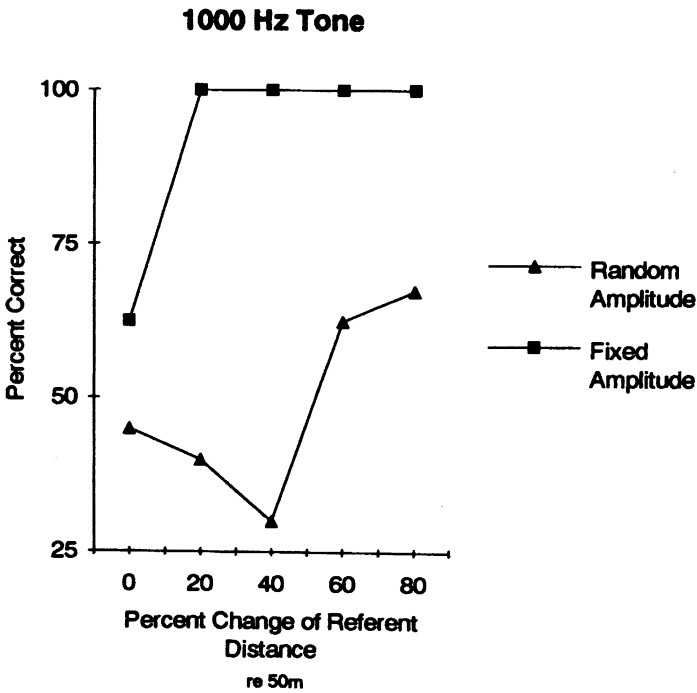
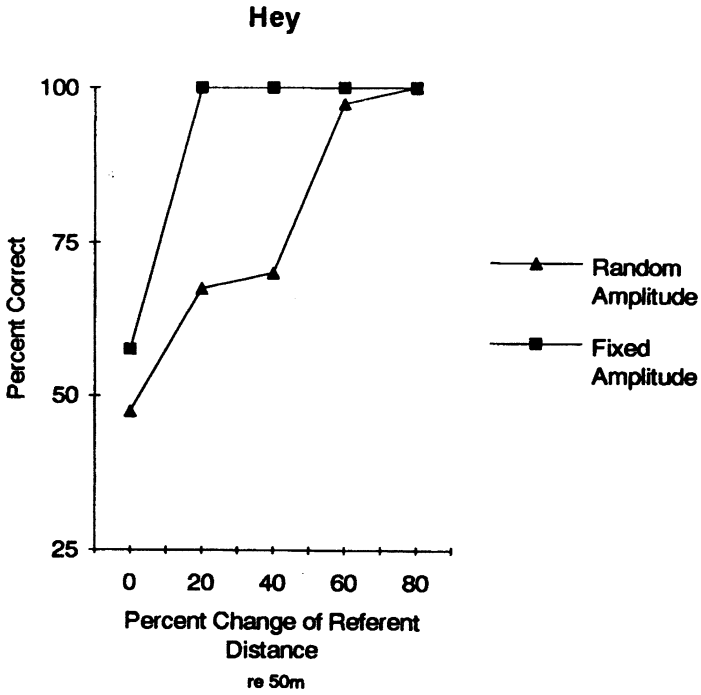
FIGURE 5.19. Effects of DCN lesions on spatial acuity. Psychometric functions summarize the detection scores of three cats for directional changes in the median plane (A) and horizontal plane (B). Responses to catch trials (XTs) are indicated by the symbols to the left of the functions. The physical dimensions of the speakers prevented testing at angular separations that were less than  $6^\circ$ . (Adapted from May 2000.)

Each function represents the post-lesion performance of one cat for directional changes in the median plane (A) and horizontal plane (B). The test stimuli were mid-frequency noise bursts. Although all of the subjects exhibited orientation deficits after the DCN lesion, none of the subjects showed signs of impaired spatial discrimination. These results confirm that DCN lesions do not lead to general hearing deficits and further support the idea that spatial acuity and directional identification are based on different auditory cues and processing pathways.

## 6. The Perception of Acoustic Proximity

The perception of acoustic distance, or acoustic proximity, has received very little formal study. Brown (1994) measured the minimal perceptible change in acoustic distance for human listeners in a forest habitat at a reference distance of 50 m. Using the speech utterance “hey” and a 1-kHz tone for the stimuli, it was determined that subjects would use changes in loudness, or sound amplitude, if the magnitude of the stimulus at its source was held constant as distance was varied. However, if signal amplitude was adjusted to compensate for changes in distance (and if random amplitude fluctuations were introduced), subjects were able to perceive changes in acoustic proximity only for the spectrally complex speech stimulus. This fact indicates that human listeners used changes in sound quality as described in Section 2.6 to detect changes in acoustic distance. Figure 5.20 shows that human listeners could perceive a 10% change in acoustic distance when the source level was fixed for both the tone and speech stimulus. This finding shows that loudness, or auditory image constancy, is an important cue for the perception of changes in acoustic proximity when it is available for processing (the amplitude of the signal is fixed). The detection of a 10% change in acoustic distance in a forested site compares closely with distance-discrimination thresholds of about 6% for reference distances of 6 to 49 m on an open athletic field (Strybel and Perrott 1984), and with distance-discrimination thresholds of about 6% for reference distances of 1 to 2 m in an anechoic room (Ashmead et al. 1990). The scattering of sound in the forested habitat will change the rate of sound attenuation with respect to distance relative to that in open environments (Waser and Brown 1986). Sound propagation is complicated because the elevation of the source and receiver, and the frequency of the signal have strong effects. Nevertheless, signals in the speech range, at the elevation of the human head, tend to be propagated better in forested than in open habitats (Waser and Brown 1986; Brown et al. 1995). That is, in forested compared to open habitats, a greater change in propagation distance will be required to produce a unit change in the level of the signal, and these acoustic influences likely account for the difference in the thresholds reported in open field and anechoic environments compared to that observed in forested environments.

Under most natural situations, sound amplitude is not the only available cue



for the perception of a change in acoustic proximity. When the amplitude of the signal is adjusted to compensate for changes in transmission distance, and when the amplitude of the signal varies randomly trial-to-trial, loudness, or auditory image constancy, is no longer a viable cue. Nevertheless, human subjects are still able to perceive changes in acoustic proximity when tested with a complex speech stimulus. The data in Figure 5.20 show that subjects could perceive a change of 44% of the reference distance under these conditions. In an anechoic room, Ashmead et al. (1990) reported that human listeners could detect changes in distance of about 16% at reference distances of 1 to 2 m when the amplitude of the test and reference stimuli were equated. It is likely that spectral changes and reverberation were the most prominent cues underlying the perception of changes of distance.

The ability to perceive changes in sound quality associated with changes in acoustic distance has been measured in blue monkeys (*Cercopithecus mitis*). Figure 5.21 shows that blue monkeys can detect a change in proximity of 54% for the pyow vocalization broadcast in their natural habitat. This finding suggests that reflection of the wave front by tree trunks and other surfaces and frequency-specific attenuation may change or distort acoustic signals in a manner that provides a reliable and perceptually useful index of acoustic distance. It is conceivable that organisms residing in various habitats may have developed signals that are particularly well suited to permit listeners to ascertain the distance to the vocalizer. Furthermore, it is possible that some calls possess an acoustic structure that makes it possible to detect small changes in the proximity of the vocalizer, while other calls may tend to obscure the available distance cues.

## 7. Conclusion

Mammals have a sense of the azimuth, elevation, and distance of the source of acoustic events. However, the resolution of sound position is not equal in all three coordinates. The available data suggests that for most mammals the acuity of resolution of sound source azimuth is greater than that for elevation, and the acuity of resolution for sound source elevation is greater than that for distance. Hence, the minimal audible change in acoustic locus for azimuth, elevation, and distance may be described by the surface of an ellipsoid, a three-dimensional figure oriented such that the width is less than the height, which in turn, is less



FIGURE 5.20. The minimum perceptible change in proximity for human listeners. The test signals were the word “hey” (*top*) and a 1-kHz tone (*bottom*). The reference distance was 50 m. Testing was conducted in a forested habitat. The *triangles* indicate detection when the intensity of the signal is held constant; the *squares* indicate detection when the level of the signal is randomized and adjusted to compensate for changes in loudness with distance (From Brown 1994. Reprinted with permission.)

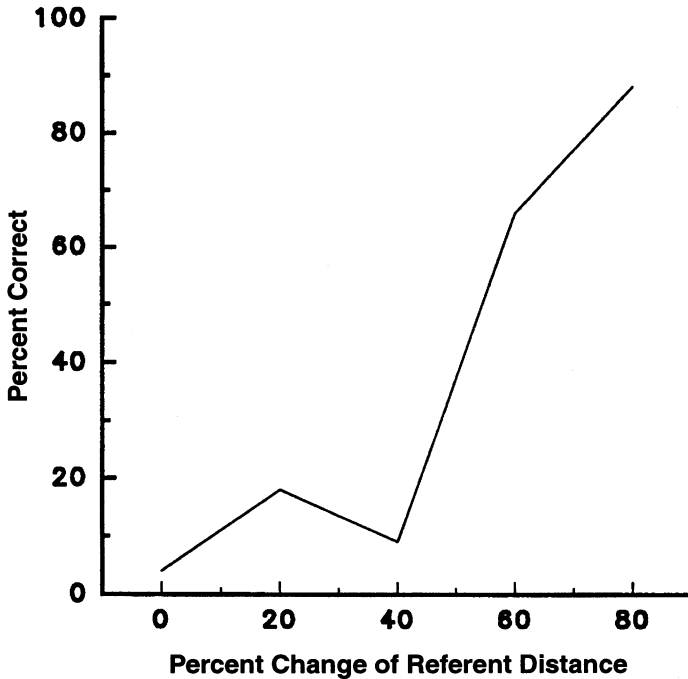


FIGURE 5.21. The minimum perceptible change in proximity in blue monkeys (*Cercopithecus mitis*). The test signal was the pyow vocalization. The reference distance was 50 m. The signal was broadcast and rerecorded at transmission distances of 5, 10, 20, 30, 40, and 50 m. Broadcasts were conducted at an elevation of 7 m in Kibale forest in western Uganda. The amplitude of the test signal was randomized between trials, and adjusted to compensate for changes in loudness with distance. (From Brown 1994. Reprinted with permission.)

than the length. A theoretical three-dimensional minimal perceptible change in locus ellipsoid is illustrated in Figure 5.22.

All three coordinates of sound source localization are important biologically. However, because the cues that underlie the perception of azimuth, elevation, and distance are so dissimilar, it is possible that subjects may experience abnormalities or disorders that impair perception in one dimension, yet leave relatively intact perception in the other two dimensions. Furthermore, it is possible that the ecology and life history of different species have led to enhanced sensitivity for localization in one coordinate relative to that in another. Terrestrial species may have been selected to maximize acuity for source azimuth, while marine organisms and arboreal species may have been selected for enhanced acuity for source elevation, and forest-living species may have been selected for greater acuity for source distance. Researchers have generated a wealth of studies of the comparative perception of sound source azimuth, and have only begun

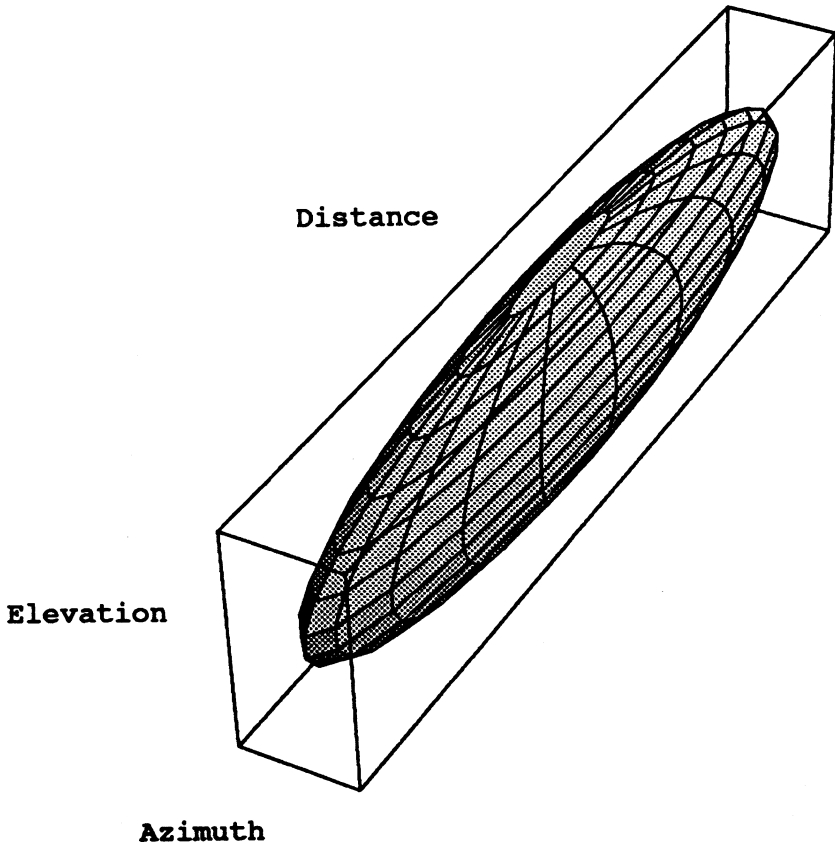


FIGURE 5.22. A theoretical volume describing the minimum perceptible change in locus of a broad-bandwidth sound. The reference locus is the center of the ellipsoid and the just perceptible change in locus in any direction is given by the distance from the center to any point on the surface of the volume. In the ellipsoid drawn here the resolution for changes in azimuth are two times that for changes in elevation, and eight times that for changes in distance. The actual dimensions of the volume describing the minimally perceptible change in space would be influenced by the acoustics of the habitat (test environment) and the temporal and spectral complexity of the test signal. (From Brown 1994. Reprinted with permission.)

to study the localization of source elevation or distance. The methodology for good comparative studies of auditory perception are well established, and we encourage researchers to focus greater attention on the elevation and distance coordinates of sound source position.

## References

- Anderson DJ (1973) Quantitative model for the effects of stimulus frequency upon the synchronization of auditory nerve discharges. *J Acoust Soc Am* 54:361–364.
- Aronson E, Rosenbloom S (1971) Space perception in early infancy: perception within a common auditory-visual space. *Science* 172:1161–1163.
- Ashmead DH, LeRoy D, Odom RD (1990) Perception of the relative distances of nearby sound sources. *Percept Psychophys* 47:326–331.
- Ayrapet'yants ES, Konstantinov AI (1974) Echolocation in nature. An English translation of the National Technical Information Service, JPRS 63326-1-2.
- Batra R, Kuwada S, Fitzpatrick DC (1997) Sensitivity of temporal disparities of low- and high-frequency neurons in the superior olivary complex. I. Heterogeneity of responses. *J Neurophysiol* 78:1222–1236.
- Batteau DW, Plante RL, Spencer RH, Lyle WE (1965) Localization of sound: Part 5. Auditory perception. U.S. Navy Ordnance Test Station Report, TP 3109, part 5.
- Beecher MD, Harrison JM (1971) Rapid acquisition of an auditor location discrimination in rats. *J Exp Anal Behav* 16:193–199.
- Beer CG (1969) Laughing gull chicks: recognition of their parent's voices *Science* 166:1030–1032.
- Beer CG (1970) Individual recognition of voice in social behavior of birds. In: Lehrman DS, Hinde RA, Shaw E (eds), *Advances in the Study of Behavior*, Vol. 3. New York: Academic Press, pp. 27–74.
- Blauert J (1997) *Spatial Hearing, The Psychophysics of Human Sound Localization*. Cambridge, MA: MIT Press.
- Brown CH (1982a) Ventriloquial and locatable vocalizations in birds. *Z Tierpsychol* 59:338–350.
- Brown CH (1982b) Auditor localization and primate vocal behavior. In: Snowdon CT, Brown CH, Petersen MR (eds), *Primate Communication*. Cambridge: Cambridge University Press, pp. 144–164.
- Brown CH (1989) The active space of blue monkey and grey-cheeked mangabey vocalizations. *Anim Behav* 37:1023–1034.
- Brown CH (1994) Sound localization. In: Fay RR, Popper AN (eds), *Comparative Hearing: Mammals*. New York: Springer-Verlag, pp. 57–96.
- Brown CH, Gomez R (1992) Functional design features in primate vocal signals: the acoustic habitat and sound distortion. In: Nishida T, McGrew WC, Marler P, Pickford M, de Waal FMB (eds), *Topics in Primatology*, Vol. 1: Human Origins. Tokyo: University of Tokyo Press, pp. 177–198.
- Brown CH, May BJ (1990) Sound Localization and binaural processes. In: Berkley MA, Stebbins WC (eds), *Comparative Perception*, Vol. I: Basic Mechanisms. New York: John Wiley & Sons, pp. 247–284.
- Brown CH, Waser PM (1988) Environmental influences on the structure of primate vocalizations. In: Todt D, Geodeking P, Symme D (eds), *Primate Vocal Communication*. Berlin: Springer-Verlag, pp. 51–66.
- Brown CH, Beecher MD, Moody DB, Stebbins WC (1978a) Localization of pure tones in Old World monkeys. *J Acoust Soc Am* 63:1484–1492.
- Brown CH, Beecher MD, Moody DB, Stebbins WC (1978b) Localization of primate calls by Old World monkeys. *Science* 201:753–754.
- Brown CH, Beecher MD, Moody DB, Stebbins WC (1979) Locatability of vocal signals

- in Old World monkeys: design features for the communication of position. *J Comp Physiol Psychol* 93:806–819.
- Brown CH, Beecher MD, Moody DB, Stebbins WC (1980) Localization of noise bands by Old World monkeys. *J Acoust Soc Am* 68:12–132.
- Brown CH, Schessler T, Moody DB, Stebbins WC (1982) Vertical and horizontal sound localization in primates. *J Acoust Soc Am* 72:1804–1811.
- Brown CH, Gomez R, Waser PM (1995) Old World monkey vocalizations: adaptation to the local habitat? *Anim Behav* 50:945–961.
- Butler RA (1969) Monaural and binaural localization of noise bursts vertically in the medial sagittal plane. *J Audit Res* 9:230–235.
- Butler RA, Belendiuk K (1977) Spectral cues utilized in the localization of sound in the median sagittal plane. *J Acoust Soc Am* 61:1264–1269.
- Cain WS, Marks LE (1971) Stimulus and Sensation. Boston: Little, Brown.
- Carlile S, Pralong D (1994) The location-dependent nature of perceptually salient features of the human head-related transfer functions. *J Acoust Soc Am* 95:3445–3459.
- Casseday JH, Neff WD (1973) Localization of pure tones. *J Acoust Soc Am* 54:365–372.
- Casseday JH, Neff WD (1975) Auditory localization: role of auditory pathways in brain stem of the cat. *J Neurophysiol* 38:842–858.
- Chambers RE (1971) Sound localization in the hedgehog (*Paraechinus hypomelas*). Unpublished master's thesis, Florida State University, Tallahassee.
- Clements M, Kelly JB (1978a) Directional responses by kittens to an auditory stimulus. *Dev Psychobiol* 11:505–511.
- Clements M, Kelly JB (1978b) Auditor spatial responses of young guinea pigs (*Cavi porcellus*) during and after ear blocking. *J Comp Physiol Psychol* 92:34–44.
- Coleman PD (1963) An analysis of cues to auditory depth perception in free space. *Psychol Bull* 60:302–315.
- Davis KA (2002) Evidence of a functionally segregated pathway from dorsal cochlear nucleus to inferior colliculus. *J Neurophysiol* 87:1824–1835.
- Davis K, Ramachandran R, May BJ (2003) Auditory processing of spectral cues for sound localization in the inferior colliculus. *JARO* 4:148–163.
- Delgutte B, Joris PX, Litovsky RY, Yin TC (1999) Receptive fields and binaural interactions for virtual-space stimuli in the cat inferior colliculus. *J Neurophysiol* 81:2833–2851.
- Denzinger A, Schnitzler, HU (1998) Echo SPL, training experience, and experimental procedure influence the ranging performance in the big brown bat, *Eptesicus fuscus*. *J Comp Physiol* 183:213–224.
- Devor A (2000) Is the cerebellum like cerebellar-like structures? *Brain Res Brain Res Rev* 34:149–156.
- Fay RR (1988) Hearing in Vertebrates: A Psychophysics Databook. Winnetka, IL: Hill-Fay Associates.
- Feng AS, Gerhardt HC, Capranica RR (1976) Sound localization behaviour of the green tree frog (*Hyla cinerea*) and the barking tree frog (*H. gratiiosa*). *J Comp Physiol* 107:241–252.
- Firestone FA (1930) The phase differences and amplitude ratio at the ears due to a source of pure tones. *J Acoust Soc Am* 2:260–270.
- Gamble EA (1909) Intensity as a criterion in estimating the distance of sounds. *Psychol Rev* 16:416–426.



- Gardner MB (1973) Some monaural and binaural facets of median plane localization. *J Acoust Soc Am* 54:1489–1495.
- Gottlieb G (1965) Imprinting in relation to parental and specie identification by avian neonates. *J Comp Physiol Psychol* 59:345–356.
- Hall JW, Haggard MP, Fernandes MA (1984) Detection in noise by spectro-temporal pattern analysis. *J Acoust Soc Am* 76:50–56.
- Harris JD (1972) A florilegium of experiments on directional hearing. *Acta Oto-Laryngol Suppl* 298:1–26.
- Harrison JM (1990) Simultaneous auditory discriminations. *J Exp Anal Behav* 54:45–51.
- Harrison JM (1992) Avoiding conflicts between the natural behavior of the animal and the demands of discrimination experiments. *J Acoust Soc Am* 92:1331–1345.
- Harrison JM, Downey P (1970) Intensity changes at the ear as a function of azimuth of a tone: a comparative study. *J Acoust Soc Am* 56:1509–1518.
- Harrison JM, Irving R (1966) Visual and nonvisual auditory systems in mammals. *Science* 154:738–743.
- Harrison JM, Downey P, Segal M, Howe M (1971) Control of responding by location of auditory stimuli: rapid acquisition in monkey and rat. *J Exp Anal Behav* 15:379–386.
- Hartley RVL, Fry TC (1921) The binaural localization of pure tones. *Physiol Rev* 18:431–442.
- Hebrank J, Wright D (1974) Spectral cues used in the localization of sound sources in the median plane. *J Acoust Soc Am* 56:1829–1834.
- Heffner HE, Heffner RS (1984) Sound localization in large mammals: localization of complex sounds by horses. *Behav Neurosci* 98:541–555.
- Heffner HE, Heffner RS (1985) Sound localization in wild Norway rats (*Rattus norvegicus*). *Hear Res* 19:151–155.
- Heffner RS, Heffner HE (1982) Hearing in the elephant (*Elephas maximus*): absolute sensitivity, frequency discrimination, and sound localization. *J Comp Physiol Psychol* 96:926–944.
- Heffner RS, Heffner HE (1983) Hearing in large mammals: horses (*Equus caballus*) and cattle (*Bos taurus*). *Behav Neurosci* 97:299–309.
- Heffner RS, Heffner HE (1985) Auditory localization and visual fields in mammals. *Neurosci Abstr* 11:547.
- Heffner RS, Heffner HE (1987) Localization of noise, use of binaural cues, and a description of the superior olivary complex in the smallest carnivore, the least weasel (*Mustel nivalis*). *Behav Neurosci* 101:701–708.
- Heffner RS, Heffner HE (1988a) Sound localization in a predatory rodent, the northern grasshopper mouse (*Onychomys leucogaster*). *J Comp Psychol* 102:66–71.
- Heffner RS, Heffner HE (1988b) Sound localization and the use of binaural cues in the gerbil (*Meriones unguiculatus*). *Behav Neurosci* 102:422–428.
- Heffner RS, Heffner HE (1988c) The relation between vision and sound localization acuity in mammals. *Soc Neurosci Abstr* 14:1096.
- Heffner RS, Heffner HE (1988d) Sound localization acuity in the cat: effect of signal azimuth, signal duration, and test procedure. *Hear Res* 36:221–232.
- Heffner RS, Heffner HE (1989) Sound localization, use of binaural cues, and the superior olivary complex in pigs. *Brain Behav Evol* 33:248–258.
- Heffner RS, Heffner HE (1992) Visual factors in sound localization in mammals. *J Comp Neurol* 317:219–232.

- Heffner HE, Masterton RB (1980) Hearing in glires: domestic rabbit, cotton rat, feral house mouse, and kangaroo rat. *J Acoust Soc Am* 68:1584–1599.
- Heffner RS, Masterton RB (1990) Sound localization: brainstem mechanisms. In: Berkley MA, Stebbins WC (eds), *Comparative Perception, Vol. I: Basic Mechanisms*. New York: John Wiley & Sons, pp. 285–314.
- Heffner RS, Richard MM, Heffner HE (1987) Hearing and the auditory brainstem in a fossorial mammal, the pocket gopher. *Neurosci Abstr* 13:546.
- Heffner RS, Heffner HE, Koay G (1995) Sound localization in chinchillas. II. Front/back and vertical localization. *Hear Res* 88:190–198.
- Henning GB (1974) Detectability of interaural delay in high-frequency complex waveforms. *J Acoust Soc Am* 55:84–90.
- Hofman PM, Van Riswick JG, Van Opstal AJ (1998) Relearning sound localization with new ears. *Nat Neurosci* 1:417–421.
- Hornbostel EM (1923) Beobachtungen über ein- und zweiohrigs Hören. *Psychol Forsch* 4:64–114.
- Houben D, Gourevitch G (1979) Auditory lateralization in monkeys: an examination of two cues serving directional hearing. *J Acoust Soc Am* 66:1057–1063.
- Huang AY, May BJ (1996a) Sound orientation behavior in cats. II. Mid-frequency spectral cues for sound localization. *J Acoust Soc Am* 100:1070–1080.
- Huang AY, May BJ (1996b) Spectral cues for sound localization in cats: effects of frequency domain on minimum audible angles in the median and horizontal planes. *J Acoust Soc Am* 100:2341–2348.
- Imig TJ, Poirier P, Irons WA, Samson FK (1997) Monaural spectral contrast mechanism for neural sensitivity to sound direction in the medial geniculate body of the cat. *J Neurophysiol* 78:2754–2771.
- Imig TJ, Bibikov NG, Poirier P, Samson FK (2000) Directionality derived from pinna-cue spectral notches in cat dorsal cochlear nucleus. *J Neurophysiol* 83:907–925.
- Johnson DH (1980) The relationship between spike rate and synchrony in responses of auditory nerve fibers to single tones. *J Acoust Soc Am* 68:1115–1122.
- Kanold PO, Young ED (2001) Proprioceptive information from the pinna provides somatosensory input to cat dorsal cochlear nucleus. *J Neurosci* 21:7848–7858.
- Kaufman L (1979) *Perception: The World Transformed*. New York: Oxford University Press.
- Kelly JB (1980) Effects of auditory cortical lesions on sound localization in the rat. *J Neurophysiol* 44:1161–1174.
- Kiang NYS, Watanabe T, Thomas EC, Clark LF (1965) *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve*. Cambridge, MA: MIT Press.
- King AJ, Parsons CH (1999) Improved auditory spatial acuity in visually deprived ferrets. *Eur J Neurosci* 11:3945–3956.
- King AJ, Hutchings ME, Moore DR, Blakemore C (1988) Developmental plasticity in the visual and auditory representations in the mammalian superior colliculus. *Nature* 332:73–76.
- Klump RG, Eady HR (1956) Some measurements of interaural time difference thresholds. *J Acoust Soc Am* 28:859–860.
- Knudsen EI (1983) Early auditory experience aligns the auditory map of space in the optic tectum of the barn owl. *Science* 222:939–942.
- Knudsen EI, Brainard MS (1991) Visual instruction of the neural map of auditory space in the developing optic tectum. *Science* 253:85–87.

- Knudsen EI, Knudsen PF (1985) Vision guides the adjustment of auditory localization in young barn owls. *Science* 230:545–548.
- Knudsen EI, Knudsen PF (1990) Sensitive and critical periods for visual calibration of sound localization by barn owls. *J Neurosci* 10:222–232.
- Knudsen EI, Konishi M (1978) A neural map of auditory space in the owl. *Science* 200:795–797.
- Knudsen EI, Blasdel GG, Konishi M (1979) Sound localization by the barn owl measured with the search coil technique. *J Comp Physiol* 133:1–11.
- Knudsen EI, Knudsen, PF, Esterly SD (1984) A critical period for the recovery of sound localization accuracy following monaural occlusion in the barn owl. *J Neurosci* 4:1012–1020.
- Knudsen EI, Esterly SD, du Lac S (1991) Stretched and upside-down maps of auditory space in the optic tectum of blind-reared owls: acoustic basis and behavioral-coorelates. *J Neurosci* 11:1727–1747.
- Kuhn GF (1977) Model for the interaural time difference in the azimuthal plane. *J Acoust Soc Am* 62:157–167.
- Kuhn GF (1979) The effect of the human torso, head and pinna on the azimuthal directivity and on the median plane vertical directivity. *J Acoust Soc Am* 65:(S1), S8(A).
- Kuhn GF (1987) Physical acoustics and measurements pertaining to directional hearing. In: Yost W, Gourevitch G (eds), *Directional Hearing*. New York: Academic Press, pp. 3–25.
- Kujala T, Alho K, Paavilainen P, Summala H, Näätänen R (1992) Neural plasticity in processing of sound location by the early blind: an event-related potential study. *Electroencephalogr Clin Neurophysiol* 84:469–472.
- Kujala T, Alho K, Kekoni J, Hamalainen H, Reinikainen K, Salonen O, Standertskjold-Nordenstam CG, Näätänen R (1995) Auditory and somatosensory event-related brain potentials in early blind humans. *Exp Brain Res* 104:519–526.
- Kujala T, Alho K, Näätänen R (2000) Cross-modal reorganization of human cortical functions. *Trends Neurosci* 23:115–120.
- Kulkarni A, Colburn HS (1998) Role of spectral detail in sound-source localization. *Nature* 396:747–749.
- Langendijk EH, Bronkhorst AW (2000) Fidelity of three-dimensional-sound reproduction using a virtual auditory display. *J Acoust Soc Am* 107:528–537.
- Lessard N, Paré M, Lepore F, Lassonde M (1998) Early-blind humans subjects localize sound sources better than sighted subjects. *Nature* 395:278–280.
- Makous JC, Middlebrooks JC (1990) Two-dimensional sound localization by human listeners. *J Acoust Soc Am* 87:2188–2200.
- Martin RL, Webster WR (1987) The auditory spatial acuity of the domestic cat in the interaural horizontal and median vertical planes. *Hear Res* 30:239–252.
- Martin RL, Webster WR (1989) Interaural sound pressure level differences associated with sound-source locations in the frontal hemifield of the domestic cat. *Hear Res* 38:289–302.
- Masters WM, Raver KA (2000) Range discrimination by big brown bats (*Eptesicus fuscus*) using altered model echoes: implications for signal processing. *J Acoust Soc Am* 107:625–637.
- Masterton B, Heffner H, Ravizza R (1969) The evolution of human hearing. *J Acoust Soc Am* 45:966–985.
- Masterton RB, Granger EM, Glendenning KK (1994) Role of acoustic striae in hearing: mechanism for enhancement of sound detection in cats. *Hear Res* 73:209–222.

- May BJ (2000) Role of the dorsal cochlear nucleus in the sound localization behavior of cats. *Hear Res* 148:74–87.
- May BJ, Huang AY (1996) Sound orientation behavior in cats. I. Localization of broadband noise. *J Acoust Soc Am* 100:1059–1069.
- May BJ, Huang AY (1997) Spectral cues for sound localization in cats: a model for discharge rate representations in the auditory nerve. *J Acoust Soc Am* 101:2705–2719.
- May B, Moody DB, Stebbins WC, Norat MA (1986) Sound localization of frequency-modulated sinusoids by Old World monkeys. *J Acoust Soc Am* 80:776–782.
- May BJ, Huang AY, Aleszczyk CM, Hienz RD (1995) Design and conduct of sensory experiments for domestic cats. In: Klump G, Dooling RJ, Fay RR, Stebbins WC (eds), *Methods in Comparative Psychoacoustics*. Basel: Birkhäuser, pp. 95–108.
- McFadden D, Pasanen EG (1976) Lateralization at high frequencies based on interaural time differences. *J Acoust Soc Am* 59:634–639.
- McGregor PK, Krebs JR (1984) Sound degradation as a distance cue in great tit (*Parus major*) song. *Behav Ecol Sociobiol* 16:49–56.
- McGregor PK, Krebs JR, Ratcliffe LM (1983) The reaction of great tits (*Parus major*) to playback of degraded and undegraded songs: the effect of familiarity with the stimulus song type. *Auk* 100:898–906.
- McGregor P, Horn AG, Todd MA (1985) Are familiar sounds ranged more accurately? *Percept Mot Skills* 61:1082.
- Mendelson MJ, Haith MM (1976) The relation between audition and vision in the human newborn. *Monographs of the Society for Research in Child Development*, 41, Serial No. 167.
- Menzel CR (1980) Head cocking and visual perception in primates. *Anim Behav* 28:151–159.
- Mershon DH, Bowers JN (1979) Absolute and relative cues for the auditory perception of egocentric distance. *Perception* 8:311–322.
- Middlebrooks JC (1992) Narrow-band sound localization related to external ear acoustics. *J Acoust Soc Am* 92:2607–2624.
- Middlebrooks JC (1999) Individual differences in external-ear transfer functions reduced by scaling in frequency. *J Acoust Soc Am* 106:1480–1492.
- Middlebrooks JC, Green DM (1990) Directional dependence of interaural envelope delays. *J Acoust Soc Am* 87:2149–2162.
- Middlebrooks JC, Makous JC, Green DM (1989) Directional sensitivity of sound-pressure levels in the human ear canal. *J Acoust Soc Am* 86:89–108.
- Mills AW (1958) On the minimum audible angle. *J Acoust Soc Am* 30:237–246.
- Mills AW (1960) Lateralization of high frequency tones. *J Acoust Soc Am* 32:132–134.
- Molitor SC, Manus PB (1997) Evidence for functional metabotropic glutamate receptors in the dorsal cochlear nucleus. *J Neurophysiol* 77:1889–1905.
- Moore BCJ (1999) Modulation minimizes masking. *Nature* 397:108–109.
- Moore CN, Casseday JH, Neff WD (1974) Sound localization: the role of the commissural pathways of the auditory system of the cat. *Brain Res* 82:13–26.
- Moore DR (1991) Anatomy and physiology of binaural hearing. *Audiology* 30:125–134.
- Muir D, Field J (1979) Newborn infants orient to sounds. *Child Dev* 50:431–436.
- Musicant AD, Chan JCK, Hind JE (1990) Direction-dependent spectral properties of cat external ear: new data and cross-species comparisons. *J Acoust Soc Am* 87:757–781.
- Neff WD, Casseday JH (1977) Effects of unilateral ablation of auditory cortex on monaural cat's ability to localize sound. *J Neurophysiol* 40:44–52.
- Nelken I, Young ED (1994) Two separate inhibitory mechanisms shape the responses of

- dorsal cochlear nucleus type IV units to narrowband and wideband stimuli. *J Neurophysiol* 71:2446–2462.
- Nelken I, Kim PJ, Young ED (1997) Linear and nonlinear spectral integration in type IV neurons of the dorsal cochlear nucleus. II. Predicting responses with the use of nonlinear models. *J Neurophysiol* 78:800–811.
- Nelken I, Rotman Y, Bar Yosef O (1999) Responses of auditory-cortex neurons to structural features of natural sounds. *Nature* 397:154–157.
- Norberg RA (1977) Occurrence and independent evolution of bilateral ear asymmetry in owls and implications on owl taxonomy. *Philos Trans R Soc Lond B* 282:375–408.
- Nyborg W, Mintzer D (1955) Review of sound propagation in the lower atmosphere. U.S. Air Force WADA Tech. Rept 54–602.
- Payne RS (1962) How the barn owl locates its prey by hearing. *Living Bird* 1:151–159.
- Perrott DR, Ambarsoom H, Tucker J (1987) Changes in head position as a measure of auditory localization performance: auditory psychomotor coordination under monaural and binaural listening conditions. *J Acoust Soc Am* 85:2669–2672.
- Populin LC, Yin TC (1998a) Behavioral studies of sound localization in the cat. *J Neurosci* 18:2147–2160.
- Populin LC, Yin TC (1998b) Pinna movements of the cat during sound localization. *J Neurosci* 18:4233–4243.
- Potash M, Kelly J (1980) Development of directional responses to sounds in the infant rat (*Rattus norvegicus*). *J Comp Physiol Psychol* 94:864–877.
- Pralong D (1996) The role of individualized headphone calibration for the generation of high fidelity virtual auditory space. *J Acoust Soc Am* 100:3785–3793.
- Pratt CC (1930) The spatial character of high and low tones. *J Exp Psychol* 13:278–285.
- Ramachandran R, May BJ (2002) Functional segregation of ITD sensitivity in the inferior colliculus of decerebrate cats. *J Neurophysiol* 88:2251–2261.
- Ramachandran R, Davis KA, May BJ (1999) Single-unit responses in the inferior colliculus of decerebrate cats. I. Classification based on frequency response maps. *J Neurophysiol* 82:152–163.
- Rauschecker JP, Harris LR (1983) Auditory compensation of the effects of visual deprivation in the cat's superior colliculus. *Exp Brain Res* 50:69–83.
- Rauschecker JP, Kniepert U (1994) Auditory localization behaviour in visually deprived cats. *Eur J Neurosci* 6:149–160.
- Ravizza RJ, Masterton RB (1972) Contribution of neocortex to sound localization in opossum (*Didelphis virginiana*). *J Neurophysiol* 35:344–356.
- Rayleigh JWS (1876) Our perception of the direction of a sound source. *Nature (Lond)* 14:32–33.
- Rayleigh JWS (1945) *The Theory of Sound*, 2nd ed. New York: Dover.
- Renaud DL, Popper AN (1975) Sound localization by the bottlenose porpoise *Tursiops truncatus*. *J Exp Biol* 63:569–585.
- Rheinlaender J, Gerhardt HC, Yager DD, Capranica RR (1979) Accuracy of phonotaxis by the green treefrog (*Hyla cinerea*). *J Comp Physiol* 133:247–255.
- Rice JJ, May BJ, Spirou GA, Young ED (1992) Pinna-based spectral cues for sound localization in cat. *Hear Res* 58:132–152.
- Roder B, Teder-Salejarvi W, Sterr A, Rosler F, Hillyard SA, and Neville HJ (1999) Improved auditory spatial tuning in blind humans. *Nature* 400:162–166.
- Rose JE, Brugge JF, Anderson DJ, Hind JE (1967) Phase-locked responses to low-

- frequency tones in single auditory nerve fibers of the squirrel monkey. *J Neurophysiol* 30:769–793.
- Searle CL, Braida LD, Cuddy DR, Davis MF (1975) Binaural pinna disparity: another auditory localization cue. *J Acoust Soc Am* 57:448–455.
- Shaw EA (1974a) Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. *J Acoust Soc Am* 56:1848–1861.
- Shaw EAG (1974b) The external ear. In: Keidel WD, Neff WD (eds), *Handbook of Sensory Physiology*, Vol. V/I. Berlin: Springer-Verlag, pp. 455–490.
- Simpson WE, Stanton LD (1973) Head movement does not facilitate perception of the distance of a source of sound. *Am J Psychol* 86:151–159.
- Spigelman MN, Bryden MP (1967) Effects of early and late blindness on auditory spatial learning in the rat. *Neuropsychologia* 5:267–274.
- Spiro GA, Young ED (1991) Organization of dorsal cochlear nucleus type IV unit response maps and their relationship to activation by bandlimited noise. *J Neurophysiol* 66:1750–1768.
- Spiro GA, May BJ, Wright DD, Ryugo DK (1993) Frequency organization of the dorsal cochlear nucleus in cats. *J Comp Neurol* 329:36–52.
- Spiro GA, Davis KA, Nelken I, Young ED (1999) Spectral integration by type II interneurons in dorsal cochlear nucleus. *J Neurophysiol* 82:648–663.
- Strybel TZ, Perrott DR (1984) Discrimination of relative distance in the auditory modality: the success and failure of the loudness discrimination hypothesis. *J Acoust Soc Am* 76:318–320.
- Sutherland DP, Masterton RB, Glendenning KK (1998) Role of acoustic striae in hearing: reflexive responses to elevated sound-sources. *Behav Brain Res* 97:1–12.
- Terhune JM (1974) Directional hearing of the harbor seal in air and water. *J Acoust Soc Am* 56:1862–1865.
- Thompson GC, Masterton RB (1978) Brain stem auditory pathways involved in reflexive head orientation to sound. *J Neurophysiol* 41:1183–1202.
- van Schaik A, Jin C, Carlile S (1999) Human localisation of band-pass filtered noise. *Int J Neural Syst* 9:441–446.
- von Békésy GV (1938) Über die Entstehung der Entfernungsempfindung beim Hören. *Akust Z* 3:21–31. [Available in English in Weaver EG (ed), *Experiments in Hearing*. New York: John Wiley, 1960, pp. 301–313.]
- Waser PM (1977) Sound localization by monkeys: a field experiment. *Behav Ecol Sociobiol* 2:427–431.
- Waser PM, Brown CH (1986) Habitat acoustics and primate communication. *Am J Primatol* 10:135–154.
- Watkins AJ (1978) Psychoacoustical aspects of synthesized vertical locale cues. *J Acoust Soc Am* 63:1152–1165.
- Wenzel EM, Arruda M, Kistler DJ, Wightman FL (1993) Localization using nonindividualized head-related transfer functions. *J Acoust Soc Am* 94:111–23.
- Wetheimer M (1961) Psychomotor coordination of auditory and visual space at birth. *Science* 134:1692.
- Wettschurek RG (1973) Die absoluten Unterschiedswellen der Richtungswahrnehmung in der Medianebene beim natürlichen Hören, sowie beim Hören über ein Kunstkopf-Übertragungssystem. *Acustica* 28:197–208.
- Whitehead JM (1987) Vocally mediated reciprocity between neighboring groups of mated howling monkeys, *Alouatta palliata palliata*. *Anim Behav* 35:1615–1627.

- Whittington DA, Hepp-Reymond MC, Flood W (1981) Eye and head movements to auditory targets. *Exp Brain Res* 41:358–363.
- Wightman FL, Kistler DJ (1989a) Headphone simulation of free-field listening. I: Stimulus synthesis. *J Acoust Soc Am* 85:858–867.
- Wightman FL, Kistler DJ (1989b) Headphone simulation of free-field listening. II: Psychophysical validation. *J Acoust Soc Am* 85:868–878.
- Wightman FL, Kistler DJ (1992) The dominant role of low-frequency interaural time differences in sound localization. *J Acoust Soc Am* 91:1648–1661.
- Wightman FL, Kistler DJ (1999) Resolution of front-back ambiguity in spatial hearing by listener and source movement. *J Acoust Soc Am* 105:2841–2853.
- Wiley RH, Richards DG (1978) Physical constraints on acoustic communication in the atmosphere: implications for the evolution of animal vocalization. *Behav Ecol Sociobiol* 3:6–94.
- Xu L, Furukawa S, Middlebrooks JC (1999) Auditory cortical responses in the cat to sounds that produce spatial illusions. *Nature* 399:688–691.
- Yin TC, Kuwada S (1983) Binaural interaction in low-frequency neurons in inferior colliculus of the cat. II. Effects of changing rate and direction of interaural phase. *J Neurophysiol* 50:1000–10019.
- Young ED, Spirou GA, Rice JJ, Voigt HF (1992) Neural organization and responses to complex stimuli in the dorsal cochlear nucleus. *Philos Trans R Soc Lond B Biol Sci* 336:407–413.
- Young ED, Rice JJ, Tong SC (1996) Effects of pinna position on head-related transfer functions in the cat. *J Acoust Soc Am* 99:3064–3076.
- Zakarauskas P, Cynader MS (1993) A computational theory of spectral cue localization. *J Acoust Soc Am* 94:1323–1331.