Annu. Rev. Psychol. 1991. 42:135–59

SOUND LOCALIZATION BY HUMAN LISTENERS

John C. Middlebrooks

Departments of Neuroscience and Surgery (ENT), University of Florida, Gainesville, Florida 32610

David M. Green

Psychoacoustics Laboratory, Department of Psychology, University of Florida, Gainesville, Florida 32611

KEY WORDS: sound localization, directional hearing, binaural hearing, monaural localization, directional transfer function

CONTENTS

INTRODUCTION

The task of localizing a sound source presents a challenge to the integrative capabilities of the nervous system. In the visual system, locations of objects in the visual world are focused by the optics of the eye directly onto the retina,

0066-4308/91/0201-0135$02.00

and maps of visual space are found in structures of the central nervous system that preserve the intrinsic topography of the retina. Inasmuch as the wavelength of sound is six orders of magnitude greater than that of light, an acoustical focusing mechanism analogous to the optics of the eye is not feasible, and a different mechanism must be employed. The sound wave generated by an external source is diffracted by its interaction with the head and external ears. The resulting changes in the temporal and intensive characteristics of the acoustical stimulus provide cues about the locus of a sound relative to the head. The goal of this chapter is to characterize the acoustical cues for sound localization and to assess the strategies by which subjects utilize these cues to determine the location of a sound source.

Before beginning the review, we must comment on some general features of practically all studies of sound localization. A common objective of many studies is to test the efficacy of some stimulus aspect (interaural level difference, interaural temporal difference, spectral shape, etc) in revealing the locus of the source. With all scientific studies, a degree of abstraction occurs. An important restriction in most behavioral experiments on sound localization is that the potential responses are restricted to a few locations (e.g. a subject must select from a small number of loudspeakers located on a particular plane). Another common restriction is a limitation in the bandwidth or duration of the acoustic stimulus. Although a result might clearly establish the effectiveness of a particular cue in a certain experimental setting, there often is no evidence to indicate the importance of that cue in a more realistic setting. For example, listeners might correctly localize source A 2° to the right of source B, given only an interaural intensity difference at 4720 Hz. That same cue might be totally ignored in determining the site of a broadband sound source that is unrestricted in its possible location. The restricted experiment provides information about the *sensitivity* of the listener to a particular cue. It does not provide information about the larger, and more interesting, issue concerning the synthesis of a central image corresponding to the external sound source. In reviewing studies of sound localization, then, one must be aware that different researchers will champion different cues as most important in localizing sound sources. Each claim is demonstrably valid in some restricted setting. As the set of potential cues becomes better understood, the next item on the agenda is to determine how these different cues are weighted and integrated in a setting in which multiple cues are present. To the extent that it is possible, we try to address this larger issue here.

Sound localization has traditionally received considerable attention in behavioral and physiological studies in animals, and those areas have been the topics of recent reviews (Masterton & Imig 1984; Phillips 1985). Moreover, physiologists and anatomists have made great progress in understanding the mechanisms by which the nervous system detects specific cues for sound localization (Rose et al 1966; Boudreau & Tsuchitani 1968; Goldberg & Brown 1969). In recent years, increasing attention has been focused on localization by human listeners. Evidence of this interest is provided by the recent publication of a book edited by Yost & Gourevitch (1987) that contains chapters by the participants in a symposium on directional hearing held earlier by the Acoustical Society of America. Additional interest has been stimulated by a meeting in 1988 of the National Research Council's Committee on Bioacoustics and Biomechanics (CHABA), the topic of which was sound localization by humans. Part of this renewed interest undoubtedly stems from the 1983 translation of Blauert's classic work on *Spatial Hearing* (Blauert 1974). This review attempts to integrate some of the more recent experimental results with the classic studies. We give only passing attention to experiments using stimuli presented over headphones (e.g. studies of masking level difference and lateralization) in order to focus on issues of localization in a free sound field. Furthermore, since most of the results to be discussed were obtained in anechoic environments, we pay little attention to issues such as localization in reverberant rooms (Blauert 1983; Hartmann 1983) and the precedence effect (Zurek 1980; Gaskell 1983; Clifton 1987).

For convenience in presentation, we draw a somewhat arbitrary distinction between localization in the horizontal and vertical dimensions. To some extent, this distinction is dictated by the design of most previous localization studies. That is, in most studies, the sound sources have been restricted to one of two cardinal planes, either the horizontal plane, defined roughly by the two ear canals and the tip of the nose, or the median (mid-sagittal) plane. The distinction between horizontal and vertical localization also appears to be justified by differences in the principal spatial cues for horizontal and vertical localization (i.e. interaural difference cues vs spectral cues). One can influence these cues differentially by plugging one ear or by filling the convolutions of the pinnae. We begin by reviewing recent studies of localization in which stimulus locations were varied freely in both horizontal and vertical dimensions; we then consider the available spatial cues and behavioral performance specifically in the horizontal compared to the vertical dimension. We pay separate attention to the perception of sound source distance, to the perception of stimulus motion, and to dynamic spatial cues provided by movements of the head.

TWO-DIMENSIONAL SOUND LOCALIZATION

When a subject is presented with a source that can appear at any position about the subject's head, and that source presents the full complement of acoustical spatial cues, how effectively does the listener use that information to synthesize the locus of a sound source? The two dimensions in this localization task are *azimuth* and *elevation*. In this review, localizations are given in a double-pole coordinate system (see Knudsen 1982) in which azimuth is defined as the angle given by the sound source, the center of the listener's head, and the median plane; this is the angle in the horizontal dimension. Elevation is defined as the angle given by the sound source, the center of the head, and the horizontal plane. The origin (0°, 0°) is straight in front of the subject.

Information about sound localization in the two-dimensional space about the listener is available in the recent studies by Oldfield & Parker (1984a), Wightman & Kistler (1989b), and Makous & Middlebrooks (1990). These studies had two basic methodological features in common. First, the stimuli had broadband spectra. The broad bandwidth provided the subjects with means to resolve the spatial ambiguities that are present within any narrow frequency band. Second, the locations of sound sources were varied freely in both the horizontal and vertical dimensions. This avoided a limitation of most previous localization studies in which much of the spatial ambiguity inherent in the localization task was removed by constraining the possible source locations to a single plane. Absolute levels of localization performance varied substantially among these three studies, owing largely to differences in the stimulus bandwidths and durations, differences in the amount of variability in spectra between trials, and differences in the methods by which subjects were required to report the apparent locations of sound sources. Despite the variation in absolute levels of performance, the three studies convey a consistent picture of the general spatial dependence of localization performance.

Figure 1 shows localization data from one subject in the study by Makous & Middlebrooks (1990). Stimulus locations and the subject's responses are drawn on the surface of an imaginary sphere, 1.2 m in radius, centered on the subject's head. Asterisks indicate the locations of stimuli, and open circles indicate the response locations. Sizes of errors and response variability were smallest for stimuli directly in front of the subject and increased at more peripheral stimulus locations. The smallest errors in that study, averaged across trials and across subjects, were about 2° and 3.5° in azimuth and elevation, respectively, increasing to average errors of as much as 20° for some rear locations. Across elevation, the variability in the judgments of azimuth was relatively constant for any given azimuth and, similarly, across azimuth, the variability in the judgments of elevations was relatively constant for any given elevation. The scatter of responses in azimuth was smaller than that in elevation for stimuli near the frontal midline, but the opposite was true for more peripheral locations.

Most sound localization studies have noted the occurrence of front/back confusions, in which a stimulus in front of the subject is localized to the rear, or vice versa. When front/back confusions occur, the azimuthal component of

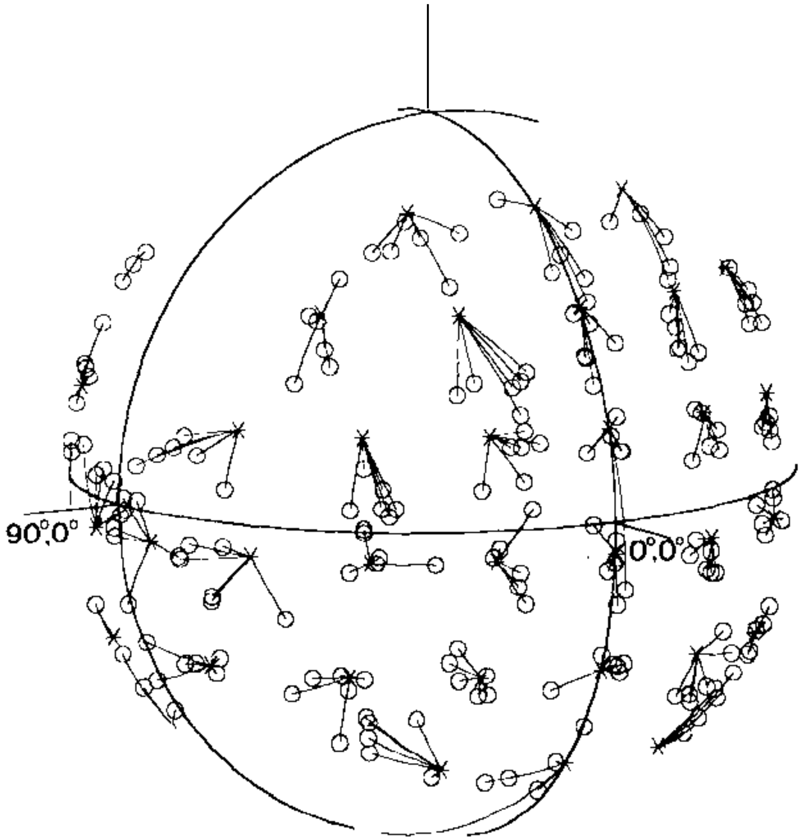

Figure 1 Localization of a broadband sound. Stimulus and response locations are drawn on the surface of an imaginary sphere as if looking in toward the subject from point 30° to the subject's right and elevated 10°. Asterisks indicate the stimulus location and open circles indicate five responses to that stimulus given by a single subject. For clarity, stimulus locations are shown in 20° steps of azimuth and elevation, although data were collected in 10° steps. The data are from the study by Makous & Middlebrooks (1990).

the stimulus and response locations tend to lie in mirror symmetry with respect to the interaural axis. The pattern of elevation errors is less regular (Makous & Middlebrooks 1990). The frequency of occurrence of front/back confusions tends to increase as the bandwidth of the stimulus is decreased (Butler 1986). For example, Makous & Middlebrooks (1990) encountered front/back confusions in 2–10% of localizations of broadband stimuli, depending on the subject, whereas subjects made front/back confusions in more than 20% of the trials when localizing one-octave noise bands (Burger 1958).

HORIZONTAL LOCALIZATION

No study of sound localization in the horizontal dimension has escaped the influence of a series of observations made by Lord Rayleigh near the end of the 19th century and reported in his Segwick lecture in 1906 (Rayleigh 1907). Rayleigh attempted to account for localization in terms of interaural difference cues. He appreciated that when a sound was presented from the side, a listener's head would interrupt the path from the source to the far ear. The far ear would be effectively shadowed and an interaural difference in sound pressure level (ILD) would result. The amount of shadowing depends on the wavelength of the sound compared with the dimensions of the head. At high frequencies, the shadow amounts to a difference of as much as 35 dB between the two ears for a source located at the side (Middlebrooks et al 1989). For frequencies below about 1000 Hz, however, the sound wavelength can be several times larger than the head and, as Rayleigh was able to compute, the ILD would be negligible. Rayleigh "reluctantly" came to accept that "When a pure tone of low pitch is recognized as being on the right or the left, the only alternative to the intensity theory is to suppose that the judgment is founded upon the difference of *phases* at the two ears" (Rayleigh 1907). Rayleigh confirmed the sensitivity of listeners to interaural phase differences (IPDs) by presenting to the two ears a pair of tuning forks that were tuned to slightly different frequencies. The slightly mistuned forks produced a steadily varying phase difference, which produced the sensation of an auditory image that moved back and forth between the two ears. Rayleigh found that the sensitivity to phases declined with increasing frequency, with an upper limit in his experiments around 770 Hz. The notion that spatial information is derived at high frequencies from ILDs and at low frequencies from IPDs is often referred to as the "duplex" theory of sound localization.

Studies of localization in the horizontal plane have produced results consistent with the duplex theory. In an experiment by Stevens & Newman (1936), sounds were presented from a loudspeaker mounted on a boom 12 feet long, and subjects attempted to report the angular location of the sound source. Although broadband stimuli (i.e. clicks and noise bursts) could be localized with reasonable accuracy, subjects often made front/back confusions when attempting to localize sinusoidal stimuli. Discounting the front/back confusions, there was a clear frequency dependence to the localization performance. Specifically, the magnitudes of errors peaked around 3000 Hz and declined at higher and lower frequencies. A similar frequency dependence in localization errors was observed by Sandel and colleagues (1955), who required subjects to adjust the position of a broadband noise source to correspond to the apparent source of a sinusoid: Performance was worst for sinusoids around 1500–3000 Hz. Interpreted with reference to the duplex theory, the accuracy of localization of low frequencies would be attributed to IPDs and that at high frequencies to ILDs. In the range around 1500 to 3000 Hz, stimuli are too high in frequency to provide usable phase cues and too long in wavelength to provide adequate ILDs.

Mills (1958) measured spatial acuity by asking subjects to distinguish between two loudspeakers separated in the horizontal dimension. Consistent with measurements of localization accuracy, the "minimum audible angle" measured in this way was maximal around 3000 Hz. Mills noted that, when the loudspeakers were placed symmetrically around 90° to the side, the minimum audible angle exceeded the largest angle that Mills could test; in this configuration, he noted that the minimum audible angle task is equivalent to a front/back discrimination.

Accuracy in sound localization in the horizontal dimension can be accounted for largely in terms of the spatial dependence of interaural difference cues and the sensitivity of subjects to those cues. The minimum thresholds for ILDs in tones are less than 1 dB (Mills 1960). Hafter and colleagues (1977) have shown that the threshold for changing ILD varies by no more than a factor of two over a broad range of ILDs. The spatial dependence of ILDs has been measured directly by recording from the two ear canals with miniature microphones while presenting sounds at varying locations around the head (Searle et al 1975; Middlebrooks et al 1989); ILD values can also be estimated from other studies that used monaural recordings by assuming that the head is bilaterally symmetrical (e.g. Shaw 1974; Mehrgardt & Mellert 1977). In the results of Middlebrooks and colleagues (1989), maximum ILDs measured near 90° azimuth were around 20 dB at the lowest frequency tested, 4 kHz, and rose to as high as around 35 dB at 10 kHz. When averaged across any perceptually relevant bandwidth, ILDs are roughly proportional to the sine of the azimuthal angle and are roughly constant across elevation at any constant azimuth. Thus, one might expect the sizes of errors in localization and the sizes of minimum audible angles to increase monotonically with increasing azimuth. The ILD produced by a tone (or a single component of a complex sound), however, can show substantial local variation with source location (Middlebrooks et al 1989). Consistent with that view, Mills (1958) found that minimum audible angles for tonal stimuli failed to show an orderly dependence on azimuth for frequencies greater than 4 kHz and for azimuths greater than 30°.

The spatial dependence of IPDs has been estimated by treating the ears as two points on the surface of a rigid sphere and determining, geometrically, the difference in the paths from a sound source to the two points (Woodworth 1938). This analysis gives:

$$\tau = (r/c)\,[\theta + \sin(\theta)] \qquad 1.$$

where τ is the interaural delay, c is the speed of sound in air, r is the radius of the sphere, and θ is the azimuthal angle of the sound source in radians. Kuhn (1977, 1987) made a more acoustically based analysis of the delay between two points on a rigid sphere and measured the interaural delays on a mannequin. At frequencies above about 2 kHz, the interaural delay given by Equation 1 corresponds well to measured values. At lower frequencies (below 500 Hz), the acoustical prediction and the measured values are larger than those predicted by this equation by a factor of 3/2.

When tonal stimuli are presented over headphones, human listeners show sensitivity to interaural differences in the ongoing phase of the stimulus for frequencies up to about 1000 to 1300 Hz (Licklider et al 1950; Zwislocki & Feldman 1956). The minimum audible angle for low-frequency tones, predicted from sensitivity to phase differences and the spatial dependence of phase differences, corresponds well with measured minimum audible angles and localization performance (Mills 1958). It is somewhat surprising that humans fail to detect interaural phase differences at higher frequencies, given that significant phase coding has been shown in the auditory nerve of other primates at frequencies up to 5 kHz (Rose et al 1967). The maximum interaural delay for an average human is around 700 μsec, which is equal to the period of a 1400 Hz tone. At higher frequencies, interaural phase would be ambiguous by steps of 360°, and phase leads and lags would be ambiguous even at frequencies as low as 700 Hz, at which the maximum interaural delay corresponds to 180° of phase. Thus, one possibility is that the human auditory system simply rejects phase information at frequencies at which that information would be ambiguous. That explanation does not seem to apply to the barn owl, however, which exhibits both neurophysiological and behavioral sensitivity to interaural phase differences at frequencies greater than 7 kHz (Moiseff & Konishi 1981). At those high frequencies, phase ambiguity is apparent in both neurophysiological and behavioral data, but the phase difference continues to be a salient localization cue. When an owl is presented with a stimulus that contains two or more frequency components, sufficient delay information is available to resolve the phase ambiguity, thus enabling accurate localization in azimuth based on interaural time differences (Knudsen 1984).

Despite the insensitivity of humans to ongoing phase differences in tones above 1000 to 1300 Hz, listeners can detect interaural delays in the *envelopes* of high-frequency sounds. Sensitivity to envelope delays has been demonstrated for highpass and high-frequency bandpass transients (Yost et al 1971; Hafter & Dye 1983), for bandpass noise (Klumpp & Eady 1956; Trahiotis & Bernstein 1986; Amenta et al 1987), for sinusoidally amplitude-modulated (SAM) tones (Henning 1980; Nuetzel & Hafter 1981), and for two-tone complexes (McFadden & Pasanen 1976; McFadden & Moffitt 1977). There is some indication that envelope delay sensitivity declines for carrier frequencies above about 4 kHz or for modulation frequencies above about 500 Hz. The detection of an interaural delay in a two-tone complex probably entails a different mechanism from that used by the barn owl, since the owl is sensitive to phase differences in each of the components of the complex, whereas the human listener can detect a delay only when two or more components are present. There is some question about whether envelope delays in high-frequency sounds provide humans with usable localization cues. The directional dependence of envelope delays suggests that the physical difference would provide a reliable cue to horizontal location (Middlebrooks & Green 1990). When carriers greater than about 1.6 kHz are presented over headphones, however, the extent of laterality provided by an interaural delay is small (Blauert 1982; Bernstein & Trahiotis 1985; Trahiotis & Bernstein 1986). Moreover, the modulation depth of a signal processed through a filter as wide as a critical band might be too small to permit detection of envelope delays (discussed by Middlebrooks & Green 1990).

VERTICAL LOCALIZATION

To the extent that the head and ears are symmetrical, a stimulus presented at any location on the median plane should produce no interaural differences and, thus, interaural differences should provide no cue to the vertical locations of sounds on the median plane. Similarly, any point off this median plane falls on a "cone of confusion," as Woodworth (1938) called it, upon which the interaural differences, either ILD or IPD, are constant. Although an actual set of spatially ambiguous points might differ from an idealized cone, it is clear that the spatial information provided by an interaural difference cue within a restricted band of frequency is spatially ambiguous, particularly along a roughly vertical and front/back dimension.

Batteau (1967, 1968) was one of the first to emphasize that the external ear, specifically the pinna, could be a source of spatial cues that might account for localization. He suggested that sound reflections within the convolutions of the pinna might produce spatial cues that could be used to disambiguate the potential locations of the source. That these cues could rival in importance the interaural difference cues was supported by behavioral data from an experiment reported by Fisher & Freedman (1968). They had subjects discriminate the locations of eight loudspeakers located at 45° angles about the listener in the horizontal plane. Given the paucity of potential source locations, they found that the accuracy of the localization judgments was about the same with one ear occluded as with both ears open. Localization suffered only when a 10-cm tube inserted in the ear canal effectively bypassed the pinna and deprived the listeners of its sound reflections.

Although Batteau's model was formulated in the *time* domain, the convolutions of the pinna would create echoes lasting only a few microseconds, so most current theories interpret the pinna as producing changes in the *spectrum* of the sound source that reaches the tympanic membrane. The influence of the pinna is to produce multiple paths to the ear canal, among them a direct path and a reflection from the *cavum concha* of the pinna. The addition of a direct signal with a delayed (reflected) version of the same signal produces a "comb-filtered" spectrum containing a characteristic pattern of peaks and notches. The length of the reflected path varies with the elevation of the sound source, so particular spectral features tend to vary in frequency according to elevation. We refer to patterns of spectral features associated with particular locations as "spectral shape cues." Because these cues are a product of the acoustics of the pinna, they often are referred to as "pinna cues." Finally, since spectral shape cues do not require an interaural comparison, they often are referred to as "monaural cues," although, at least for a sound source on the median plane, essentially equivalent information is available from each ear.

Several lines of evidence indicate that spectral shape cues are the major cues for vertical localization. For example, accurate vertical localization is observed only when the stimulus has a broad bandwidth (Roffler & Butler 1968; Gardner & Gardner 1973; Butler & Helwig 1983) and contains energy at high frequencies. The cutoff frequency for the high frequency requirement is subject to some debate (Roffler & Butler 1968; Hebrank & Wright 1974b), but it seems that frequencies above 4 kHz are important. Vertical localization is prevented when the convolutions of the pinnae are occluded (Roffler & Butler 1968; Gardner & Gardner 1973; Oldfield & Parker 1984b), thus altering the ear's directional transfer function. Horizontal localization is relatively unimpaired by such a manipulation, but the pinna does aid in resolving front/back ambiguity (Musicant & Butler 1984a; Oldfield & Parker 1984b). Vertical localization is almost as good when listening with a single ear as with binaural listening (Hebrank & Wright, 1974a; Oldfield and Parker, 1986). Finally, as discussed below, vertical localization is sensitive to manipulations of the source spectrum.

Measurements of the transfer function of the external ear have revealed particular spectral features that vary systematically with source location (Butler 1987; Humanski & Butler 1988; Hebrank & Wright 1974b; Watkins 1978). Blauert (1969/1970) has emphasized the presence of "boosted bands," which are peaks in the transfer function associated with front, overhead, or rear source locations. Butler (1987) defines a roughly equivalent concept as the "covert peak"—a band of frequency that produces the greatest level when presented from a particular source location. He has compiled a "spatial referent map" in which particular covert peaks correspond to specific locations on the median or horizontal plane. Among the most prominent features of the external ear transfer functions, at least on visual inspection of the spectra, are relatively narrowband notches that shift systematically in elevation with sound source elevation (Hebrank & Wright 1974b; Bloom 1977a; Watkins 1978; Butler 1987; Humanski & Butler 1988). Transfer functions generally are similar across different subjects, although the spectra tend to show a downward shift in frequency associated with increasing physical size of subjects (Middlebrooks et al 1989).

The apparent elevation of a sound source can be influenced by modifying the source spectrum. In a widely quoted experiment, Blauert (1969/1970) recorded from a subject's ear canals while presenting a noise from two source locations, either directly in back or directly in front of the listener. He then electronically modified both sources, so that each mimicked the spectrum associated with the opposite source position. The listener's judgments of the source's locus were solely determined by the source spectrum and not influenced by its actual location. Thus, the listener localized the source as behind if it had the "back" spectrum, even if played from the front speaker. In a separate experiment, Blauert presented one-third-octave noise bands from the median plane and showed that the probability of a subject's responding that a sound source was in front, overhead, or behind was determined more by the center frequency of the stimulus than by actual source locations. A similar result was reported by Butler & Helwig (1983).

Bloom (1977b) examined the spatial character of spectral notches. He presented stimuli from a fixed location and asked listeners to adjust the center frequency of a band-reject filter so that the perceived elevation most closely matched one of a set of sources located at various elevations in the median plane. The center frequencies selected by the listeners tended to correlate with the frequencies of the notches in the external ear transfer functions measured for stimuli located at corresponding elevations. Similarly, Hebrank & Wright (1974b) found that certain band-reject, band-pass, and high-pass filter characteristics tended to correlate with particular perceived elevations in the median plane. Watkins (1978) processed noise with a delay-and-add system that was intended to mimic the acoustical reflections within the external ear. These signals were presented to listeners over headphones, and the listeners adjusted a mechanical pointer to indicate the perceived source elevation. The reported elevations corresponded well with those predicted by mechanisms proposed by Batteau (1967).

Any model of localization based on spectral cues must acknowledge the importance of a priori information regarding the source spectrum. That is, the presence of a particular spectral feature at the tympanic membrane could be the result of the transfer characteristic of the ear, or it could be a feature that was present in the source spectrum. The experiments described above indicate that the auditory system cannot effectively distinguish the two. Blauert states that his model of vertical localization applies "for nearly all usual signals, as long as their power density curve is to some extent smooth." It remains to be tested how "smooth" a signal must be to permit accurate vertical localization.

This section on vertical localization began with an assumption that the two ears are essentially bilaterally symmetrical and that interaural differences provide no cue to elevation. This assumption may not be strictly valid. Measurements by Searle and colleagues (1975) and by Middlebrooks and colleagues (1989) have demonstrated that, at least within narrow bands of frequency, ILDs along the median plane can be as large as about 10 dB and that ILDs can vary with source elevation at a constant source azimuth. There is some controversy over whether or not ILD cues actually contribute to localization in elevation, particularly on the median plane. For example, Searle and colleagues (1976) devised a model based on statistical decision theory to account for results of published localization trials. Their model predicts that interaural difference ("interaural pinna") cues actually contribute more than monaural spectral shape cues to localization in the median plane. Moreover, several groups have found that plugging one ear can produce a substantial decrement in performance in a median-plane localization task (Butler 1969; Gardner 1973; Ivarsson et al 1980). One must be cautious in interpreting such a result, since the presence of an earplug would cause normal, spectral shape cues to occur in association with unfamiliar interaural difference cues, thus impairing localization. Hebrank & Wright (1974a) found that, initially following the placement of an earplug, subjects showed impaired localization in the median plane. After a short period of training, however, the ability to localize a source in the median plane was nearly the same for monaural or binaural listening. Currently, the evidence for a major contribution of ILDs to vertical localization, at least in the median plane, is not compelling.

MONAURAL LOCALIZATION

Although most investigators agree that two ears are better than one in localizing sounds in space, there is considerable evidence showing that some localization can be achieved with a single ear. Evidence on this issue comes from as early a study as that of Angell & Fite (1901) and from a more recent study by Fisher & Freedman (1968). Since most studies have produced the "monaural" listener by occluding the entrance to one ear canal, we should comment briefly on the controls that must be employed in these studies. A high-quality earplug can produce 30-40 dB attenuation at low frequencies and somewhat greater attenuation at higher frequencies. If the sound level is greater than 40 dB sensation level at the occluded ear, then, although ILDs are severely altered, interaural temporal comparison might be relatively unaffected. Thus, it is critical to conduct the localization tests with the source at a low sensation level, so that the occluded ear will be completely nonfunctional. In the studies we review, this control was exercised.

Oldfield & Parker (1986) measured two-dimensional localization under monaural conditions. Subjects localized a white-noise source that was positioned in 10° intervals ranging from −40° to +40° in elevation and 0° to 180° in azimuth on the left side. In the monaural condition, the absolute error in elevation (12°) was only slightly greater than in the binaural condition (9°). The absolute error in azimuth, however, was barely on the same scale: 30–40° for monaural listening compared to 5°–10° for binaural listening.

The Oldfield & Parker results suggest that binaural cues are primarily responsible for determining the source's azimuth and that monaural, presumable spectral shape, cues are primarily responsible for resolving the source's elevation. On the other hand, Butler and colleagues (Butler 1986; Butler & Flannery 1980; Belendiuk & Butler 1975; Musicant & Butler 1984b) have emphasized that monaural cues also contribute to localization in azimuth. Belendiuk & Butler (1975) established that when a source contains energy above 4 kHz, subjects wearing a plug in one ear can achieve better than chance localization of one of five loudspeakers located in 15° intervals on a horizontal plane in the quadrant between the unoccluded ear and straight ahead. Musicant & Butler (1984a) also showed that monaural spectral cues help resolve front/back confusions in the horizontal plane of the unoccluded ear. Butler & Flannery (1980) and Musicant & Butler (1984b, 1985) showed that narrow bands of noise (1 kHz wide) appear to move back to front in azimuth as the center frequency of the band is varied. Although responses varied among subjects, some subjects showed the following trends. As center frequencies shift from 4 to 10 kHz, the reported azimuth moves from front to back, repeating that trend for frequencies centered between 10 to 14 kHz. In those experiments, the sound was presented from only a few of the many potential loudspeaker locations. The speaker locations were visible to the listener. Whether such constraints were necessary to achieve these localization judgments is unknown.

Again, we remind the reader of the caveat expressed at the beginning of this review. There is little doubt that some of these cues can operate in certain restricted listening situations. Strong biases may develop, however, if the responses are restricted to a limited number of speaker locations. For example, Butler & Planert (1976) show an almost identical distribution of choices among five loudspeakers located in the median plane when the spacing between successive loudspeakers was 15° and when it was 7.5°. The central issue is not whether these cues play some role in certain experiments but whether they contribute significantly to sound localization under more realistic conditions.

A recent experiment tested the relative salience of ILD cues and spectral shape cues for localization in azimuth (Middlebrooks 1990). Subjects localized 1/6-octave bandpass sound sources that were varied simultaneously in azimuth and elevation. Center frequencies range between 6 and 14 kHz. Subjects made substantial and reproducible errors in localization that were largely confined to the vertical dimension. Accuracy of localization in azimuth was near that observed with broadband stimuli. Moreover, the ILDs associated with the response location were roughly the same as those produced by the stimulus. These results suggest that the bandpass filtering operation introduced erroneous spectral shape cues to elevation. The filter probably also influenced spectral cues for azimuth, but the ILD cues for azimuth, unaffected by the filter, were sufficient to support near-normal accuracy in azimuth. That is, the auditory system appears to favor binaural cues over spectral shape cues for azimuth, but must rely on spectral shape cues for elevation.

DISTANCE PERCEPTION

The listener's ability to localize the distance of a sound source is certainly not very good and has been the subject of minimal research. The scholarly review article by Coleman (1963) still provides the best summary of the potential distance cues. Little is known about the effectiveness of any of these cues. If the source intensity is known, then distance can be judged on the basis of sound intensity at the listener's ear. Gardner (1969) showed that the distance of a person speaking in a conversational tone of voice in an anechoic chamber can be judged with some accuracy. The apparent distance of the same speech delivered from a speaker is almost entirely determined by the sound level of the loudspeaker. Thus, familiarity with the signal source is clearly an important variable. In an anechoic environment, the intensity of a source diminishes 6 dB for each doubling of distance. If a change in loudness is the cue for distance, then the just-discriminable change in distance should be predictable from the just-noticeable difference in intensity. Strybel & Perrott (1984) have confirmed that expectation for sources beyond 3 m. For closer distances, much greater changes in distances must be made before the listener can achieve a successful discrimination. Simpson & Stanton (1973) have shown that head motion does not facilitate better distance judgments. Butler et al (1980) listened to sounds over headphones and judged their apparent distance. The apparent distance of the source increased as the low-frequency part of the spectrum increased. Finally, Holt & Thurlow (1969) judged sources located outdoors in a grassy park. The sources ranged in distance from 32 to 64 feet. Noise bursts, 0.5 sec in duration and presented once per second for 10 sec, were adjusted for equal level at the listeners' ears. The estimates of distance, when averaged across listeners and trials, correlated significantly with the actual source distances when the sources were located at one side, but not when placed in front of the listeners.

There is some evidence to suggest that the listener can better estimate the distance of a sound source if the surrounding environment is not anechoic. In an ordinary room, a distant source produces sound energy that reaches the listener's ears via direct and indirect paths. Differences in the ratio of these two energies might produce perceptible differences in the quality of the source as a function of distance. This cue to source distance, however, is strongly influenced by the specific reflections of the particular listening environment. For example, Mershon & Bowers (1979) have demonstrated that estimates of the distance of a wideband noise source (sounded only once for 5 sec) are significantly different from chance in a reverberant room. Apparently these effects are not robust. In an earlier experiment (Mershon & King 1975), the judgments of distance were significantly different from chance for the higher intensity levels but not for the lower intensity level. There are no systematic studies of how the ratio of direct to indirect sound energy influences the judgments of source distance.

MOTION DETECTION

The bulk of the data discussed in this review concerns the ability to identify the location of a stationary sound source or to resolve the locations of two stationary sources. An additional stimulus attribute to be considered is the change of location—that is, the motion of a sound source. We refer here to change in the azimuth and/or elevation of the source, not the change in source distance, which can be cued by a change in sound pressure or by a Doppler shift in the stimulus frequency. In the visual system, there is good evidence for neural systems specialized for the detection of motion (Hubel & Wiesel 1962; Newsome et al 1986). As yet there is no compelling evidence for motion-sensitive systems in the auditory system. The problem is that there are two intepretations for any observation of sensitivity to source motion. One interpretation is that the auditory system is sensitive to dynamic aspects of localization cues—for example, changing ILDs or IPDs. That is, source velocity might be a "directly perceived attribute of moving stimuli" (Lappin et al 1975). A second, more parsimonious interpretation is that the nervous system simply measures the sound source location at two distinct times and interprets a change in location as motion; this has been called the "snapshot theory." The difficulty in resolving these alternatives stems from the fact that most studies of motion detection tend to confound the attributes of duration, velocity, and net change in location. For example, an increase in sound source velocity will result in an increase in the total distance traversed by a stimulus of a given duration.

In studies of motion perception, listeners have typically been asked to discriminate between directions of motion (Harris & Sergeant 1971; Perrott & Tucker 1988) or to discriminate between a stationary and a moving sound source (Perrott & Musicant 1977). Although most studies have used sound sources in actual motion, some have simulated motion by systematically varying the levels of sinusoids presented from two loudspeakers (Grantham 1986; Saberi & Perrott 1990). Thresholds have been measured for duration, velocity, and change in location. All these thresholds, however, can be expressed in terms of a minimum audible movement angle (MAMA), which is the smallest net change in location of a moving stimulus that can be detected under some specified set of conditions. The MAMA shares several properties with the minimum audible angle for static sources. Specifically (*a*) MAMAs in azimuth are smallest for stimuli around 0° azimuth and increase with increasing azimuth (Harris & Sergeant 1971; Grantham 1986). A MAMA often cannot be measured for tonal stimuli around 90° azimuth, since the task is confounded by front/back confusions. (*b*) MAMAs are smaller for broadband than for tonal stimuli (Harris & Sergeant 1971; Saberi & Perrott 1990). (*c*) When measured with tonal stimuli, MAMAs are largest (i.e. performance is worst) for a range of frequencies around 1300–2000 kHz (Perrott & Tucker 1988), the range over which minimum audible angles are largest, owing to the lack of effective spatial cues (Mills 1958).

It has been a common observation that MAMAs are as much as several times larger than static minimum audible angles measured under comparable conditions. For example, Perrott & Musicant (1977) measured a MAMA of 8.3° near 0° azimuth for a 500-Hz tone, whereas the minimum audible angle measured by Mills at 800 Hz was closer to 1°. Perrott & Musicant (1977) showed that, over a range of sound source velocities from 2.8–360°/sec, the MAMA increases linearly with increasing velocity; Saberi & Perrott (1990), looking with finer resolution in the frequency dimension, recently showed that the MAMA actually only begins to grow at velocities around 10°/sec. The finding that MAMAs increase with source velocity is consistent with Grantham's (1986) observation that discrimination deteriorates for stimulus durations less than a minimum "integration time" of about 150–300 msec. That is, the angle traversed by a sound source within the time required for an accurate localization judgment increases with increasing source velocity. At the low end of the frequency range, Saberi & Perrott (1990) also reported that MAMAs can show an increase when source velocities *decrease* below about 1.8°/sec. This decrease in performance probably occurs because, at such low velocities, the duration of the stimulus (5–10 sec) exceeds the storage time of the auditory system. Perrott & Musicant (1977), in a separate experiment, asked listeners to indicate the apparent position of a moving source at stimulus onset and offset. The onset judgments were consistently displaced toward the direction of movement, and the displacement increased with increasing source velocity. This result might also be explained by the presence of a minimum integration time.

None of the results reviewed above distinguishes between a "snapshot" theory and a theory invoking specialized motion-sensitive systems. Perrott & Marlborough (1989) recently tried to test the snapshot theory directly. They presented a noise burst from a source that moved in azimuth in front of the subject at a velocity of 20°/sec, and the subject was required to discriminate the direction of the movement. In one condition, the stimulus was presented continuously. The MAMA was about 1°, corresponding to a stimulus duration of 50 msec. In a second condition, the source sounded in two discrete 10-msec bursts—that is, marking the beginning and end of the loudspeaker traverse. The silent interval between bursts was adjusted to find a threshold for discrimination of movement direction. According to the authors' interpretation, the snapshot theory would predict equal angles measured under the two conditions. The observation was that thresholds measured in the second condition were about 50% larger. Perrott & Marlborough (1989) conclude that "the information arriving after the onset and before the offset of the signal does contribute to the resolution of motion"; that is, the listener uses spatial information collected throughout the traverse of the sound source. Alternative explanations are possible. Pollack & Rose (1967) showed that the average error in localization of a stationary target increases about 60% when the duration of a noise source is shortened from 50 to 20 msec. Thus, one can argue that two 10-msec sounds do not provide sufficient information to localize reliably the beginning and end of the sound source traverse. That is, longer "snapshots" must be provided to equal the 50-msec continuous condition.

In vision, it is common to claim that a motion detector exists, since prolonged exposure to one direction of motion results in an increased threshold for motion in the same direction and no change in the threshold for motion in the opposite direction (Sekuler & Pantle 1967; Pantle & Sekuler 1968). No similar effect has been observed in audition, although Grantham & Wightman (1979), using earphones, have reported a "motion aftereffect" similar in form to the visual "waterfall effect." Grantham (1989) has repeated this experiment using an actual moving source in an anechoic room. He shows that part of the effect is a simple response bias, but there is also a small component showing loss of sensitivity to motion in the adapted direction. Because the auditory sense, as a whole, shows considerably less adaptation than vision, a more appropriate test might be borrowed from Lappin and colleagues (1975). One could construct four distinct moving stimuli by selecting two distances of travel, say 2° and 4°, and two stimulus durations, say 100 msec and 200 msec. Discriminations between two pairs of stimuli

(4°/100 msec vs 2°/200 msec and 2°/100 msec vs 4°/200 msec) involve identical distance and duration cues. One of the pairs (4°/100 msec vs 2°/200 msec) exhibits a maximal velocity difference, whereas the latter pair (2°/100 msec vs 4°/200 msec) has no velocity difference. If sound source velocity is a distinct cue, then performance on the latter discrimination should be substantially worse than on the former.

DYNAMIC CUES FOR LOCALIZATION

Sustained sounds are localized best when the head is turned to face the direction of the source. Head turning occurs spontaneously on the part of practically all subjects seeking to determine the locus of the sound source (Thurlow & Runge 1967). Such scanning is also effective with monaural listening (Perrott et al 1987). Less clear is the contribution of head motion to the localization of brief sounds that permit only a slight change in head orientation. Is the information gained from the two or more head positions integrated, and does it greatly improve the process of disambiguating the source's location? Two views exist.

The earlier view was nicely summarized by Wallach (1939, 1940) nearly five decades ago. He begins the description of his classic experiments by observing that a given interaural time or intensity difference can originate from "a geometrical locus which has the shape of a cone" (Wallach 1939). Wallach is very clear on how this ambiguous information is resolved. "One obtains the cues for a number of lateral angles for the same sound direction by turning one's head while the sound is being given. Geometrically, a sequence of lateral angles obtained in this manner completely determines a given direction, . . . and the experiments to be reported indicate that the perception of sound direction actually works on this principle" (Wallach 1939).

Wallach's experiments demonstrated that certain confusions in localization can be created by moving the locus of the source with the head. For example, if the source is held in the subject's median plane as the subject turns his/her head, then the subject's only reasonable conclusion is that the source is directly overhead. In fact, Wallach found that of the subjects who could localize over their heads (an appreciable proportion of subjects cannot localize an overhead source), all reported the impression that the source was directly overhead (90° elevation) when the actual source was located straight ahead of the subject on the horizontal plane (0° elevation).

Wallach realized that the geometric assumptions are simplistic in nature. In particular, they ignore the fact that the pinnae are asymmetric and respond differently to sources presented in the front and back of the head. He observed that front/back ambiguities are generally resolved under ordinary conditions "on the basis of the pinna factor alone, i.e., without head movement. . . ." But he maintains that the systematic confusions generated in his experiments demonstrate that head motions are the primary means of disambiguating potential confusions. His experiments demonstrate that the "pinna factor" is invariably overcome by head motions, and this "indicates quite clearly its [the pinna's] subordinate role."

Experiments designed to test directly the importance of head motion have shown this cue to be remarkably weak. Thurlow & Runge (1967) studied the effect of induced (nonvoluntary) head motions. Although in some conditions the errors in localization were smaller by a statistically significant amount, the reduction in error was often less than 30%. Fisher & Freedman (1968) show no significant influence of head motion in localizing eight loudspeakers located at 45° angles about the listener on the horizon. Modifying the pinna had a more profound influence on the errors in localization than did restricting head motion.

Pollack & Rose (1967) carried out a series of five studies to investigate the role of head motion in the localization of sounds presented on the equatorial plane. Only one condition demonstrated a clear improvement in localization with head motion—when the sound was initially to the side and the stimulus was present until the head was turned to face the source. One of the more interesting results was from a condition in which the duration of the sound was systematically varied. Localization accuracy improved slightly if we compare conditions where head motion was allowed or restricted. The average error was 10–15% less with head motion. But when no head motion was allowed, changing the duration from 0.03 sec to 1 sec reduced the average error from 10° to 2°, an improvement of 500%!

The exact status of head motion and its effects on localization are far from clear. In light of all the evidence, a defensible argument is that unless the sound duration is sufficient to allow the listener to turn to face the source, thereby obtaining the optimum static localization cues, moving one's head may indeed be a poor strategy for improving the accuracy of localizing short-duration sources. If a source does not allow such extensive search time, then there is essentially no evidence to suggest that the information gained from two head locations (and the information gained from two cones of confusion) is substantially better than the information gained from a single head position.

Blauert provides the most diplomatic, if somewhat tautological, means of resolving the conflicting evidence. He maintains that there is a hierarchy of cues, and head motion can be superordinate if such information is available. "If information obtained by means of head movements is evaluated, it overrides information derived from monaural signal characteristics" (Blauert 1983).

SIMULATING EXTERNAL SOURCES OVER HEADPHONES

If one really understood all the cues responsible for locating external sound sources, then one should be able to simulate that same set of cues using headphones. This simple, but surprisingly daunting, challenge is the essential premise of the research program initiated by Wightman & Kistler at the University of Wisconsin. In their papers (Wightman & Kistler 1989a,b; Wightman et al 1987), they outline a program of research to compare sound localization in the free field and under headphone listening. The first step in this process is to measure directional transfer functions for multiple sound source locations around the listener. Six loudspeakers face the listener on a circular arc that is rotated about the subject in 15° steps of azimuth. Speakers are spaced from −36° to +54° elevation, in 18° steps. A broadband stimulus (200–14,000 Hz) is presented from each of 144 different sound locations, and recordings are made with miniature microphones located at fixed positions near both tympanic membranes. Dividing the recorded directional spectra by the source spectrum produces the directional transfer function of the ear. Directional spectra can then be reproduced over headphones, after compensating for the transfer functions of the headphones located on the subject's head. As they show in their first article (Wightman & Kistler 1989a), these simulated spectra, measured at the tympanic membrane, are the same as those produced by the free field source to within a few decibels in magnitude and a few degrees in phase over the frequency range measured. The measurements also tend to agree, on average, with previous measurements made by Shaw (1974) and by Mehrgardt & Mellert (1977) under similar circumstances.

Given the catalog of directional transfer functions established for each listener, Wightman & Kistler next compared the localization judgments given by the same subject, in free-field and under headphone listening. The subject indicated the apparent locus of the loudspeaker by simply calling out its location in degrees of elevation and azimuth. The general conclusion is that "The data . . . from the headphone condition are nearly identical to the data from the free-field condition" (Wightman & Kistler 1989b). Thus, to within the error of measurement associated with these experiments, the transfer functions contain all the information needed to simulate the localization of sound sources in the free field.

The enormous power of this technique is that the transfer functions can now be modified in various ways to determine the relative importance of different cue classes. For example, the level or phase cues can be set to point in one direction, while allowing the other cues to retain their natural values, thus placing different cues in opposition. How will localization change in such altered situations? So far the results are still somewhat incomplete and have only been presented at professional meetings (Wightman et al 1989). One consistent finding is that in situations in which interaural time cues point to a lateral location and interaural level cues point straight ahead, the time cues appear to dominate the localization judgment. The priority given to the time cues is mildly surprising, since duplex theory would say that time cues are effective in this experiment only from 200 Hz to about 1000 Hz (i.e. 2.3 octaves), whereas interaural level cues are available from about 2000 Hz up to 14,000 Hz (2.8 octaves).

In addition to their theoretical work, Wightman & Kistler are also involved in investigating whether this basic technique can be used to produce a practical auditory display device. The object of this system is to provide an acoustic signal under headphone listening that maintains an apparently stable location even if the head is moved. The system, being developed in collaboration with Wenzel and Foster, uses a head position indicator (ISOTRAK) to determine the roll, pitch, and yaw angles of the listener's head. The head position information is used to determine which transfer function should be used in the headphone presentation to maintain an apparently stable source location. A prototype of the system has been built and was demonstrated at the annual meeting (1988) of the National Research Council's Committee on Bioacoustics and Biomechanics (CHABA) and at the fall, 1988 meeting of the Acoustical Society of America (Wenzel et al 1988). A similar device has been shown to materially reduce errors associated with judgments of source azimuth in a simulation of an operational setting (Sorkin et al 1989).

SUMMARY

In keeping with our promise earlier in this review, we summarize here the process by which we believe spatial cues are used for localizing a sound source in a free-field listening situation. We believe it entails two parallel processes:

1. The azimuth of the source is determined using differences in interaural time or interaural intensity, whichever is present. Wightman and colleagues (1989) believe the low-frequency temporal information is dominant if both are present.
2. The elevation of the source is determined from spectral shape cues. The received sound spectrum, as modified by the pinna, is in effect compared with a stored set of directional transfer functions. These are actually the spectra of a nearly flat source heard at various elevations. The elevation that corresponds to the best-matching transfer function is selected as the locus of the sound. Pinnae are similar enough between people that certain

general rules (e.g. Blauert's boosted bands or Butler's covert peaks) can describe this process.

Head motion is probably not a critical part of the localization process, except in cases where time permits a very detailed assessment of location, in which case one tries to localize the source by turning the head toward the putative location. Sound localization is only moderately more precise when the listener points directly toward the source. The process is not analogous to localizing a visual source on the fovea of the retina. Thus, head motion provides only a moderate increase in localization accuracy.

Finally, current evidence does not support the view that auditory motion perception is anything more than detection of changes in static location over time.

ACKNOWLEDGMENTS

We wish to acknowledge the support of the National Institutes of Health, the Office of Naval Research, and the A. P. Sloan Foundation.

Literature Cited

Amenta, C. A. III, Trahiotis, C., Bernstein, L. R., Nuetzel, J. M. 1987. Some physical and psychological effects produced by selective delays of the envelope of narrow bands of noise. *Hearing Res.* 29:147–61

Angell, J. R., Fite, W. 1901. The monaural localization of sound. *Psychol. Rev.* 8:225–46

Batteau, D. W. 1967. The role of the pinna in human localization, *Proc. R. Soc. London Ser. B* 168:158–80

Batteau, D. W. 1968. Listening with the naked ear. In *The Neuropsychology of Spatially Oriented Behavior,* ed. S. J. Freedman, pp. 109–33. Homewood, IL: Dorsey Press

Belendiuk, K., Butler, R. A. 1975. Monaural location of low-pass noise bands in the horizontal plane. *J. Acoust. Soc. Am.* 58:701–5

Bernstein, L. R., Trahiotis, C. 1985. Lateralization of sinusoidally amplitude-modulated tones: effects of spectral locus and temporal variation. *J. Acoust. Soc. Am.* 78:514–23

Blauert, J. 1969/1970. Sound localization in the median plane. *Acustica* 22:205–13

Blauert, J. 1974. *Raumliches Hören.* Stuttgart: S. Hirzel

Blauert, J. 1982. Binaural localization: multiple images and application in room- and electroacoustics. In *Localization of Sound: Theory and Application,* ed. R. W. Gatehouse. Groton, CT: Amphora

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

Bloom, P. J. 1977a. Creating source elevation illusions by spectral manipulation. *J. Audio Eng. Soc.* 25:560–65

Bloom, P. J. 1977b. Determination of monaural sensitivity changes due to the pinna by use of minimum-audible-field measurements in the lateral plane. *J. Acoust. Soc. Am.* 61:820–28

Boudreau, J. C., Tsuchitani, C. 1968. Binaural interaction in the cat superior olive S segment. *J. Neurophysiol.* 31:442–54

Burger, J. F. 1958. Front-back discrimination of the hearing system. *Acustica* 8: 301–2

Butler, R. A. 1969. Monaural and binaural localization of noise bursts vertically in the median sagittal plane. *J. Audit. Res.* 9:230–35

Butler, R. A. 1986. The bandwidth effect on monaural and binaural localization. *Hearing Res.* 21:67–73

Butler, R. A. 1987. An analysis of the monaural displacement of sound in space. *Percept. Psychophys.* 41:1–7

Butler, R. A., Flannery, R. 1980. The spatial attributes of stimulus frequency and their role in monaural localization of sound in the horizontal plane. *Percept. Psychophys.* 28:449–57

Butler, R. A., Helwig, C. C. 1983. The spatial attributes of stimulus frequency in the median sagittal plane and their role in sound localization. *Am. J. Otolaryngol.* 4:165–73

Butler, R. A., Levy, E. T., Neff, W. D. 1980. Apparent distance of sounds recorded in echoic and anechoic chambers. *J. Exp. Psychol.: Hum. Percept. Perform.* 6:745–50

Butler, R. A., Planert, N. 1976. The influence of stimulus bandwidth on localization of sound in space. *Percept. Psychol:.* 19:103–8

Clifton, R. K. 1987. Breakdown of echo suppression in the precedence effect. *J. Acoust. Soc. Am.* 82:1834–35

Coleman, P. D. 1963. An analysis of cues to auditory depth perception in free space. *Psychol. Bull.* 60:302–15

Fisher, H. G., Freedman, S. J. 1968. The role of the pinna in auditory localization. *J. Audit. Res.* 8:15–26

Gardner, M. B. 1969. Distance estimation of 0° or apparent 0°-oriented speech signals in anechoic space. *J. Acoust. Soc. Am.* 45:47–53

Gardner, M. B. 1973. Some monaural and binaural facets of median plane localization. *J. Acoust. Soc. Am.* 54:1489–95

Gardner, M. B., Gardner, R. S. 1973. Problem of localization in the medial plane: effect of pinnae cavity occlusion. *J. Acoust. Soc. Am.* 53:400–08

Gaskell, H. 1983. The precedence effect. *Hearing Res.* 11:277–303

Goldberg, J. M., Brown, P. B. 1969. Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *J. Neurophysiol.* 32:613–36

Grantham, D. W. 1986. Detection and discrimination of simulated motion of auditory targets in the horizontal plane. *J. Acoust. Soc. Am.* 79:1939–49

Grantham, D. W. 1989. Motion aftereffects with horizontally moving sound sources in the free field. *Percept. Psychophys.* 45:129–36

Grantham, D. W., Wightman, F. L. 1979. Auditory motion aftereffects. *Percept. Psychophys.* 26:403–8

Hafter, E. R., Dye, R. H. 1983. Detection of interaural differences of time in trains of high-frequency clicks as a function of interclick interval and number. *J. Acoust. Soc. Am.* 73:644–51

Hafter, E. R., Dye, R. H., Nuetzel, J. M., Aronow, H. 1977. Difference threshold for interaural intensity. *J. Acoust. Soc. Am.* 61:829–33

Harris, J. D., Sergeant, R. L. 1971. Monaural/binaural minimum audible angles for a moving sound source. *J. Speech and Hearing Res.* 14:618–29

Hartmann, W. M. 1983. Localization of sound in rooms. *J. Acoust. Soc. Am.* 74:1380–91

Hebrank, J., Wright, D. 1974a. Are two ears necessary for localization of sound sources on the median plane? *J. Acoust. Soc. Am.* 56:935–38

Hebrank, J., Wright, D. 1974b. Spectral cues used in the localization of sound sources on the median plane. *J. Acoust. Soc. Am.* 56:1829–34

Henning, G. B. 1980. Some observations on the lateralization of complex waveforms. *J. Acoust. Soc. Am.* 68:446-54

Holt, R. E., Thurlow, W. R. 1969. Subject orientation and judgement of distance of a sound source. *J. Acoust. Soc. Am.* 46:1584–85

Hubel, D. H., Wiesel, T. N. 1962. Receptive fields, binocular interactions and functional architecture in the cat's visual cortex. *J. Physiol.* 160:106–54

Humanski, R. A., Butler, R. A. 1988. The contribution of the near and far ear toward localization of sound in the sagittal plane. *J. Acoust. Soc. Am.* 83:2300–10

Ivarsson, C., De Ribaupierre, Y., De Ribaupierre, F. 1980. Functional ear asymmetry in vertical localization. *Hearing Res.* 3:241–47

Klumpp, R. G., Eady, H. R. 1956. Some measurements of interaural time difference thresholds. *J. Acoust. Soc. Am.* 28:859–60

Knudsen, E. I. 1982. Auditory and visual maps of space in the optic tectum of the owl. *J. Neurosci.* 2:1177–94

Knudsen, E. I. 1984. Synthesis of a neural map of auditory space in the owl. In *Dynamic Aspects of Neocortical Function*, ed. G. M. Edelman, W. E. Gall, W. M. Cowan, pp. 375–96. New York: John Wiley

Kuhn, G. F. 1977. Model for interaural time differences in the azimuthal plane. *J. Acoust. Soc. Am.* 62:157–67

Kuhn, G. F. 1987. Physical acoustics and measurements pertaining to directional hearing. See Yost & Gourevitch 1987, pp. 3–25

Lappin, J. S., Bell, H. H., Harm, O. J., Kottas, B. 1975. On the relation between time and space in the visual discrimination of velocity. *J. Exp. Psychol.: Hum. Percept. Perform.* 1:383–94

Licklider, J. C. R., Webster, J. C., Hedlum, J. M. 1950. On the frequency limits of binaural beats. *J. Acoust. Soc. Am.* 22:468–73

Makous, J. C., Middlebrooks, J. C. 1990. Two-dimensional sound localization by human listeners. *J. Acoust. Soc. Am.* 87:2188–2200

Masterton, R. B., Imig, T. J. 1984. Neural mechanisms for sound localization. *Annu. Rev. Physiol.* 46:275–87

McFadden, D., Moffitt, C. M. 1977. Acoustic integration for lateralization at high frequencies. *J. Acoust. Soc. Am.* 61:1604–8

McFadden, D., Pasanen, E. G. 1976. Lateralization at high frequencies based on interaural time differences. *J. Acoust. Soc. Am.* 59:634–39

Mehrgardt, S., Mellert, V. 1977. Transformation characteristics of the external human ear. *J. Acoust. Soc. Am.* 61:1567–76

Mershon, D. H., Bowers, J. N. 1979. Absolute and relative cues for the auditory perception of egocentric distance. *Perception* 8:311–22

Mershon, D. H, King, L. E. 1975. Intensity and reverberation as factors in the auditory perception of egocentric distance. *Percept. Psychophys.* 18:409–15

Middlebrooks, J. C. 1990. Two-dimensional localization of narrowband sound sources. *Assoc. Res. Otolaryngol.* 13:109 (Abstr.)

Middlebrooks, J. C., Green, D. M. 1990. Directional dependence of interaural envelope delays. *J. Acoust. Soc. Am.* 87:2149–62

Middlebrooks, J. C., Makous, J. C., Green, D. M. 1989. Directional sensitivity of sound-pressure levels in the human ear canal. *J. Acoust. Soc. Am.* 86:89–108

Mills, A. W. 1958. On the minimum audible angle. *J. Acoust. Soc. Am.* 30:237–46

Mills, A. W. 1960. Lateralization of high-frequency tones. *J. Acoust. Soc. Am.* 32:132–34

Moiseff, A., Konishi, M. 1981. Neuronal and behavioral sensitivity to binaural time differences in the owl. *J. Neurophysiol.* 1:40–48

Musicant, A. D., Butler, R. A. 1984a. The influence of pinnae-based spectral cues on sound localization. *J. Acoust. Soc. Am.* 75:1195–1200

Musicant, A. D., Butler, R. A. 1984b. The psychophysical basis of monaural localization. *Hearing Res.* 14:185–90

Musicant, A. D., Butler, R. A. 1985. Influence of monaural spectral cues on binaural localization. *J. Acoust. Soc. Am.* 77:202–8

Newsome, W. T., Mikami, A., Wurtz, R. H. 1986. Motion selectivity in macaque visual cortex. III. Psychophysics and physiology of apparent motion. *J. Neurophysiol.* 55:1340–51

Nuetzel, J. M., Hafter, E. R. 1981. Lateralization of complex wave-forms: spectral effects. *J. Acoust. Soc. Am.* 69:1112–18

Oldfield, S. R., Parker, S. P. A. 1984a. Acuity of sound localisation: a topography of auditory space. I. Normal hearing conditions. *Perception* 13:581–600

Oldfield, S. R., Parker, S. P. A. 1984b. Acuity of sound localisation: a topography of auditory space. II. Pinna cues absent. *Perception* 13:601–17

Oldfield, S. R., Parker, S. P. A. 1986. Acuity of sound localisation: a topography of auditory space. III. Monaural hearing conditions. *Perception* 15:67–81

Pantle, A. J., Sekuler, R. W. 1968. Velocity-sensitive elements in human vision: initial psychophysical evidence. *Vis. Res.* 8:445–50

Perrott, D. R., 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.* 82:1637–45

Perrott, D. R., Marlborough, K. 1989. Minimum audible movement angle: marking the end points of the path traveled by a moving sound source. *J. Acoust. Soc. Am.* 85:1773–75

Perrott, D. R., Musicant, A. D. 1977. Minimum auditory movement angle: binaural localization of moving sound sources. *J. Acoust. Soc. Am.* 62:1463–66

Perrott, D. R., Tucker, J. 1988. Minimum audible movement angle as a function of signal frequency and the velocity of the source. *J. Acoust. Soc. Am.* 83:1522–27

Phillips, D. P. 1985. Progress in neurophysiology of sound localization. *Annu. Rev. Psychol.* 36:245–74

Pollack, I., Rose, M. 1967. Effect of head movement on the localization of sounds in the equatorial plane. *Percept. Psychophys.* 2:591–96

Rayleigh, Lord. 1907. On our perception of sound direction. *Philos. Mag.* 13:214–32

Roffler, S. K., Butler, R. A. 1968. Factors that influence the localization of sound in the vertical plane. *J. Acoust. Soc. Am.* 43:1255–59

Rose, J. E., Brugge, J. F., Anderson, D. J., Hind, J. E. 1967. Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *J. Neurophysiol.* 30:769–93

Rose, J. E., Gross, N. B., Geisler, C. D., Hind, J. E. 1966. Some neural mechanism in the inferior colliculus of the cat which may be relevant to localization of a sound source. *J. Neurophysiol.* 29:288–314

Saberi, K., Perrott, D. R. 1990. Minimum audible movement angle thresholds as a function of sound source trajectory. *J. Acoust. Soc. Am.* In press

Sandel, T. T., Teas, D. C., Feddersen, W. E., Jeffress, L. A. 1955. Localization of sound from single and paired sources. *J. Acoust. Soc. Am.* 27:842–52

Searle, C. L., Braida, L. D., Cuddy, D. R., Davis, M. F. 1975. Binaural pinna disparity: another auditory localization cue. *J. Acoust. Soc. Am.* 57:448–55

Searle, C. L., Braida, L. D., Davis, M. F., Colburn, H. S. 1976. Model for auditory localization. *J. Acoust. Soc. Am.* 60:1164–75

Sekuler, R. W., Pantle, A. J. 1967. A model for after effects of seen movement. *Vis. Res.* 88:1–11

Shaw, E. A. G. 1974. Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. *J. Acoust. Soc. Am.* 56:1848–61

Simpson, W. E., Stanton, L. D. 1973. Head movement does not facilitate perception of the distance of a source of sound. *Am. J. Psychol.* 86:151–60

Sorkin, R. D., Wightman, F. L., Kistler, D. S., Elvers, G. C. 1989. An exploratory study of the use of movement-correlated cues in an auditory head-up display. *Hum. Factors* 31:161–66

Stevens, S. S., Newman, E. B. 1936. The localization of actual sources of sound. *Am. J. Psychol.* 48:297–306

Strybel, T. Z., Perrott, D. R. 1984. Discrimination of relative distance in the auditory modality: the success and failure of the loudness discrimination hypothesis. *J. Acoust. Soc. Am.* 76:318–20

Thurlow, W. R., Runge, P. S. 1967. Effect of induced head movements on localization of direction of sounds. *J. Acoust. Soc. Am.* 42:480–88

Trahiotis, C., Bernstein, L. R. 1986. Lateralization of bands of noise and sinusoidally amplitude-modulated tones: effects of spectral locus and bandwidth. *J. Acoust. Soc. Am.* 79:1950–57

Wallach, H. 1939. On sound localization. *J. Acoust. Soc. Am.* 10:270–74

Wallach, H. 1940. The role of head movements and vestibular and visual cues in sound localization. *J. Exp. Psychol.* 27:339–68

Watkins, A. J. 1978. Psychoacoustical aspects of synthesized vertical locale cues. *J. Acoust. Soc. Am.* 63:1152–65

Wightman, F. L., Kistler, D. J. 1989a. Headphones simulation of free-field listening. I. Stimulus synthesis. *J. Acoust. Soc. Am.* 85:858–67

Wightman, F. L., Kistler, D. J. 1989b. Headphone simulation of free-field listening. II. Psychophysical validation. *J. Acoust. Soc. Am.* 85:868–78

Wightman, F. L., Kistler, D., Arruda, M. 1989. The hierarchy of sound localization cues revealed by experimcnts in a simulated free-field. *Assoc. Res. Otolaryngol.* 12:65 (Abstr.)

Wightman, F. L., Kistler, D. J., Perkins, M. E. 1987. A new approach to the study of human sound localization. See Yost & Gourcvitch 1987, pp. 26–48

Woodworth, R. S. 1938. *Experimental psychology.* New York: Holt

Yost, W. A., Gourevitch, G., eds. 1987. *Directional Hearing.* New York: Springer-Verlag

Yost, W. A., Wightman, F. L., Green, D. M. 1971. Lateralization of filtered clicks. *J. Acoust. Soc. Am.* 50:1526-31

Zurek, P. M. 1980. The precedence effect and its possible role in the avoidance of interaural ambiguities. *J. Acoust. Soc. Am.* 67:952–64

Zwislocki, J., Feldman, R. S. 1956. Just noticeable differences in dichotic phase. *J. Acoust. Soc. Am.* 28:860–64

Wenzel, E., Wightman, F., Kistler, D., Foster, S. 1988. Acoustic origins of individual differences in sound localization behavior. *J. Acoust. Soc. Am.* 84 (Suppl 1): S79